12-1-2014

Metadata determination for a cadaveric collection

Shamsi Berry

Shamsi Rebecca Berry

Candidate

Biomedical Sciences

Department

This thesis is approved, and it is acceptable in quality and form for publication:

Approved by the Thesis Committee:

Philip J. Kroth, MD, MS , Chairperson

Heather J. H. Edgar, PhD

Gary Hatch, MD

Teddy Warner, PhD

METADATA DETERMINATION FOR A CADAVERIC COLLECTION

by

SHAMSI REBECCA BERRY

BACHELOR OF SCIENCE, UNIVERSITY OF CALIFORNIA, DAVIS, 1998
MASTER OF SCIENCE, UNIVERSITY OF NEW MEXICO, 2003
DOCTORATE OF PHILOSOPHY, UNIVERSITY OF NEW MEXICO, 2011

THESIS

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Master of Science in Biomedical Sciences
Concentration: Clinical Research

The University of New Mexico
Albuquerque, New Mexico

December 2014

Dedication

To my husband, Phillip Berry, thank you for all of your love and support.

Acknowledgements

I would like to thank my advisor, Dr. Philip Kroth, for his support and guidance throughout the MSCR program and my Biomedical Informatics Fellowship. Additionally, I thank the other members of my thesis committee, Drs. Heather Edgar, Gary Hatch and Teddy Warner, for their invaluable assistance and advice.

I also thank all of my family and friends for their support and love.

Metadata Determination for a Cadaveric Collection

By

Shamsi Rebecca Berry

BS, Anthropology, University of California, Davis, 1998

MS, Anthropology, University of New Mexico, 2003

PhD, Anthropology, University of New Mexico, 2011

MS, Biomedical Sciences, University of New Mexico, 2014

Abstract

The number, size and complexity of research databases continue to grow unabated. Yet, despite significant advances in information technology, scientists continue to struggle with "data wrangling" or issues of finding, sharing, and reusing data, often for unanticipated future purposes. Most research databases are designed ad hoc, by native investigators for specific research functions and immediate needs with little to no input from database designers, informaticians, or other subject matter experts. Little thought is usually given to the implications of future data retrieval, sharing and reuse. When the needs of future investigators or research requirements change, the original design of the database can become a significant barrier to meeting unanticipated needs and can impede future scientific discovery. Designing research databases to anticipate future needs is a significant challenge given there is no

universally acknowledged standard or guideline for researchers to follow when designing research databases.

The New Mexico Office of the Medical Investigator (OMI) received a grant from the Department of Justice in 2010. The OMI's research sought to determine if Computed Tomography (CT) scans could supplement or supplant traditional autopsies. A by-product of this research was over 6,000 full-body, three dimensional, high resolution scans on every decedent that underwent a traditional autopsy. There were no plans to reuse this treasure trove of scans and associated health information.

A Modified Delphi Method was used to create a Minimum Data Set for a research database of full-body, three dimensional cadaveric images. A Snowball Sampling Method was also performed to evaluate the quality of the metadata produced by the Delphi expert group.

Fifty-nine metadata variables were recommended for inclusion in the Minimum Data Set, which only included 44% of the original ad hoc variables. As a result the Minimum Data Set is thought to be applicable and relevant to more research domains and studies than the original set of metadata variables selected by the native database designers. The Snowball Validation Method verified the 59 variables selected by the Delphi expert group and suggested 3 additional fields not included in the Delphi set.

Using a larger group of experts produced 56% more metadata variables than the database designers had created ad hoc. This suggests that a modified Delphi Method that queries a broad domain of experts beyond what is typically done for immediate needs is superior. The Snowball Validation Methods can also work well to check the

validity of the Delphi design process. These methods can produce a Minimum Data Set of metadata variables that is more “future-proof” than those typically created by local, native investigators alone.

Table of Contents

Chapter 1: Introduction

As data increases and technology evolves the ability to organize and manage information becomes of greater importance. In a large number of cases data is stored in some type of database rather than separate files. Databases are an efficient and effective method for storage, retrieval of information and potential reuse. There are many reasons why databases are created, which depend on the goal of the research or project, including: documentation or administrative purposes, organization of information, and to enable research. The reason(s) for creating a database influences the information captured, how it is recorded and who has access. For most research databases, the goal is to gather information regarding a specific scientific query, investigation or task. As such, the majority of selected database fields relates to the specific project and is not always applicable to other investigators.

The searchable fields present in a database are the metadata of the database, or the data about data. The quality of the set of metadata is inextricably linked to the quality and ease of data retrieval and therefore the value of the data in the future. The process of determining optimal metadata sets is problematic since it is difficult to predict all potential future uses of a database and therefore the form and number of metadata fields to include in the database's design. How can a database designer create metadata in a way that optimizes the value of the data not only for immediate use but also future, unspecified, uses? If allowed, researchers would request all possible data variables to be present in a database. In reality, resource constraints often limit the sophistication of metadata design. Choosing the "wrong" metadata fields or selecting

too few can significantly reduce the future value of the data over time. Likewise, defining too many fields utilizes valuable resources inefficiently and can reach a point of diminishing returns. In effect it is similar to pre-coordinating or post-coordinating a terminology standard. Pre-coordination is when database fields are pre-combined into one variable in anticipation of searching on that compound heading. Post-coordination, on the other hand, stores all data in discrete database fields allowing a researcher to combine them in any manner. There are benefits and limitations to both approaches. Pre-coordination involves the possibility of being too detailed and creating a plethora of database fields combining all of the possible variables (known as combinatorial explosion). Post- coordination can also be a negative method to use as it puts a burden on the user to create the necessary combination of database fields which can become fairly complex and require a great deal of knowledge of the data domain. There must be a balance between combinatorial explosion and limiting retrieval by under designing.[1, 2] Likewise, the challenge in database design is to define the smallest metadata set that will produce the most value in the future. Therefore, the process of designing the selected variables and optimal minimal dataset is critical to maximize the value of research data over time.

Chapter 2: Background

Metadata

Metadata is the structured information that supports additional functions or actions about an object, topic or person.[3] Quality metadata enables the user to retrieve information with high sensitivity and specificity whereas poor quality metadata risks missing relevant items in a database.[4] The use of appropriate and high quality metadata, therefore, facilitates information retrieval, efficient searching, maintenance, understanding, interoperability and reuse.[5, 6]

In the current technologically advanced and expanding information environment there are numerous new opportunities for data collection and reuse for research purposes. For example, medical images are being created in vast quantities every day in hospitals, doctor offices, imaging facilities, and coroner/medical examiner offices. As the number of images being created and stored continues to grow, a few facilities are incorporating plans for selected images to be reused by investigators and educators. Additionally, there is an expansion in image modalities and quality, which increase the number and size of the images created.[7] In 2009, 227.9 CT scans and 91.2 MRI scans were performed per 1,000 people, with greater than 50% occurring in hospitals.[8] In the United States with a population of 305 million, an estimated 65 million CT and 28 million MRI scans were performed in 2009. At present, the majority of these images are stored in PACS (Picture Archiving and Communication System), encoded using DICOM standards, and the Abbreviated Injury Scale but without detailed health and lifestyle

related metadata.[7, 8] Although medical images are being created every day for diagnostic and therapeutic purposes, very few research image databases exist from routine or exploratory procedures. Often these stored images consist of only one organ or region of the body, with no associated information.[9]

Metadata Selection

The selection of metadata, both the number and content of the fields, will impact the effectiveness of retrieval. As a result, it is vital to select the appropriate metadata for a research database. Malaxa and Douglas[6] stress that the selection of metadata is a trade-off between discoverability and cost, where the greater the number of metadata fields, the greater the chance of discovering the relevant resource but the higher the cost. In addition, it is possible to plateau on the usefulness of the metadata fields. (See Figure 1). Metadata, therefore, needs to be accurate, complete and cost-effective in order to make it most useful to future investigators. In addition, there must be some flexibility present in the design so that new metadata fields can be added, as necessary, making the database adaptable to unanticipated circumstances or "future-proof."[6]

Within the realm of medical images, information can be indexed in two possible ways: semantic (text-based and visual) or content-based. Currently techniques that utilize content based indexing only cover shape, texture, color, segmentation, distance or angles of the image itself.[10] This information is usually derived by software tools that scan the image and record detectable image characteristics in a standardized manner.

These techniques can be used in databases where the queries involve searching on those specific variables, but do not extend to uses such as determining information not present in the image. Also, in cases where there is limited content or where the creativity and intelligence of a human is required, semantic or text-based indexing is the preferred method for annotating an image.[4] In these cases, metadata can be standardized and applied consistently to each image to improve the retrieval, storage and processing.

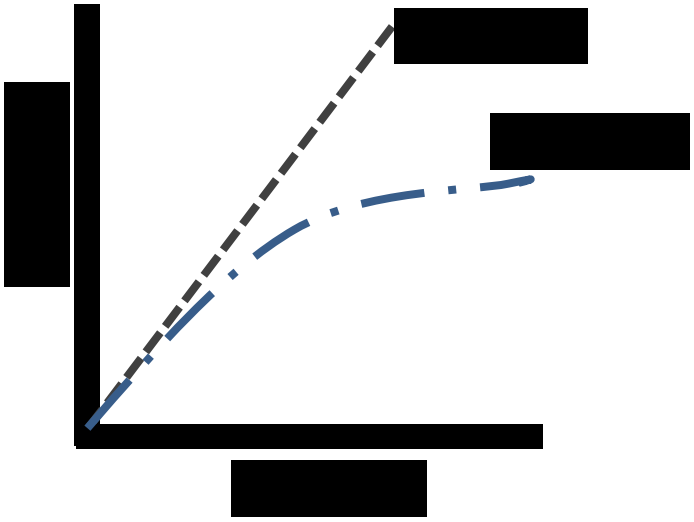

Figure 1: The marginal value of additional metadata used for the initial design of a minimum dataset asymptotically approaches a limit

The type of information associated with an image can include content-independent, content-dependent and content-descriptive metadata. Content independent metadata refers to data that is not directly concerned with the image, but is related. For example, the file format and date are content-dependent variables. In contrast, content dependent metadata refers to low-level features, such as color, shape, and texture. Lastly content descriptive metadata refers to the relationship between the

image and the real-world entity. Content descriptive metadata therefore includes variables that describe what is seen in the image itself, but cannot be automatically detected. For example, the tumor type, and age of an individual cannot be determined from scanning the image, but it is descriptive of the image. The selection of these variables will impact the effectiveness of retrieval and therefore it is vital to select the appropriate types of metadata to associate with images.[10]

Minimum Data Set Creation

Individual metadata elements can be combined to form a set of data for an image or object. The Health Information Policy Council[11] defined a Uniform Minimum Health Data Set as a "set of items [or elements] of information with uniform definitions and categories, concerning a specific aspect or dimension of the health care system, which meets the essential needs of multiple data users." This concept has been applied to multiple disciplines as a Minimum Data Set (MDS), which allows for interoperability of data between investigators.[12]

MDSs have been created for a multitude of databases, especially within health care. The Nursing MDS aims to establish and standardize the important and necessary data recorded by nurses. Developers sought to allow comparison of nursing data across multiple situations, locations, describe care received, project workflow allocations and stimulate research in health care domains.[13] Other major disciplines using a MDS to standardize retrieval of vital information include: genetics,[14] nursing homes,[15] and rare and orphaned diseases.[16, 17] The metadata for each database depends on the intent of

its creation. For example, within the MDS for rare and orphaned diseases, common data elements (metadata fields) include demographic data, contact information, diagnoses, family history, reproductive history, anthropometric data, patient outcomes, biospecimens data, and communication preferences.[18, 19] Each MDS serves to facilitate the goals of the creators and in some cases, assist with reuse of the data. Unfortunately no standardized MDS has been devised allowing for reuse of images to conduct health, anthropology, dental, informatics, demographic and forensic research.

Given the importance of the metadata fields (both singularly and as a MDS) to a database's usefulness it is surprising that the process for metadata creation has not been standardized. There has, however, been work by the IEEE (Institute of Electrical and Electronics Engineers) creating the IEEE Learning Object Metadata standard (LOM).[20] This work addresses the structure of the metadata (minimal set of attributes allowing interoperability and "findability") and not the content of those chosen data variables.

The selection of the metadata content can involve multiple approaches, including determination through: 1) the resource author (the individual requiring the database for their research), 2) a metadata specialist or 3) a collaborative procedure[4]. Research has shown that many resource authors lack the knowledge and skills of indexing and therefore generate insufficient metadata.[3] Furthermore, resource authors generally lack the awareness of the benefits in funding and investigating in a quality metadata design. This in turn diminishes the future discovery of pertinent records and

generates inadequate results. The same outcomes apply to using a metadata specialist as they may have little domain knowledge and have difficulty designing the metadata appropriate to likely future uses.[21] Greenberg and Robertson[21] suggested that the best quality metadata is obtained through a collaborative process. The exact model for the collaboration depends on the investigative team and their resources. Models include, but are not limited to, the Delphi method, in which there is no direct interaction, and the Nominal Group Technique, which includes a round-robin iterative discussion, all with participant numbers ranging up to over 100 experts.[22]

Metadata, once determined, can be associated with an image in two ways, both of which have benefits and drawbacks. The data can be imbedded in the image file itself or stored in a separate database that is linked to the actual image. The time and resources for embedding data into an image file is relatively costly and time consuming. However, the advantage of embedded metadata helps assure the metadata is always associated with the image. The benefit of a separate database with a link to the image is usually less costly and time consuming than embedding the information. The database method also has the benefit of allowing simple and complex queries to be performed and supports the de-identification of images. Lastly, the use of a database to record the metadata has the advantage of allowing faster retrospective upgrades since only the metadata is altered rather than the entire image files. The database can remain available to researchers while a copy is upgraded with new information or metadata field(s). Modifying embedded data would result in the images not being available for research or unpredictable results until all of the images had been upgraded. See Table 1

for the benefits and limitations of each method of associating metadata and images. The use of the database to house the MDS with a link to the images is usually financially beneficial, efficient, and more flexible in design allowing for linkage to de-identified images and more effectual metadata upgrades in the future.

Table 1-Advantages and Disadvantages to Methods of Associating Metadata and Images

Method	Advantages	Disadvantages
Data Imbedded in Image	• Information always associated with image	• High cost • Large amount of time to integrate data into image file • Difficult to retrospectively alter metadata • Metadata upgrade would involve disabling database until complete
Data Stored in database	• Lower cost • Less time required • Faster retrospective upgrades with database copy replacement	• Only linked to image through link

Assessment of a Minimum Data Set

Once a MDS has been determined there should also be an evaluation of its quality. Hillman[23] defines the seven domains that compromise metadata quality: completeness, accuracy, provenance, conformance to expectations, logical consistency and coherence, timeliness and accessibility. Previous evaluations of MDSs involve a multitude of assessments to analyze these attributes. However, evaluation of metadata is still not a standardized procedure.[23]

Dushay and Hillman[24] determined that the primary errors found in the selection of metadata fall into three categories: missing data, confusing data and insufficient data.

This assessment relates to the data within the fields and not necessarily the fields selected for the database, however, the ideas hold true for metadata fields as well. Assessments of a MDS must therefore consider whether chosen metadata fields will be complete, clear, and sufficient for the intended research uses. Goossen and colleagues[25] also identified five characteristics for a MDS: 1) relevant data must be identified, 2) the data must be defined accurately, 3) all possible values must be elucidated, 4) patient data (or other data) must be able to be documented using the chosen variables and the identified values, and 5) data must be able to be combined, coded and used for its intended purpose. Goossen and colleagues'[25] and Dushay and Hillman's[24] characteristics are complementary and both seek to identify metadata that serves the intended purpose accurately and completely. Laws and Sullivan[26], among others, have conducted a survey of the quality of the metadata associated in a database. The evaluations explore the utility of the metadata fields selected for inclusion in regards to its importance and usefulness in the database.

Chapter 3: Methods

Research Context

The New Mexico Office of the Medical Investigator (OMI) is a state-wide, centralized medical examiner's office for the entire state of New Mexico. With a few exceptions, any individual who dies in the state in a sudden, violent, untimely, or unexpected manner, and any person who is found dead and the cause of death is unknown, is routed to the OMI for a possible autopsy. In 2010, 5,249 deaths were processed by the OMI, which accounted for 35% of the total deaths within the state and closely mirrored the ethnic and racial composition of the state.[27, 28] See Table 2 for OMI's 2010 demographic data.[27] Many people assume that the decedents sent to the OMI for a post-mortem examination have all succumbed to a violent death. However, the vast majority of these cases are from natural causes (58%) and only 51% of the total cases sent to the OMI undergo autopsy. Of those autopsied, roughly 25% died from natural causes, 35% from accidents, 17% from suicides, 13.5% from unknown causes and 9.5 % from homicides.[27] As a result, the sample, although skewed, is more representative of the state's population than seems readily apparent from a medical examiner's "sample."

Table 2- OMI demographic data (2010)

Race	Percentage of OMI cases
European American	87.3%
Asian/Pacific Islander	0.8%
African American	2.4%
American Indian	8.7%
Ethnicity	
Hispanic	30.2%
Non-Hispanic	69.8%
Age	
0-19	6.3%
20-64	51.7%
65 +	42.0%

The Center for Forensic Imaging at the OMI was awarded a grant in 2010 from the National Institute of Justice to evaluate the utility of postmortem CT scans to supplant or supplement traditional autopsy. For this research endevor, every cadaver that undergoes an autopsy at the OMI receives a high resolution, head-to-toe Computed Tomography (CT) scan. This produced thousands of whole-body 3-D CT images – a treasure trove for a variety of research domains – but unfortunately, without the associated metadata to allow investigators to efficiently identify images of interest for specific research purposes. With the vast amount of data in health care, such as in the case with the OMI, there is a need for the curation of these data for both education and research.[12] The incorporation of a comprehensive annotation schema to this database would facilitate future research of the CT images and associated health and lifestyle information by a wide variety of investigators. This database will be a unique resource due to its size, metadata, high quality 3-D images, and diverse population.

Research Aims

The specific aims of this research are to:

1. Determine the Minimum Data Set (MDS) to associate with the CT scans in a database of 3-D, whole-body, human, cadaveric images developed at the New Mexico Office of the Medical Investigator (OMI). The MDS should enable investigators to efficiently and effectively search for images from the database that meet the inclusion criteria of his/her study.
2. Determine how well the MDS facilitates meaningful retrieval of relevant images within the database of 3-D cadaveric images for specific research purposes. The MDS should enable investigators to conduct a potential future investigation in their field.

Unfortunately no MDSs have been created for medical images to facilitate research in multiple domains. As a result, we created a MDS using consensus group methods most appropriate for creating and validating the optimal MDS.

Design

In order to accomplish the primary aim of creating a MDS, Greenberg and Robertson's[21] suggestion to use an iterative consensus method was followed to eliminate the biases produced from a database creator alone or metadata specialist without domain knowledge. In addition, an electronic consensus method was selected as it is beneficial as it does not require the expenses of in-person meetings and can therefore include participation from experts living virtually anywhere in the world. In

this instance, the Delphi method, developed by the RAND Corporation during the 1950's for defense research,[29] is the most appropriate method to undertake as it allows for electronic data collection. This method involves soliciting experts in certain domain(s) to obtain a convergence of opinions traditionally concerning forecasting, goal setting or policy setting.[30] The Delphi method is well established and has been applied to multiple subjects within a wide range of domains and more recently has been implemented electronically. The method allows for anonymous participation of experts through an iterative process. This process usually involves 2 to 5 iterative rounds of questionnaires in which individual participants are asked to provide their opinions and then review the information provided by the whole panel. This process provides an opportunity for the individual participant to re-evaluate their original feedback in light of others' responses and evaluate other's suggestions. The underlying principle of the Delphi method, like all consensus methods, is that the group generated solution is usually better than each of the experts would have been able to formulate independently.[31] Due to the varying nature of each consensus panel, the level of consensus is determined after each round of the Delphi method.[32]

Once a MDS is determined through a collaborative, iterative process it should be validated by additional experts not involved with its creation.[23, 25] Although surveys and questionnaires have been used to validate a MSD, a snowball sampling method was selected here to take advantage of the knowledge and connections of our experts. Experts involved in the Delphi selection process were asked to supply two to three names and emails of additional experts within their domain they thought would be well

qualified to validate the metadata fields. The validation process sought to evaluate the MDS using Hillman's[23] evaluation domains: completeness, accuracy, conformance to expectations, logical consistency and timeliness. The experts in the validation phase were asked if the database fields selected are adequate for future research and if there are additional fields that have been omitted by the original panel. This method does not take into account the majority of Goossen and colleagues'[25] five characteristics for a MDS as the process outlined here does not specify how the selected fields will be collected or coded with terminology standards. The method outlined here is simply for the determination of which database fields experts wish to use for research now and in the future. The next step of the project after the MDS design is completed will be to create the metadata database, choose and create terminology standards, and determine how data collection can be efficiently implemented.

Institutional Review Board approval was granted from the University of New Mexico Human Subjects Research Review Committee on June 10, 2013 to determine and validate a MDS for full-body, 3-D, cadaveric CT images housed at the New Mexico OMI (Human Research Protections Office #13-229). The Delphi method's (part 1- the Design Phase) goal was to recruit 100 individuals. Experts were recruited based on the research qualifications and expertise in a variety of scientific domains where they could foreseeably use the CT database of cadaveric images for a research project. For the validation portion of the project (part 2- the Validation Phase), experts were suggested by participants in part 1. The only inclusion criterion for Part 2 was that suggested

individuals needed to have research that was amenable to using the database. The methods used in part 1 and 2 are shown in Figure 2.

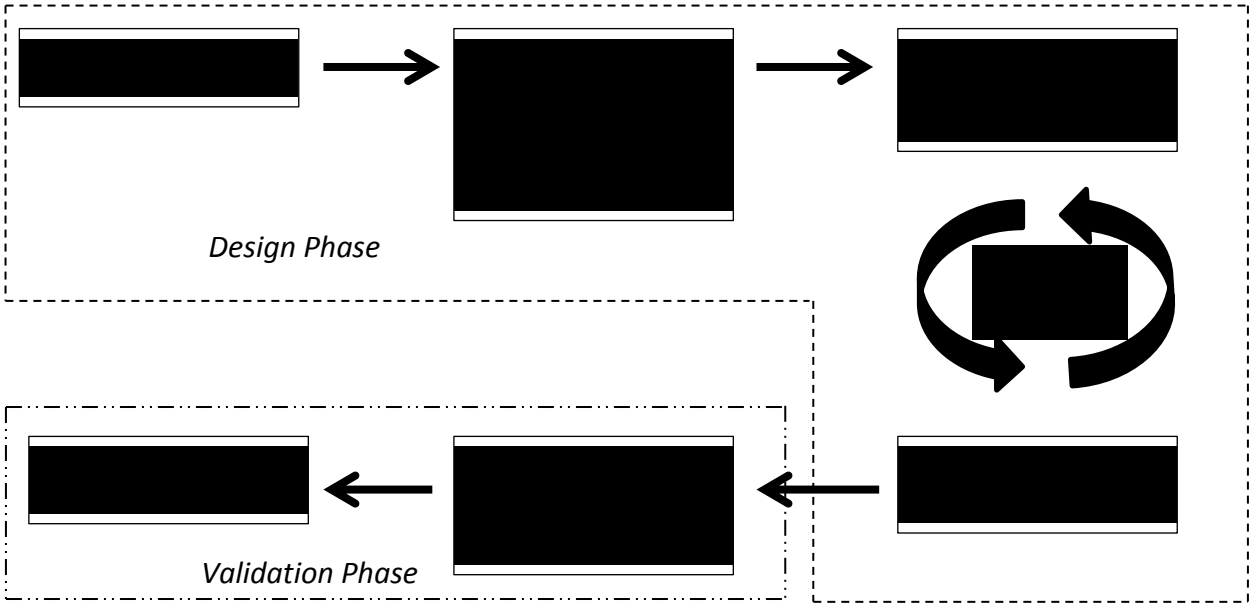

Figure 2: Method for determining and validating a Minimum Data Set

Expert Determination

For the Design Phase, a list of research domains was created by Drs. Shamsi Daneshvari, Philip J. Kroth, Heather H. J. Edgar, Gary M. Hatch and Teddy Warner. The group believed these selected domains would be amenable to using the scans and associated data for future research. Within each of the domains, participants were selected as experts through a detailed search of the literature, as suggested by Cuhls.[32] In addition, Drs. Daneshvari, Kroth, Edgar, Hatch and Warner suggested participants that were experts within their respective disciplines as well as individuals to contact for suggestions.

For the Validation Phase, the Design Phase participants will be asked to recommend 2 to 3 experts that they believe may use the full-body, 3-D, cadaveric CT scan and associated health information database. These individuals can be people they know personally or by professional reputation when an individual investigator recommends a member of the Design Phase, the person was excluded from the Validation Phase group.

Questionnaire Creation

A "straw man" draft preliminary questionnaire was created, with input from Drs. Daneshvari, Edgar and Kroth, using REDCap,[33] a secure web-based survey and data capture program for experts to participate in the Delphi process. By creating a straw man list of database fields, experts had a starting point in which they could react to at the beginning of the consensus process. Additionally, this reduced the burden of the experts by providing a list of what the questionnaire creators believed were basic fields that each participant would likely recommend as candidate metadata terms (e.g. age, gender). If these straw man terms were not wanted for inclusion in the MDS by the participants, the Delphi process eliminated them. The first questionnaire provided a basic set of metadata candidate variables in five categories: personal characteristics, lifestyle, health, occupation, and other.

Follow-up questionnaire(s) in the Design Phase allowed participants to revise the groups' and their own ideas. This process continued until the investigators believed saturation was reached (i.e. additional design iterations were unlikely to provide any further benefit). The final questionnaire within the Design Phase asked participants to

rate the suggested database fields in terms of importance of inclusion in the MDS (e.g., from 0 = not important at all to 10 = absolutely essential to include). Through this process a consensus was arrived at in regards to the minimum required database fields that constitute the MDS. The standard for consensus was determined by the investigators and informed by the data as the series of questionnaires proceed.

For both the Design and Validation Phases, a one-page recruitment letter was mailed first class to potential experts as well as sent electronically to their institution e-mail; that letter also had enclosed a one-page abbreviated consent form. See Appendix A. Because this project collects only non-sensitive data, a waiver for a signed documentation of informed consent was approved by the University of New Mexico IRB. Experts had the option of completing each survey round with a hardcopy or with an electronic questionnaire sent to experts 1-2 weeks after the initial recruitment letter.

The Validation Phase questionnaire also asked participants to rate the suggested database fields (e.g., from 0 = not important at all to 10 = absolutely essential to include). In addition, the experts from the Validation Phase were asked to provide any additional absolutely essential fields that the Design Phase participants did not identify. This questionnaire was also provided to participants in an electronic REDCap survey.

Chapter 4: Results

Part 1: Design Phase

A total of 72 experts were sent a letter and/or email requesting participation in this research. The research domains surveyed include: medicine, biological anthropology, forensics, radiology, biomedical informatics, dentistry, epidemiology, growth and development, demography, health disparities, secular change, population variation, pathology, odontology, health economy, nursing informatics and chronic pain. In total 42 participants completed the questionnaire (the experts self-identified from the domains listed in Table 3). Only 2 participants responded to the paper letter. The remaining 40 participants did not respond until they were emailed a reminder to participate. The letters were mailed at the end of September to coincide with the beginning of the fall school schedule. However, the majority of individuals asked to participate were extremely busy during the fall semester and therefore the questionnaire remained open until the end of November (3 months total). There was a 58.3% respondent rate from the experts solicited, which indicates a strong interest among scientists to participate in metadata determination.

Table 3: Summary of the self-identified research domains in the Design Phase

Experts' Self-Identified Research Domains	Count
Anthropology	17
Informatics	9
Medicine	5
Epidemiology/Public Health	4
Other Biomedical Research	3
Dentistry	2
Health Services Research	2

The first questionnaire contained 50 straw man database fields for experts to evaluate. (See Appendix B). If a field was suggested for elimination by the expert, they were asked to provide a reason. At the end of each section (personal characteristics, lifestyle, health, occupation, and other), experts were asked what additional database fields they would like to have included in the MDS. The experts suggested a wide variety of additional database fields and consensus was defined as 60% agreement for Round 1. This level of consensus was selected since 41/50 variables (82%) had over an 80% agreement. In this first round, only three variables (last name, first name, and current residence address) were eliminated from the list.

The results were summarized and similar suggestions combined, thereby creating the questionnaire for Round 2 with roughly 120 database fields for experts to evaluate. Thirty-three participants (78.6% of round 1 participants) responded to round 2 of the Delphi questionnaire. (See Appendix C). The second round was finished within a 2-week period. Because agreement on inclusion of the database fields was extremely high consensus was defined as 93%, which eliminated a large number of database fields (see Figure 2). If a lower cut-off point had been selected the number of variables would have increased by 50. As a result, a 93% cut-off point was selected in order to keep the number of database fields within a range feasible for implementation (i.e. under 60). See Table 4 for the 17 variables with a 100% consensus level in round 2.

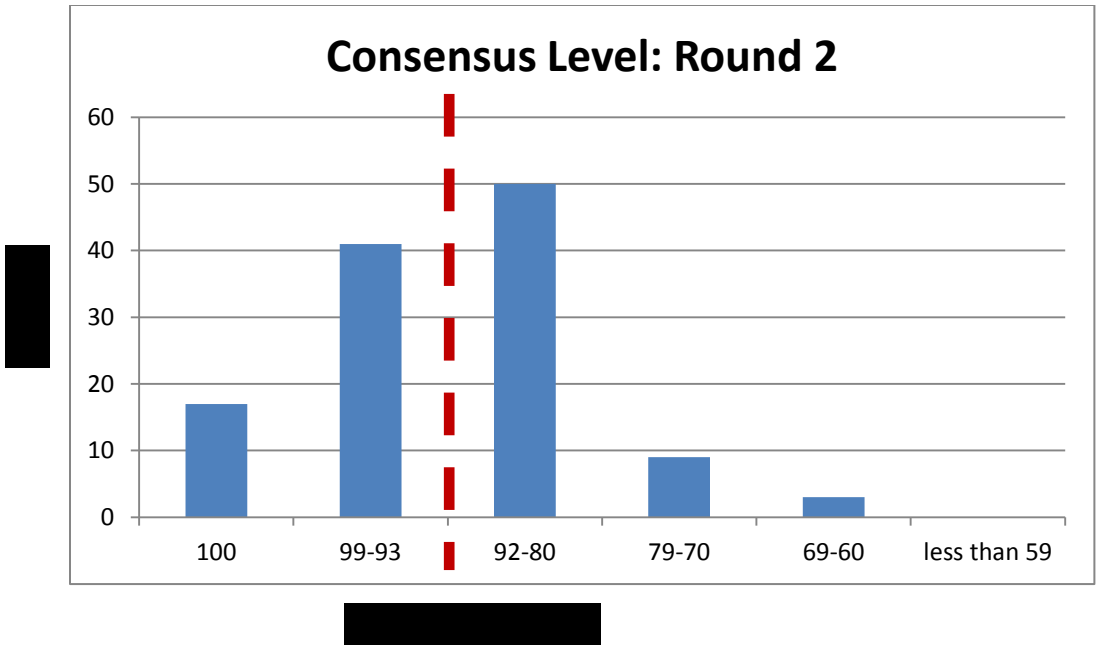

Figure 2. Consensus cut-off point for round 2

Table 4. Variables with 100% consensus in round 2

Variables selected
Number of live births
History of radiation therapy
Primary cause of death
Time delay between death and CT scan
CT scanner settings
Current residence zip code
Sex/gender
Medical diagnoses
Dental health as an adult
Contributing cause of death
Name of person entering information into database
Location of death
Country of origin
Current medications
Current occupation
Manner of death
Environmental conditions of cadaver

After elimination of database fields with less than a 93% consensus, one additional field was added back into the potential MDS. Normal height was on the cusp of being included as a variable since it had 90.3% agreement between the participants and without its inclusion, cadaveric height (96.7% agreement) would be a less useful

variable. A total of 59 database fields remained after Round 2 of the Delphi Method determination of the MDS.

The last round (Round 3 for this research project) asked experts to rate the remaining database fields in terms of importance as consensus of over 93% had been reached (see Appendix D). Twenty-two participants from Round 2 participated (67%) and ranked the remaining 59 database fields in order of importance in the MDS for future research. In addition, they were asked to supply 2 to 3 additional experts to evaluate the MDS they had created (Snowball Method). See Figure 3 for Design Phase participation.

The ranking of all the 59 MDS database fields allows for future down-scaling of variables if funding for the full set is not obtained. As such, each expert was asked whether the variable should be included in the MDS and to rank its importance. This allowed a point to be set, in the future, based on resources available to fund collection and data entry of MDS data. For Round 3, no variables had 100% agreement. However, 22 variables had over 95% agreement and 41 variables had over 90%. All 59 variables are included if the cut-off point is at 76% agreement. Rankings of importance are given for each variable on a scale of 0 to 10 from 0 = not important at all to 10 = absolutely essential. See Table 5 for the varying levels of agreement and the associated number of selected MDS variables. See Appendix E for the list of the actual variables selected by MDS size. Depending on the funding awarded for the creation of the Cadaveric Image

Database MDS, the consensus and ranking rates can be used to scale the project accordingly.

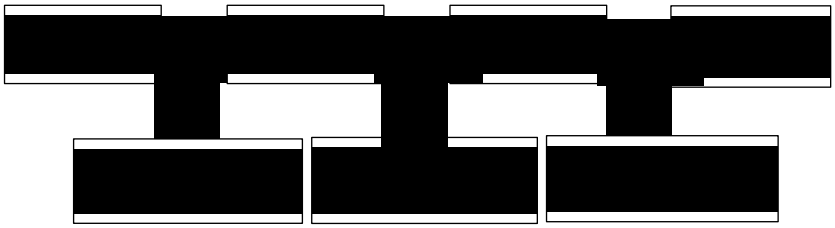

Figure 3: Design Phase expert participation for each round of questionnaires

Table 5. Varying consensus ranking and number of variables in MDS.

Consensus rate (and ranking)	Number of variables in MDS
76%	59
90%	41
95%	22
95% and a median rank of 9	5
95% and a median rank of 8	15
95% and a median rank of 7	21

Part 2: Validation Phase

A total of 34 experts were suggested by 15 Design Phase participants using the Snowball Method. Six of those suggested had already participated in the Design Phase and were eliminated. Two suggested experts for the Validation Phase had been on the Design phase list but not responded to the questionnaires and therefore were included. See Appendix F for the validation questionnaire. One of the Snowball Sample participants also requested that three additional experts be emailed the questionnaire, resulting in a second wave of the Snowball Sample.[34] As a result, a total of 31 experts were emailed a letter of participation, consent form, and link to the REDCap database for validation of the MDS. Seventeen (54%) of participants responded and ranked the

database fields in order of importance to future research (from 0 = not important at all to 10 = absolutely essential). See Table 6 for the self-identified domains of each participant.

Table 6. Expert's self-identified research domains for the Validation Phase.

Experts' Self-Identified Research Domains	Count
Anthropology	9
Medicine	2
Forensics	2
Dentistry	1
Interprofessional Collaboration	1
Medical Devices	1

No variables were selected for deletion from the MDS during the validation Phase. The level of consensus was lower during this portion; however, the majority of variables had greater than 70% consensus. Only one variable had consensus as low as 46.2% (Length of military service), demonstrating that the variables selected by the Design Phase participants were thorough in the selection process. See Appendix G for the consensus level, ranking scores and 95% confidence interval of the MDS during the Validation Phase.

The validation Phase also allowed for additional variables not included in the MDS to be elucidated. The majority of the variables suggested by the Validation Phase participants can be incorporated into variables already included in the MDS as optional responses or inferred from other variables. Fourteen variables were suggested for addition by the Validation Phase participants, with only 21% (3) not included or inferred

from the original MDS. See Table 7 for the list of suggested variables for inclusion in the MDS.

Table 7. Variables suggested by the Validation Phase participants for inclusion in the MDS.

Variables Suggested for Inclusion	MDS variable it can be inferred from	MDS variable it can be an additional response	Number of participants suggesting change
Absence/presence of removable dental implants		Implanted Devices	1
Occupation of parents	Childhood socioeconomic status		1
Income of parents	Childhood socioeconomic status		1
Income of decedent	Adult socioeconomic status		1
Exercise habits	Habitual activity		1
How consistent was exercise	Habitual activity		1
Was the individual an athlete	Habitual activity		1
Presence of amputations		Major surgeries	1
Presence of surgical implants		Implanted devices	1
Trauma present at death	History of broken bones, primary cause of death, and contributing cause of death		2
Age	Date of death and Date of birth		1
Maxilo-facial skeletal category			1
Dental occlusion category			1
Organ weights			1

Chapter 5: Discussion

The response rate was excellent with over 50% for Round 1 (58%). The participants were mostly academicians that are extremely busy and notoriously averse to email surveys. The high response rate indicates a great deal of interest in this specific application to create an extraordinary research database and only an acknowledgment by the experts of the importance of a good MDS design of a database for research. By providing a well-structured, efficient, and convenient way to participate, a sufficient number of barriers have been removed to make participation more attractive and worthwhile than other methods such as a face-to-face meeting.

Metadata and the optimization of the MDS are essential to the future value of any research database. This is especially true in the realm of image databases. The technology to search on the images themselves is in its infancy[35] and the discovery of specific images of interest relies heavily on the quality of the metadata design. Without sufficient metadata, images will be significantly less discoverable, making the sensitivity and specificity of a search or query to decrease markedly. As a result, metadata is a vital and yet a complicated concept for medical images that requires a thoughtful balance between discoverability of relevant images and the resources necessary to design and construct a sufficient MDS at the outset.

Metadata can be determined by metadata experts, the research team or through collaborative efforts with national and international experts. Using a consensus method with experts in the field is the best approach to improve the quality and completeness

of the chosen variables as well as eliminate bias.[21] This study used an electronic collaborative model, an electronic Delphi Method, to ascertain the MDS for a full-body, 3-D, cadaveric image database. Perhaps more important than the method used to design the MDS is the validation of the MDS in order to verify that the original group of experts did not suffer from "groupthink" or overlook important fields during the determination process.

The Validation Phase group of participants included both individuals that the first group (Design Phase) of experts knew personally as well as others whose work they admired or thought was amenable to using this database. The validation group of experts was selected using a snowball sampling technique to take advantage of the professional network of the experts in the Design Phase group. This was not only convenient and helpful to use, but it also helped to reduce potential bias the authors introduced into the initial expert selection process of the Design Phase group.

There are multiple benefits to this method for MDS creation and validation. The most valuable benefit is that by using an electronic consensus method to determine a MDS there is input from a large number of experts from multiple domains. If a meeting is held in person, the number of individuals that are invited is dependent upon the funding and the other limitations imposed by requiring travel (e.g. time, schedule availability). By using an electronic method, especially with the aid of the REDCap survey tool, a very large number of experts can be invited to participate, most likely larger and more diverse than could be scheduled and physically travel to a face-to-face meeting.

The use of the REDCap questionnaire also eased the data management burden of the investigators. This tool allows creation of questionnaires with multiple question formats, such as: yes/no, multiple choice options, open ended responses, and a visual analog scales. This provides a wide breadth of options for investigators. The tool is especially useful for determining the level of consensus as it provides a graphical representation of the summary of participants' responses (Figure 4 and 5). Furthermore, REDCap provides a way to email all participants a unique hypertext link to take the survey. REDCap allows investigators to track participation by documenting who has begun, finished or not responded to the questionnaire.

Furthermore, using a consensus method is strengthened by including a rigorous validation of the selected MDS. In the majority of current research databases, database fields are not validated before implementation, let alone by a group of experts that are suggested by the first panel.

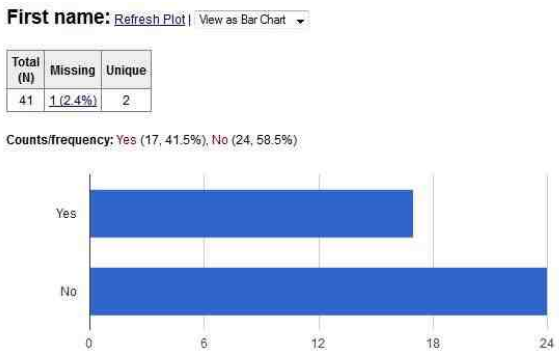

Figure 4: Example of the REDCap graphical and statistical summary of a variable with a low level of consensus.

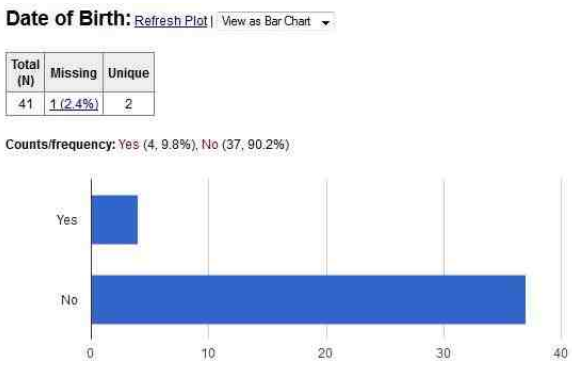

Figure 5: Example of the REDCap graphical and statistical summary of a variable with a high level of consensus.

Although this method provided many benefits, there were a few caveats. Most importantly, experts involved in research and education are extremely busy and as such, the rounds must be of the appropriate length. The survey should take no longer than 15 minutes for each individual to complete. Similarly, starting with a basic list (straw man) of database fields respected the participant's time by not requiring them to include variables that are arguably universal (e.g. age, gender). This was beneficial in minimizing the time for the expert to complete each survey. A large number of the experts I selected to participate also taught classes. A majority of participants expressed a wish to participate over semester breaks. This knowledge provides a future avenue to pursue when administering questionnaires to academicians.

A drawback to using an iterative process is that the number of participants decreased between each successive round which could potentially introduce bias. This

could have been mitigated by having a progressive monetary incentive for participation (i.e. $5.00 for Round 1 and $20.00 for Round 2 participation). However, this study, with no incentives, received over 58% participation in Round 1 during an academic semester. This indicates that researchers are willing and able to participate in metadata determination for large databases that they may use in the future.

Although this method is robust in its ability to identify potential "future-proof" metadata, it is not infallible. Not all variables are discoverable after three rounds with experts suggesting and editing metadata fields and a validation round in which additional participants can recommend further fields. No researchers mentioned marital status as a field even though it is commonly included in multiple health datasets and is a good indicator of health. [36-38]

With regards to the analysis, there are also some important unexpected results. Initially, we believed the level of consensus on database fields would be low and might take many rounds to achieve saturation due to the varying fields queried. Participants were selected from a wide array of domains. These individuals are mostly academicians and are notoriously busy and often averse to completely email surveys. However, the level of consensus was extremely high and the number of rounds to reach consensus was less than what I had originally anticipated. For Round 1, the consensus level was purposely kept low at 60% and only 3 fields were eliminated. After experts suggested a large number of additional fields, Round 2 required a much stricter cut off point in order to keep the number of database fields under 100. Our original target was for between

30 and 40 final fields, in order to keep implementation costs to a minimum. After a consensus of 93% was imposed on Round 2, 59 variables remained. For Round 3, the experts were not only asked if a field should be kept in the database but also how important it was to future research within the database. This provided us with the ability to create a sliding cut-off point depending on how many final fields we wish to include in the database since consensus was high. This may be useful if more or less funding is secured for the creation of the database and the number of database fields can be adjusted.

Some of the variables chosen as important and those eliminated are surprising. The final list of database fields (n=59) contained only 17 original variables (29%) selected by Drs. Daneshvari, Edgar and Kroth. The vast majority of final fields were suggested by the experts and validated by a separate group. This supports the value of a consensus method incorporating opinions beyond the immanent project.

Chapter 6: Conclusion

Using virtual Delphi and Snowball methodologies to obtain consensus can be an extremely beneficial tool for MDS design. These two methods require a large number of experts to weigh in and can be conducted at a relatively low cost. Furthermore, by allowing consensus among disparate researchers, some bias inherent in one individual's metadata creation can be eliminated.

Although the busy lives of experts can make scheduling of the questionnaires complicated, the first round response rate remained over 58% without any compensation. Additionally, the experts had a high level of consensus among themselves. In the future, the response rate may be increased by offering a financial compensation or conducting the questionnaires over academic breaks.

The use of a research tool, such as REDCap, is also an important factor in the success of a virtual consensus project. The program significantly reduces the data management and survey tracking burden on the investigator and produces helpful graphical summaries at various points in the project.

By including a rigorous validation phase in the process, the MDS has a better chance of being useful to a wide array of investigators both now and in the future. It is difficult to ensure any database will be "future-proof." However, the database is likely to remain more future-proof than if a metadata expert, or the original team of investigators designed the metadata. If only database creators (Daneshvari, Edgar and Kroth) had been consulted for MDS creation, over 70% of vital variables would not have

been captured. This research validates what Greenberg and Robertson[21] argue is the best method for creating a MDS -- expert group opinion.

Appendices

A. **Consent form and Letter of Participation for Design Phase**
B. **Questionnaire for Round 1 of Design Phase**
C. **Questionnaire for Round 2 of Design Phase**
D. **Questionnaire for Round 3 of Design Phase**
E. **Minimum Data Set Variables (Consensus and Ranking scores for Design Phase)**
F. **Questionnaire for Validation Phase**
G. **Consensus and Ranking scores for Validation Phase**

A. Consent form and Letter of Participation for Design Phase

The University of New Mexico Health Sciences Center
Consent to Participate in Research:
Metadata Determination for a Cadaveric Collection

You are being asked to participate in a research study that is being conducted by Shamsi Daneshvari, Ph.D., who is the Principal Investigator and Phil J. Kroth, M.D., M.S., Co-Investigator, from the Health Sciences Library and Informatics Center, and Heather Edgar, Ph.D. Co-Investigator, from the Department of Anthropology. This research is studying development of a "Minimum Data Set" for a 3-D image database. A MDS is the minimum number of variables needed to enable efficient and effective location of cases that meet the inclusion criteria for research studies. You are being asked to participate in this study because you are an expert in one or more of the following fields: anthropology, biomechanics, demography, dentistry, epidemiology, forensics, health, health disparities, human development, and medicine. About 100 experts will participate internationally.

If you agree to participate, we ask you to complete our first "Delphi Method" survey (the link to survey is included in your e-mail and/or letter) to provide your views about an initial list of variables to consider including in the Minimum Data Set (MDS). These initial variables were selected by this project's investigators from the medical, forensic, and anthropological literatures. You will be asked to suggest variables that you believe should be added to the list of variables to be included in the 3-D imaging database, and you also will be asked to suggest deletion of unnecessary listed variables. The purpose of this survey is to expand and refine the list of variables to include in the MDS to enable the database to be effectively searched for cases that meet inclusion criteria for diverse research studies in the future.

The Delphi Method requires experts to respond confidentially to 3 to 5 short, iterative surveys of about 15 minutes each spaced over 3 to 5 months (about an hour or so of total time). We will email/mail a link to each survey every 4 weeks or so, until it is clear that expert consensus has been reached. Our follow-up 3-5 surveys will allow you to revise the expert groups' ideas about variables to include and also to rate the suggested variables in terms of importance of inclusion in the MDS. You are asked to complete each survey either using the website link provided or request and complete a printed survey via mail to: Dr. Shamsi Daneshvari, UNM Health Sciences Library and Informatics Center, MSC 09 5100, 1 University of New Mexico, Albuquerque, NM 87131-0001.

This study involves only a very small risk of loss of confidentiality of survey responses and a small loss of privacy. We will make every reasonable measure to protect your privacy and confidentiality. Names, addresses and email addresses of expert participants will be stored on a secure HSC password protected server and on password protected UNM computers of the investigators. The benefit of this study is that it begins the process or creating an annotated full-body, cadaveric, 3-D image that will eventually be available to diverse researchers and educators world-wide.

We ask you to participate in this project as one of our experts. At this early stage, this is an unfunded initial study, so we are unable to provide any compensation for your contribution except our thanks. Your contributions will have a critical impact on this unique project. We know of no other efforts nationally to construct a database of cadaver images with associated data that will be made available to researchers around the world.

If you have any questions and concerns at any time about this research study, please call Shamsi Daneshvari, Ph.D., at (505) 379-3111 to discuss them. If you would like to speak with someone other than the research team, you may call the University of New Mexico Institutional Review Board at (505) 272-1129, which provides oversight for human participant research (see also: http://hsc.unm.edu/som/research/hrrc/).

Your completion and return of the enclosed survey indicates your agreement to participate in this study. Of course, you can stop your participation at any time. If you choose, you may return an email to Dr. Daneshvari (shamsi@salud.unm.edu) at any time to ask us to remove your name from our list of expert participants.

Thanks again very much for your participation in advance.

UNM HEALTH SCIENCES LIBRARY and INFORMATICS CENTER

September 30, 2013

Your expertise is needed as part of an extraordinary research opportunity developing at the University of New Mexico. New Mexico has a centralized medical examiner (ME) system affiliated with our University, where statewide decedent data are stored in one place. The situation is very different in most other states, which have county-based ME systems where data are maintained independently by each county according to its own standards. For the last 2 years, the NM ME has been performing full-body, high resolution CT scans on all cadavers autopsied (about 2,000 per year). This data source represents the extremely diverse racial and ethnic population that lives in NM, with roughly 30% Hispanic, 9% Native American, and 3% African-American individuals. This extraordinary data was collected as part of an extramurally funded research project to evaluate the efficacy of the virtual autopsy. Unfortunately, there were no provisions in the funded project design to structure this data in a way so that it could be used for scientific purposes beyond the questions specific to the funded work. A small group of investigators at UNM is trying to repurpose this treasure trove of imaging so that it can be useful in a myriad of scientific fields. As part of the first step of this process, we need your help to identify what critical patient data should be associated with each 3-D image.

More Background: The State of New Mexico's Office of the Medical Investigator (OMI) is the centralized ME office. All deaths with unknown or questionable causes that occur throughout the entire State, with a few exceptions, are processed by the OMI. Since the beginning of their extramural grant, every decedent that undergoes a postmortem examination is imaged from head to toe, with limited clinical and other information captured and stored in the OMI database. Because the sample of images is large and diverse, the images and their associated data could represent an important resource for researchers in studies of in many scientific fields including: anthropology, biomechanics, demography, dentistry, epidemiology, forensics, health, health disparities, human development, and medicine.

Our Methods: The first step in this process is to determine the optimal Minimum Data Set (MDS) to associate with the 3-D images. A MDS is the minimum number of variables needed to enable efficient and effective location of cases that meet the inclusion criteria for research studies. To achieve this aim, we must engage experts such as you, from multiple professional fields, to share their expert opinions regarding the data elements to comprise the MDS.

The creation of the MDS will be conducted using a modified consensus group method (the *Delphi Method* developed by the RAND Institute) (*1*), in which experts from a wide array of fields are solicited for their opinions. This method requires experts to respond confidentially to a sequence of 3 to 5 short, iterative surveys (~15 minutes each). We will email successive surveys every 4 weeks, or so, until it is clear that expert consensus is reached.

We hope that you will be able to participate in this project as one of our experts. At this early stage, this is an unfunded study and we are unfortunately unable to provide you any compensation for your contribution except for our thanks. Your contributions will have a critical impact on this unique project. We know of no other efforts nationally to construct such a large, structured database of 3-D cadaveric images representing such a diverse population containing research data elements amenable to research the world over.

What to do: If you agree to participate, please read and agree to the attached consent form. Information regarding how to complete the survey is contained within the consent form.

Sincerely,

Shamsi Daneshvari, Ph.D.
Biomedical Informatics Fellow
Health Sciences Library and Informatics Center

Philip J. Kroth, M.D.
Director, Biomedical Informatics, Research, Training, and Scholarship
Health Sciences Library and Informatics Center

Heather J.H. Edgar, Ph.D.
Assistant professor
Dept. of Anthropology
Curator of Human Osteology
Maxwell Museum of Anthropology

Dalkey, Norman Crolee. The Delphi Method: An Experimental Study of Group Opinion. Santa Monica, CA: RAND Corporation, 1969. http://www.rand.org/pubs/research_memoranda/RM5888

B. Questionnaire for Round 1 of Design Phase

Delphi Questionnaire to Determine the Minimum Data Set for a Cadaveric Image Database

Thank you for agreeing to particiapte in the determination of the Minimum Data Set for a cadaveric collection. Prior to ansering questions, please review the consent form that was emailed to you.

This is the first delphi method questionnaire.

The questionnaire is divided into 5 categories:Personal , lifestyle , health, occupational and other characteristics.

Please indicate if a variable should be eliminated from the Minimum Data Set (MDS) (and a reason for its elimination) within each category.

At the end of the category, list any additional variables that should be associated with the 3D, whole body, cadaveric CT scans.

This section will be used only to determine who has completed the questionnaire.

Name _____

Email address _____

Primary field of research _____

Personal Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Last name
☐ Yes
☐ No

Reason(s) for eliminating last name _____

First name
☐ Yes
☐ No

Reason(s) for eliminating first name _____

Date of Birth
☐ Yes
☐ No

Reason(s) for eliminating date of birth _____

Date of death
☐ Yes
☐ No

Reason(s) for eliminating date of death _____

Current residence address
☐ Yes
☐ No

Reason(s) for eliminating current residence address _____

Length at current residence
☐ Yes
☐ No

Reason(s) for eliminating length at current residence _____

Marital status
☐ Yes
☐ No

Reason(s) for eliminating marital status _____

Sex/gender
☐ Yes
☐ No

Reason(s) for eliminating sex/gender _____

Race (standard census categories)
☐ Yes
☐ No

Reason(s) for eliminating race _____

Hispanic ethnicity
☐ Yes
☐ No

Reason(s) for eliminating Hispanic ethnicity _____

Country of origin
☐ Yes
☐ No

Reason(s) for eliminating country of origin _____

Parents' country of origin
☐ Yes
☐ No

Reason(s) for eliminating parents' country of origin _____

Number of pregnancies (females)
☐ Yes
☐ No

Reason(s) for eliminating number of pregnancies _____

Number of live births
☐ Yes
☐ No

Reason(s) for eliminating number of live births _____

Number of living offspring
☐ Yes
☐ No

Reason(s) for eliminating number of living offspring _____

Annual income
☐ Yes
☐ No

Reason(s) for eliminating annual income _____

Highest education level
☐ Yes
☐ No

Reason(s) for eliminating highest education level _____

Handedness
☐ Yes
☐ No

Reason(s) for eliminating handedness _____

List any additional PERSONAL CHARACTERISTIC variables that should be associated with the 3-D, whole-body, cadaveric CT scans.

Additional PERSONAL CHARACTERISTIC variables to associate with the 3-D, whole-body, cadaveric CT images.

_____

(List as many as desired.)

Lifestyle Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Hobbies
☐ Yes
☐ No

Reason(s) for eliminating hobbies _____

Current exercise status
☐ Yes
☐ No

Reason(s) for eliminating current exercise status _____

Exercise histroy
☐ Yes
☐ No

Reason(s) for eliminating exercise history _____

Current smoking status
☐ Yes
☐ No

Reason(s) for eliminating current smoking status _____

Smoking history
☐ Yes
☐ No

Reason(s) for eliminating smoking history _____

Current drinking status
☐ Yes
☐ No

Reason(s) for eliminating current drinking status _____

Drinking history
☐ Yes
☐ No

Reason(s) for eliminating drinking history _____

Current drug use
☐ Yes
☐ No

Reason(s) for eliminating current drug use _____

Drug use history
☐ Yes
☐ No

Reason(s) for eliminating drug use history _____

List any additional LIFESTYLE CHARACTERISTIC variables that should be associated with the 3-D, whole-body, cadaveric CT scans.

Additional LIFESTYLE CHARACTERISTIC variables to associate with the 3-D, whole-body, cadaveric CT images.

_____

(List as many as desired.)

Health Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Medical diagnoses
☐ Yes
☐ No

Reason(s) for eliminating medical diagnoses _____

Surgical history
☐ Yes
☐ No

Reason(s) for eliminating surgical history _____

Height
☐ Yes
☐ No

Reason(s) for eliminating height _____

Current weight
☐ Yes
☐ No

Reason(s) for eliminating current weight _____

Weight history
☐ Yes
☐ No

Reason(s) for eliminating weight history _____

Childhood health status
☐ Yes
☐ No

Reason(s) for eliminating childhood health status _____

Diabetes history
☐ Yes
☐ No

Reason(s) for eliminating diabetes history _____

Family history of cancer
☐ Yes
☐ No

Reason(s) for eliminating family history of cancer _____

Cancer diagnosis
☐ Yes
☐ No

Reason(s) for eliminating cancer diagnosis _____

High blood pressure history
☐ Yes
☐ No

Reason(s) for eliminating high blood pressure history _____

History of broken bones
☐ Yes
☐ No

Reason(s) for eliminating history of broken bones _____

History of other diseases/disorders
☐ Yes
☐ No

Reason(s) for eliminating history of other diseases/disorders _____

Dental health as a child
☐ Yes
☐ No

Reason(s) for eliminating dental health as a child _____

Dental health as an adult
☐ Yes
☐ No

Reason(s) for eliminating dental health as an adult _____

List any additional HEALTH CHARACTERISTIC variables that should be associated with the 3-D, whole-body, cadaveric CT scans.

Additional HEALTH CHARACTERISTIC variables to associate with the 3-D, whole-body, cadaveric CT images.

_____

(List as many as desired.)

Occupational Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Current occupation
☐ Yes
☐ No

Reason(s) for eliminating current occupation _____

Length at current occupation
☐ Yes
☐ No

Reason(s) for eliminating length at current occupation _____

Occupation history
☐ Yes
☐ No

Reason(s) for eliminating occupation history _____

List any additional OCCUPATIONAL CHARACTERISTIC variables that should be associated with the 3-D, whole-body, cadaveric CT scans.

Additional OCCUPATIONAL CHARACTERISTIC variables to associate with the 3-D, whole-body, cadaveric CT images.

_____

(List as many as desired.)

Other Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Primary cause of death
☐ Yes
☐ No

Reason(s) for eliminating primary cause of death _____

Secondary cause of death
☐ Yes
☐ No

Reason(s) for eliminating secondary cause of death _____

Medical insurance status
☐ Yes
☐ No

Reason(s) for eliminating medical insurance status _____

List any additional OTHER CHARACTERISTIC variables that should be associated with the 3-D, whole-body, cadaveric CT scans.

Additional OTHER CHARACTERISTIC variables to associate with the 3-D, whole-body, cadaveric CT images.

_____

(List as many as desired.)

C. Questionnaire for Round 2 of Design Phase

Round 2: Delphi Questionnaire to Determine the Minimum Data Set for a Cadaveric Image Database

Thank you for agreeing to particiapte in the determination of the Minimum Data Set for a cadaveric collection. Prior to ansering questions, please review the consent form that was emailed to you.

This is the second delphi method questionnaire.

The questionnaire is divided into 5 categories:Personal , lifestyle , health, occupational and other characteristics.

Please indicate if a variable should be eliminated from the Minimum Data Set (MDS) (and a reason for its elimination) within each category.

At the end of the category, list any additional variables that should be associated with the 3D, whole body, cadaveric CT scans.

This section will be used only to determine who has completed the questionnaire.

Name _____

Email address _____

Personal Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Social security number (for linking databses- not available to researchers without IRB approval)
☐ Yes
☐ No

Reason(s) for eliminating Social Security Number _____

Last known phone number (for database linking only)
☐ Yes
☐ No

Reason(s) for eliminating last known phone number _____

Age at death
☐ Yes
☐ No

Reason(s) for eliminating age at death _____

Date of Birth
☐ Yes
☐ No

Reason(s) for eliminating date of birth _____

Year of birth
☐ Yes
☐ No

Reason(s) for eliminating year of birth _____

Date of death
☐ Yes
☐ No

Reason(s) for eliminating date of death _____

Current residence zip code
☐ Yes
☐ No

Reason(s) for eliminating current zip code _____

Length at current residence
☐ Yes
☐ No

Reason(s) for eliminating length at current residence _____

Marital status
☐ Yes
☐ No

Reason(s) for eliminating marital status _____

Sex/gender
☐ Yes
☐ No

Reason(s) for eliminating sex/gender _____

Race (standard census categories)
☐ Yes
☐ No

Reason(s) for eliminating race _____

Hispanic ethnicity
☐ Yes
☐ No

Reason(s) for eliminating Hispanic ethnicity _____

Country of origin
☐ Yes
☐ No

Reason(s) for eliminating country of origin _____

Number of years in the United States if born elsewhere
☐ Yes
☐ No

Reason(s) for eliminating number of years in the United States if born elsewhere _____

Parents' country of origin
☐ Yes
☐ No

Reason(s) for eliminating parents' country of origin _____

Hair color
☐ Yes
☐ No

Reason(s) for eliminating hair color _____

Eye color
☐ Yes
☐ No

Reason(s) for eliminating eye color _____

Sexual Orientation
☐ Yes
☐ No

Reason(s) for eliminating sexual orientation _____

Number of pregnancies (females)
☐ Yes
☐ No

Reason(s) for eliminating number of pregnancies _____

Number of live births
☐ Yes
☐ No

Reason(s) for eliminating number of live births _____

Number of living offspring
☐ Yes
☐ No

Reason(s) for eliminating number of living offspring _____

Average annual income
☐ Yes
☐ No

Reason(s) for eliminating average annual income _____

Annual household income
☐ Yes
☐ No

Reason(s) for eliminating annual household income _____

Highest education level
☐ Yes
☐ No

Reason(s) for eliminating highest education level _____

Childhood socioeconomic status
☐ Yes
☐ No

Reason(s) for eliminating childhood socioeconomic status _____

Adult socioeconomic status
☐ Yes
☐ No

Reason(s) for eliminating adult socioeconomic status _____

Preferred language
☐ Yes
☐ No

Reason(s) for eliminating preferred language _____

Handedness
☐ Yes
☐ No

Reason(s) for eliminating handedness _____

Shoe size
☐ Yes
☐ No

Reason(s) for eliminating shoe size _____

Religious affiliation
☐ Yes
☐ No

Reason(s) for eliminating religious affiliation _____

List any additional PERSONAL CHARACTERISTIC variables that should be associated with the 3-D, whole-body, cadaveric CT scans.

Additional PERSONAL CHARACTERISTIC variables to associate with the 3-D, whole-body, cadaveric CT images.

_____

(List as many as desired.)

Lifestyle Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Repetitive or habitual activities
☐ Yes
☐ No

Reason(s) for eliminating repetitive or habitual activities _____

Hobbies
☐ Yes
☐ No

Reason(s) for eliminating hobbies _____

Amount of exercising
☐ Yes
☐ No

Reason(s) for eliminating amount of exercising _____

Current exercise status
☐ Yes
☐ No

Reason(s) for eliminating current exercise status _____

Exercise history
☐ Yes
☐ No

Reason(s) for eliminating exercise history _____

History of organized sports
☐ Yes
☐ No

Reason(s) for eliminating history of organized sports _____

History of running
☐ Yes
☐ No

Reason(s) for eliminating history of running _____

History of incarceration
☐ Yes
☐ No

Reason(s) for eliminating history of incarceration _____

Current smoking status
☐ Yes
☐ No

Reason(s) for eliminating current smoking status _____

Smoking history
☐ Yes
☐ No

Reason(s) for eliminating smoking history _____

Current drinking status
☐ Yes
☐ No

Reason(s) for eliminating current drinking status _____

Drinking history
☐ Yes
☐ No

Reason(s) for eliminating drinking history _____

Current drug use
☐ Yes
☐ No

Reason(s) for eliminating current drug use _____

Drug use history
☐ Yes
☐ No

Reason(s) for eliminating drug use history _____

History of injection drug use
☐ Yes
☐ No

Reason(s) for eliminating history of injection drug use _____

Dietary pattern (vegan, vegetarian, etc.)
☐ Yes
☐ No

Reason(s) for eliminating dietary pattern _____

Time spent traveling in foreign countries
☐ Yes
☐ No

Reason(s) for eliminating time spent traveling in foreign contries _____

Time spent living in foreign countries
☐ Yes
☐ No

Reason(s) for eliminating time spent living in foreign countries _____

Mobile phone usage
☐ Yes
☐ No

Reason(s) for eliminating mobile phone usage _____

List any additional LIFESTYLE CHARACTERISTIC variables that should be associated with the 3-D, whole-body, cadaveric CT scans.

Additional LIFESTYLE CHARACTERISTIC variables to associate with the 3-D, whole-body, cadaveric CT images.

_____

(List as many as desired.)

Health Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Birth weight
☐ Yes
☐ No

Reason(s) for eliminating birth weight _____

Presence of congenital abnormalities
☐ Yes
☐ No

Reason(s) for eliminating presence of congenital abnormalities _____

Medical diagnoses
☐ Yes
☐ No

Reason(s) for eliminating medical diagnoses _____

Chronic disease history
☐ Yes
☐ No

Reason(s) for eliminating chronic disease history _____

Surgical history
☐ Yes
☐ No

Reason(s) for eliminating surgical history _____

Current medications
☐ Yes
☐ No

Reason(s) for eliminating current medications _____

Normal height
☐ Yes
☐ No

Reason(s) for eliminating normal height _____

Cadaver length (height at death)
☐ Yes
☐ No

Reason(s) for eliminating cadaver height _____

Current weight
☐ Yes
☐ No

Reason(s) for eliminating current weight _____

Cadaver weight
☐ Yes
☐ No

Reason(s) for eliminating cadaver weight _____

Weight history
☐ Yes
☐ No

Reason(s) for eliminating weight history _____

Body Mass Index value
☐ Yes
☐ No

Reason(s) for eliminating Body Mass Index values _____

Waist to hip ratio
☐ Yes
☐ No

Reason(s) for eliminating waist to hip ratio _____

Current triceps fat fold measurement
☐ Yes
☐ No

Reason(s) for eliminating current triceps fat fold measurement _____

Current subscapular fat fold measurement
☐ Yes
☐ No

Reason(s) for eliminating current subscapular fat fold measurement _____

Current mid calf fat fold measurement
☐ Yes
☐ No

Reason(s) for eliminating current mid calf fat fold measurement _____

Current ischial fat fold measurement
☐ Yes
☐ No

Reason(s) for eliminating current ischial fat fold measurement _____

Current bone density
☐ Yes
☐ No

Reason(s) for eliminating current bone density _____

Current Psoas density
☐ Yes
☐ No

Reason(s) for eliminating current psoas density _____

Childhood health status
☐ Yes
☐ No

Reason(s) for eliminating childhood health status _____

Diabetes history
☐ Yes
☐ No

Reason(s) for eliminating diabetes history _____

Family history of cancer
☐ Yes
☐ No

Reason(s) for eliminating family history of cancer _____

Cancer diagnosis
☐ Yes
☐ No

Reason(s) for eliminating cancer diagnosis _____

History of radiation therapy
☐ Yes
☐ No

Reason(s) for eliminating radiation history _____

High blood pressure history
☐ Yes
☐ No

Reason(s) for eliminating high blood pressure history _____

History of broken bones
☐ Yes
☐ No

Reason(s) for eliminating history of broken bones _____

History of facial trauma
☐ Yes
☐ No

Reason(s) for eliminating history of facial trauma _____

Presence of implanted devices
☐ Yes
☐ No

Reason(s) for eliminating presence of implanted devices _____

Missing limbs
☐ Yes
☐ No

Reason(s) for eliminating missing limbs _____

Presence of a prosthesis
☐ Yes
☐ No

Reason(s) for eliminating presence of a prosthesis _____

History of joint pain
☐ Yes
☐ No

Reason(s) for eliminating history of joint pain _____

Presence of genetic disorder(s)
☐ Yes
☐ No

Reason(s) for eliminating history of genetic disorder(s) _____

Family history of genetic disorders
☐ Yes
☐ No

Reason(s) for eliminating family history of genetic disorders _____

Presence of sensory abnormalities
☐ Yes
☐ No

Reason(s) for eliminating presence of sensory abnormalities _____

Presence of scoliosis
☐ Yes
☐ No

Reason(s) for eliminating presence of scoliosis _____

History of plastic surgery
☐ Yes
☐ No

Reason(s) for eliminating history of plastic surgery _____

Behavior health diagnosis
☐ Yes
☐ No

Reason(s) for eliminating behavior health diagnosis _____

History of other diseases/disorders
☐ Yes
☐ No

Reason(s) for eliminating history of other diseases/disorders _____

History of infectious diseases
☐ Yes
☐ No

Reason(s) for eliminating history of infectious diseases _____

History of blood transfusions
☐ Yes
☐ No

Reason(s) for eliminating history of blood transfusions _____

History of blood donations
☐ Yes
☐ No

Reason(s) for eliminating history of blood donations _____

Start date of most recent inpatient stay
☐ Yes
☐ No

Reason(s) for eliminating start date of most recent inpatient stay _____

Start date of most recent outpatient visit
☐ Yes
☐ No

Reason(s) for eliminating start date of most recent outpatient visit _____

Start date of most recent nursing home stay
☐ Yes
☐ No

Reason(s) for eliminating start date of most recent nursing home stay _____

Start date of most recent hospice stay
☐ Yes
☐ No

Reason(s) for eliminating start date of most recent hospice stay _____

Dental health as a child
☐ Yes
☐ No

Reason(s) for eliminating dental health as a child _____

Dental health as an adult
☐ Yes
☐ No

Reason(s) for eliminating dental health as an adult _____

Presence of dental caries
☐ Yes
☐ No

Reason(s) for eliminating presence of dental caries _____

History of orthodontic treatment
☐ Yes
☐ No

Reason(s) for eliminating history of orthodontic treatment _____

List any additional HEALTH CHARACTERISTIC variables that should be associated with the 3-D, whole-body, cadaveric CT scans.

Additional HEALTH CHARACTERISTIC variables to associate with the 3-D, whole-body, cadaveric CT images.

_____

(List as many as desired.)

Occupational Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Current occupation
☐ Yes
☐ No

Reason(s) for eliminating current occupation _____

Length at current occupation
☐ Yes
☐ No

Reason(s) for eliminating length at current occupation _____

Major occupation during life
☐ Yes
☐ No

Reason(s) for eliminating major occupation during life _____

Occupation history
☐ Yes
☐ No

Reason(s) for eliminating occupation history _____

Exposure to carcinogens or lethal substances
☐ Yes
☐ No

Reason(s) for eliminating exposure to carcinogens or lethal substances _____

Exposure to repetitive typing
☐ Yes
☐ No

Reason(s) for eliminating exposure to repetitive typing _____

Exposure to strenuous lifting at work
☐ Yes
☐ No

Reason(s) for eliminating exposure to strenuous lifting at work _____

Veteran status
☐ Yes
☐ No

Reason(s) for eliminating veteran status _____

Length of military service
☐ Yes
☐ No

Reason(s) for eliminating length of military service _____

List any additional OCCUPATIONAL CHARACTERISTIC variables that should be associated with the 3-D, whole-body, cadaveric CT scans.

Additional OCCUPATIONAL CHARACTERISTIC variables to associate with the 3-D, whole-body, cadaveric CT images.

_____

(List as many as desired.)

Other Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Primary cause of death
☐ Yes
☐ No

Reason(s) for eliminating primary cause of death _____

Contributing cause of death
☐ Yes
☐ No

Reason(s) for eliminating contributing cause of death _____

Manner of death
☐ Yes
☐ No

Reason(s) for eliminating manner of death _____

Time delay between death and CT scan
☐ Yes
☐ No

Reason(s) for eliminating time delay between death and CT scan _____

Location of death
☐ Yes
☐ No

Reason(s) for eliminating location of death _____

Environmental conditions of cadaver
☐ Yes
☐ No

Reason(s) for eliminating environmental conditions of cadaver _____

Method used for decedent identification
☐ Yes
☐ No

Reason(s) for eliminating method used for decedent identification _____

CT scanner settings
☐ Yes
☐ No

Identification CT scanner settings _____

Medical insurance status
☐ Yes
☐ No

Reason(s) for eliminating medical insurance status _____

Name of person entering information into database
☐ Yes
☐ No

Reason(s) for eliminating name of person entering information into database _____

List any additional OTHER CHARACTERISTIC variables that should be associated with the 3-D, whole-body, cadaveric CT scans.

Additional OTHER CHARACTERISTIC variables to associate with the 3-D, whole-body, cadaveric CT images.

_____

(List as many as desired.)

D. Questionnaire for Round 3 of Design Phase

Round 3: Delphi Questionnaire to Determine the Minimum Data Set for a Cadaveric Image Database

Thank you for agreeing to particiapte in the determination of the Minimum Data Set for a cadaveric collection. Prior to ansering questions, please review the consent form that was emailed to you.

This is the third delphi method questionnaire.

The questionnaire is divided into 5 categories:Personal , lifestyle , health, occupational and other characteristics.

Please indicate if a variable should be eliminated from the Minimum Data Set (MDS) within each category, the reason for elimination as well as rank the importance of the variable for inclusion.

This section will be used to determine who has completed the questionnaire and provide information for the next phase in this research.

Name _____

Email address _____

Please provide the names and email addresses of 2-3 researchers (that you know or know of) you believe would use or benefit from this future database. We will contact them for a short survey to validate the results from this phase of the research. _____

(If you do not know the contact information, please provid

Personal Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Date of Birth

☐ Yes
☐ No

Reason(s) for eliminating date of birth

_____

Importance of including Date of birth in the database

0- Not important — 5- Somewhat important — 10- Very Important

(Place a mark on the scale above)

Date of death

☐ Yes
☐ No

Reason(s) for eliminating date of death

_____

Importance of including date of death in the database

0- Not important — 5- Somewhat important — 10- Very Important

(Place a mark on the scale above)

Current residence zip code

☐ Yes
☐ No

Reason(s) for eliminating current zip code

_____

Importance of including current zip code in the database

0- Not important — 5- Somewhat important — 10- Very Important

(Place a mark on the scale above)

Sex/gender

☐ Yes
☐ No

Reason(s) for eliminating sex/gender

_____

Importance of including sex/gender in the database

0- Not important — 5- Somewhat important — 10- Very Important

(Place a mark on the scale above)

Race (standard census categories)

☐ Yes
☐ No

Reason(s) for eliminating race

_____

Importance of including race in the database

0- Not important — 5- Somewhat important — 10- Very Important

(Place a mark on the scale above)

Country of origin

☐ Yes
☐ No

Reason(s) for eliminating country of origin

_____

Importance of including country of origin in the database

0- Not important — 5- Somewhat important — 10- Very Important

(Place a mark on the scale above)

Number of years in the United States if born elsewhere

☐ Yes
☐ No

Reason(s) for eliminating number of years in the United States if born elsewhere

_____

Importance of including number of years in the United States if born elsewhere in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Parents' country of origin

☐ Yes
☐ No

Reason(s) for eliminating parents' country of origin

_____

Importance of including parents' country of origin in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Number of pregnancies (females)

☐ Yes
☐ No

Reason(s) for eliminating number of pregnancies

_____

Importance of including number of pregnancies (female) in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Number of live births

☐ Yes
☐ No

Reason(s) for eliminating number of live births

_____

Importance of including number of live births in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Highest education level

☐ Yes
☐ No

Reason(s) for eliminating highest education level

_____

Importance of including highest education level in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Childhood socioeconomic status

☐ Yes
☐ No

Reason(s) for eliminating childhood socioeconomic status

_____

Importance of including childhood socioeconomic status in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Adult socioeconomic status

☐ Yes
☐ No

Reason(s) for eliminating adult socioeconomic status

_____

	0- Not important	5- Somewhat important	10- Very Important
Importance of including adult socioeconomic status in the database			

(Place a mark on the scale above)

Lifestyle Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Repetitive or habitual activities

☐ Yes
☐ No

Reason(s) for eliminating repetitive or habitual activities

_____

Importance of including repetitive or habitual activities in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Current smoking status

☐ Yes
☐ No

Reason(s) for eliminating current smoking status

_____

Importance of including current smoking status in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Smoking history

☐ Yes
☐ No

Reason(s) for eliminating smoking history

_____

Importance of including smoking history in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Current drinking status

☐ Yes
☐ No

Reason(s) for eliminating current drinking status

_____

Importance of including current drinking status in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Drinking history

☐ Yes
☐ No

Reason(s) for eliminating drinking history

_____

Importance of including drinking history in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Current drug use

☐ Yes
☐ No

Reason(s) for eliminating current drug use

_____

Importance of including current drug use in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Drug use history

☐ Yes
☐ No

Reason(s) for eliminating drug use history

_____

Importance of including drug use history in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Dietary pattern (vegan, vegetarian, etc.)

☐ Yes
☐ No

Reason(s) for eliminating dietary pattern

_____

Importance of including dietary pattern in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Health Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Birth weight

☐ Yes
☐ No

Reason(s) for eliminating birth weight

_____

Importance of including birth weight in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Presence of congenital abnormalities

☐ Yes
☐ No

Reason(s) for eliminating presence of congenital abnormalities

_____

Importance of including presence of congenital abnormalities in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Medical diagnoses

☐ Yes
☐ No

Reason(s) for eliminating medical diagnoses

_____

Importance of including medical diagnoses in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Surgical history

☐ Yes
☐ No

Reason(s) for eliminating surgical history

_____

Importance of including surgical history in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Current medications

☐ Yes
☐ No

Reason(s) for eliminating current medications

_____

Importance of including current medications in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Current height

☐ Yes
☐ No

Reason(s) for eliminating current height

_____

Importance of including current height in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Cadaver length (height at death)

☐ Yes
☐ No

Reason(s) for eliminating cadaver height

_____

Importance of including cadaver height in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Current weight

☐ Yes
☐ No

Reason(s) for eliminating current weight

_____

Importance of including current weight in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Cadaver weight

☐ Yes
☐ No

Reason(s) for eliminating cadaver weight

_____

Importance of including cadaver weight in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Current bone density

☐ Yes
☐ No

Reason(s) for eliminating current bone density

_____

Importance of including bone density in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Family history of cancer

☐ Yes
☐ No

Reason(s) for eliminating family history of cancer

_____

Importance of including family history of cancer in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

History of radiation therapy

☐ Yes
☐ No

Reason(s) for eliminating radiation history

_____

Importance of including radiation therapy in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

History of broken bones

☐ Yes
☐ No

Reason(s) for eliminating history of broken bones

_____

Importance of including history of broken bones in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

History of facial trauma

☐ Yes
☐ No

Reason(s) for eliminating history of facial trauma

_____

Importance of including history of facial trauma in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Presence of implanted devices

☐ Yes
☐ No

Reason(s) for eliminating presence of implanted devices

_____

Importance of including presence of implanted devices in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Presence of genetic disorder(s)

☐ Yes
☐ No

Reason(s) for eliminating presence of genetic disorder(s)

_____

Importance of including presence of genetic disorder(s) in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Family history of genetic disorders

☐ Yes
☐ No

Reason(s) for eliminating family history of genetic disorders

_____

Importance of including family history of genetic disorder(s) in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Presence of scoliosis

☐ Yes
☐ No

Reason(s) for eliminating presence of scoliosis

_____

Importance of including presence of scoliosis in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

History of plastic surgery

☐ Yes
☐ No

Reason(s) for eliminating history of plastic surgery

_____

Importance of including history of plastic surgery in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Dental health as a child

☐ Yes
☐ No

Reason(s) for eliminating dental health as a child

_____

Importance of including dental health as a child in the database

0- Not important — 5- Somewhat important — 10- Very Important

(Place a mark on the scale above)

Dental health as an adult

☐ Yes
☐ No

Reason(s) for eliminating dental health as an adult

_____

Importance of including dental health as an adult in the database

0- Not important — 5- Somewhat important — 10- Very Important

(Place a mark on the scale above)

Presence of dental caries

☐ Yes
☐ No

Reason(s) for eliminating presence of dental caries

_____

Importance of including presence of dental caries in the database

0- Not important — 5- Somewhat important — 10- Very Important

(Place a mark on the scale above)

Occupational Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Current occupation
☐ Yes
☐ No

Reason(s) for eliminating current occupation
_____

Importance of including current occupation in the database
0- Not important | 5- Somewhat important | 10- Very Important
(Place a mark on the scale above)

Length at current occupation
☐ Yes
☐ No

Reason(s) for eliminating length at current occupation
_____

Importance of including length at current occupation in the database
0- Not important | 5- Somewhat important | 10- Very Important
(Place a mark on the scale above)

Major occupation during life
☐ Yes
☐ No

Reason(s) for eliminating major occupation during life
_____

Importance of including major occupation during life in the database
0- Not important | 5- Somewhat important | 10- Very Important
(Place a mark on the scale above)

Occupation history
☐ Yes
☐ No

Reason(s) for eliminating occupation history
_____

Importance of including occupation history in the database
0- Not important | 5- Somewhat important | 10- Very Important
(Place a mark on the scale above)

Exposure to carcinogens or lethal substances
☐ Yes
☐ No

Reason(s) for eliminating exposure to carcinogens or lethal substances
_____

Importance of including exposure to carcinogens or lethal substances in the database
0- Not important | 5- Somewhat important | 10- Very Important
(Place a mark on the scale above)

Exposure to strenuous lifting at work
☐ Yes
☐ No

Reason(s) for eliminating exposure to strenuous lifting at work
_____

Importance of including strenuous lifting in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Length of military service

☐ Yes
☐ No

Reason(s) for eliminating length of military service

_____

Importance of including length of military service in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Other Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Primary cause of death

☐ Yes
☐ No

Reason(s) for eliminating primary cause of death

_____

Importance of including primary cause of death in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Contributing cause of death

☐ Yes
☐ No

Reason(s) for eliminating contributing cause of death

_____

Importance of including contributing cause of death in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Manner of death

☐ Yes
☐ No

Reason(s) for eliminating manner of death

_____

Importance of including manner of death in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Time delay between death and CT scan

☐ Yes
☐ No

Reason(s) for eliminating time delay between death and CT scan

_____

Importance of including time delay between death and CT scan in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Location of death

☐ Yes
☐ No

Reason(s) for eliminating location of death

_____

Importance of including location of death in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Environmental conditions of cadaver

☐ Yes
☐ No

Reason(s) for eliminating environmental conditions of cadaver

_____

Importance of including environmental conditions of cadaver in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Method used for decedent identification

☐ Yes
☐ No

Reason(s) for eliminating method used for decedent identification

_____

Importance of including method used for decedent identification in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

CT scanner settings

☐ Yes
☐ No

Identification CT scanner settings

_____

Importance of including CT scanner settings in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Name of person entering information into database

☐ Yes
☐ No

Reason(s) for eliminating name of person entering information into database

_____

Importance of including name of person entering information into the database in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

E. Minimum Data Set Variables (Consensus and Ranking scores for Design Phase)

Variables ordered by consensus value.

The sliding scale is provided in groups by the two double bars below.

Variable	Consensus Level (%)	Importance Median Ranking (0-10)	95% Confidence Interval
Date of death	95.5	10.0	8.8, 10.0
Contributing cause of death	95.2	9.95	9.3, 10.0
Primary cause of death	95.5	9.9	9.2, 10.0
Date of Birth	95.5	9.85	8.6, 10.0
Medical diagnoses	95.5	9.45	8.5, 10.0
Method used for decedent identification	95.5	8.9	7.1, 10.0
Current medications	95.2	8.9	4.8, 10.0
Surgical history	95.5	8.8	7.4, 10.0
Current smoking status	95.5	8.6	7.1, 10.0
Drug use history	95.5	8.4	7.5, 9.1
Environmental conditions of cadaver	95.5	8.3	7.3, 10.0
Dental health as an adult	95.2	8.3	3.0, 9.6
Drinking history	95	8.25	7.1, 10.0
Current drug use	95.5	8.2	7.3, 9.7
History of broken bones	95.2	8.2	3.4, 10.0
Presence of implanted devices	95.5	7.85	6.1, 10.0
Presence of dental caries	95.5	7.8	3.0, 10.0
Smoking history	95	7.8	6.6, 10.0
Current drinking status	95.5	7.7	7.3, 9.7
Length at current occupation	95.5	7.4	4.7, 10.0
Current occupation	95.5	7.3	6.8, 10.0
Number of years in the US if born elsewhere	95.2	6.6	3.7, 8.8
Manner of death	90.9	10.0	8.5, 10.0
Race	90.9	9.95	7.8, 10.0
Time delay between death and CT scan	90.9	9.3	8.0, 10.0
CT scanner settings	90.9	9.1	7.8, 10.0
Cadaver length	90.9	8.4	7.3, 9.8
Occupation history	90.9	8.4	1.6, 10.0
Presence of congenital abnormalities	90.9	8.3	6.6, 9.7
Presence of genetic disorder(s)	90.9	8.3	6.3, 10.0
Major occupation during life	90.9	8.2	6.5, 10.0
Adult socioeconomic	90.9	7.9	6.9, 9.8

status			
Country of origin	90.9	7.9	6.7, 9.7
Highest education level	90.9	7.5	3.7, 9.8
Number of pregnancies	90.9	7.35	3.3, 8.5
Number of live births	90.9	6.7	3.9, 8.7
Presence of scoliosis	90.9	6.0	1.5, 9.3
Sex/gender	90.5	10.0	9.9, 10.0
Location of death	90.5	8.25	5.8, 9.9
Childhood Socioeconomic status	90.5	7.1	5.9, 10.0
Current bone density	90	7.45	6.3, 10.0
Current height	86.4	8.5	7.1, 9.6
Current weight	86.4	8.2	7.6, 10.0
Exposure to carcinogens or lethal substances	86.4	8.05	2.4, 10.0
Dental health as a child	86.4	7.5	2.4, 9.5
Name of person entering information into database	86.4	7.5	1.8, 10.0
Exposure to strenuous lifting at work	86.4	7.0	5.1, 9.3
History of radiation therapy	86.4	6.95	3.1, 9.7
Birth weight	86.4	6.9	2.4, 8.5
History of plastic surgery	86.4	6.8	2.6, 8.9
Family history of cancer	86.4	6.35	1.8, 10.0
History of facial trauma	85.7	7.55	2.3, 10.0
Current residence zip code	85.7	7.05	6.3, 8.6
Cadaver weight	84.2	8.4	6.9, 10.0
Length of military service	81.8	4.1	1.1, 9.0
Family history of genetic disorder(s)	81	8.25	3.2, 9.6
Dietary pattern	77.3	7.8	2.2, 10.0
Repetitive or habitual activities	77.3	7.7	5.9, 9.2
Parents' country of origin	76.2	6.9	3.5, 9.4

F. Questionnaire for Validation Phase

Validation of the Minimum Data Set for a Cadaveric Image Database

Thank you for agreeing to particiapte in the validation of the Minimum Data Set for a Cadaveric Image Collection. Prior to ansering questions, please review the consent form that was emailed to you.

This is the validation questionnaire for a Minimum Data Set.

The questionnaire is divided into 5 categories:Personal , lifestyle , health, occupational and other characteristics.

Please indicate if a variable should be eliminated from the Minimum Data Set (MDS) within each category as well as rate the importance of each variable.

At the end of the questionnaire, please list any essential variables that have been left out of the Minimum Data Set.

This section will be used to determine who has completed the questionnaire.

Name _____

Email address _____

Primary field of research _____

Personal Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Date of Birth

☐ Yes
☐ No

Importance of including Date of birth in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Date of death

☐ Yes
☐ No

Importance of including date of death in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Current residence zip code

☐ Yes
☐ No

Importance of including current zip code in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Sex/gender

☐ Yes
☐ No

Importance of including sex/gender in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Race (standard census categories)

☐ Yes
☐ No

Importance of including race in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Country of origin

☐ Yes
☐ No

Importance of including country of origin in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Number of years in the United States if born elsewhere

☐ Yes
☐ No

Importance of including number of years in the United States if born elsewhere in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Parents' country of origin

☐ Yes
☐ No

Importance of including parents' country of origin in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Number of pregnancies (females)

☐ Yes
☐ No

Importance of including number of pregnancies (female) in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Number of live births

☐ Yes
☐ No

Importance of including number of live births in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Highest education level

☐ Yes
☐ No

Importance of including highest education level in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Childhood socioeconomic status

☐ Yes
☐ No

Importance of including childhood socioeconomic status in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Adult socioeconomic status

☐ Yes
☐ No

Importance of including adult socioeconomic status in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Lifestyle Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Repetitive or habitual activities
☐ Yes
☐ No

Importance of including repetitive or habitual activities in the database
0- Not important | 5- Somewhat important | 10- Very Important
(Place a mark on the scale above)

Current smoking status
☐ Yes
☐ No

Importance of including current smoking status in the database
0- Not important | 5- Somewhat important | 10- Very Important
(Place a mark on the scale above)

Smoking history
☐ Yes
☐ No

Importance of including smoking history in the database
0- Not important | 5- Somewhat important | 10- Very Important
(Place a mark on the scale above)

Current drinking status
☐ Yes
☐ No

Importance of including current drinking status in the database
0- Not important | 5- Somewhat important | 10- Very Important
(Place a mark on the scale above)

Drinking history
☐ Yes
☐ No

Importance of including drinking history in the database
0- Not important | 5- Somewhat important | 10- Very Important
(Place a mark on the scale above)

Current drug use
☐ Yes
☐ No

Importance of including current drug use in the database
0- Not important | 5- Somewhat important | 10- Very Important
(Place a mark on the scale above)

Drug use history
☐ Yes
☐ No

Importance of including drug use history in the database
0- Not important | 5- Somewhat important | 10- Very Important
(Place a mark on the scale above)

Dietary pattern (vegan, vegetarian, etc.)
☐ Yes
☐ No

Importance of including dietary pattern in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Health Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Birth weight

☐ Yes
☐ No

Importance of including birth weight in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Presence of congenital abnormalities

☐ Yes
☐ No

Importance of including presence of congenital abnormalities in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Medical diagnoses

☐ Yes
☐ No

Importance of including medical diagnoses in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Surgical history

☐ Yes
☐ No

Importance of including surgical history in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Current medications

☐ Yes
☐ No

Importance of including current medications in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Current height

☐ Yes
☐ No

Importance of including current height in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Cadaver length (height at death)

☐ Yes
☐ No

Importance of including cadaver height in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Current weight

☐ Yes
☐ No

Importance of including current weight in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Cadaver weight

☐ Yes
☐ No

Importance of including cadaver weight in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Current bone density

☐ Yes
☐ No

Importance of including bone density in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Family history of cancer

☐ Yes
☐ No

Importance of including family history of cancer in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

History of radiation therapy

☐ Yes
☐ No

Importance of including radiation therapy in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

History of broken bones

☐ Yes
☐ No

Importance of including history of broken bones in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

History of facial trauma

☐ Yes
☐ No

Importance of including history of facial trauma in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Presence of implanted devices

☐ Yes
☐ No

Importance of including presence of implanted devices in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Presence of genetic disorder(s)

☐ Yes
☐ No

Importance of including presence of genetic disorder(s) in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Family history of genetic disorders

☐ Yes
☐ No

Importance of including family history of genetic disorder(s) in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Presence of scoliosis

☐ Yes
☐ No

Importance of including presence of scoliosis in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

History of plastic surgery

☐ Yes
☐ No

Importance of including history of plastic surgery in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Dental health as a child

☐ Yes
☐ No

Importance of including dental health as a child in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Dental health as an adult

☐ Yes
☐ No

Importance of including dental health as an adult in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Presence of dental caries

☐ Yes
☐ No

Importance of including presence of dental caries in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Occupational Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Current occupation

☐ Yes
☐ No

Importance of including current occupation in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Length at current occupation

☐ Yes
☐ No

Importance of including length at current occupation in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Major occupation during life

☐ Yes
☐ No

Importance of including major occupation during life in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Occupation history

☐ Yes
☐ No

Importance of including occupation history in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Exposure to carcinogens or lethal substances

☐ Yes
☐ No

Importance of including exposure to carcinogens or lethal substances in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Exposure to strenuous lifting at work

☐ Yes
☐ No

Importance of including strenuous lifting in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Length of military service

☐ Yes
☐ No

Importance of including length of military service in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Other Characteristics

Should the following variables be eliminated from the Minimum Data Set?

If the variable should be eliminated, please state the reason for elimination.

Primary cause of death

☐ Yes
☐ No

Importance of including primary cause of death in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Contributing cause of death

☐ Yes
☐ No

Importance of including contributing cause of death in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Manner of death

☐ Yes
☐ No

Importance of including manner of death in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Time delay between death and CT scan

☐ Yes
☐ No

Importance of including time delay between death and CT scan in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Location of death

☐ Yes
☐ No

Importance of including location of death in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Environmental conditions of cadaver

☐ Yes
☐ No

Importance of including environmental conditions of cadaver in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Method used for decedent identification

☐ Yes
☐ No

Importance of including method used for decedent identification in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

CT scanner settings

☐ Yes
☐ No

Importance of including CT scanner settings in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Name of person entering information into database

☐ Yes
☐ No

Importance of including name of person entering information into the database in the database

0- Not important | 5- Somewhat important | 10- Very Important

(Place a mark on the scale above)

Variables not included in current list that are essential to include

Please list any additional variables that are essential to include in the minimum dataset _____

G. Consensus and Ranking scores for Validation Phase

Variable	Consensus Level (%)	Importance Median Ranking (0-10)	95% Confidence Interval
Date of Birth	85.7	10.0	8.6, 10.0
Date of death	64.3	10.0	8.5, 10.0
Current residence zip code	50.0	2.2	0.3, 5.0
Sex/gender	78.6	10.0	9.7, 10.0
Race	78.6	10.0	8.4, 10.0
Country of origin	71.4	9.45	6.8, 10.0
Number of years in the US if born elsewhere	64.3	5.45	1.1, 9.6
Parents' country of origin	71.4	6.65	4.4, 10.0
Number of pregnancies	64.3	3.7	0, 6.2
Number of live births	64.3	1.8	0, 6.6
Highest education level	57.1	2.05	0.8, 6.1
Childhood Socioeconomic status	71.4	7.1	5.9 10.0
Adult socioeconomic status	57.1	6.0	1.6, 10.0
Repetitive or habitual activities	71.4	6.55	1.3, 10.0
Current smoking status	57.1	4.9	0.6, 7.4
Smoking history	64.3	5.7	0.8, 6.8
Current drinking status	57.1	3.85	0.1, 6.3
Drinking history	71.4	5.2	0.7, 6.4
Current drug use	69.2	5.6	2.5, 7.0
Drug use history	69.2	5.6	2.2, 7.0
Dietary pattern	71.4	2.5	1.6, 8.4
Birth weight	69.2	5.1	2.3, 8.0
Presence of congenital abnormalities	76.9	7.85	5.0, 10.0
Medical diagnoses	76.9	8.75	5.0, 10.0
Surgical history	76.9	9.7	6.7, 10.0
Current medications	76.9	6.95	1.3, 9.6
Current height	76.9	9.9	7.1, 10.0
Cadaver length	83.3	10.0	7.1, 10.0
Current weight	75.0	9.5	6.5, 10.0
Cadaver weight	84.6	9.7	6.9, 10.0
Current bone density	61.5	7.8	1.6, 10.0
Family history of cancer	53.8	4.3	1.1, 9.8
History of radiation therapy	53.8	4.1	0.6, 9.6
History of broken bones	69.2	9.9	6.8, 10.0
History of facial trauma	69.2	9.55	5.0, 10.0
Presence of implanted	75.0	9.5	2.3, 10.0

devices			
Presence of genetic disorder(s)	84.6	8.95	2.4, 10.0
Family history of genetic disorder(s)	76.9	8.0	2.0, 9.8
Presence of scoliosis	92.3	8.2	1.4, 9.7
History of plastic surgery	76.9	5.7	1.7, 8.5
Dental health as a child	58.3	7.2	2.8, 9.6
Dental health as an adult	69.2	9.2	3.7, 10.0
Presence of dental caries	76.9	8.9	2.4, 10.0
Current occupation	61.5	5.0	0.8, 7.8
Length at current occupation	61.5	2.95	0, 7.5
Major occupation during life	76.9	6.5	2.8, 10.0
Occupation history	53.8	2.35	1.3, 9.1
Exposure to carcinogens or lethal substances	69.2	5.1	1.1, 10.0
Exposure to strenuous lifting at work	76.9	5.0	1.1, 9.7
Length of military service	46.2	1.4	0, 5.0
Primary cause of death	76.9	10.0	7.9, 10.0
Contributing cause of death	76.9	10.0	8.2, 10.0
Manner of death	83.3	10.0	7.4, 10.0
Time delay between death and CT scan	69.2	8.2	5.0, 10.0
Location of death	61.5	6.0	1.8, 8.0
Environmental conditions of cadaver	69.2	9.2	5.1, 10.0
Method used for decedent identification	84.6	8.65	1.8, 10.0
CT scanner settings	83.3	10.0	9.6, 10.0
Name of person entering information into database	84.6	3.3	1.4, 7.1

References

1. Cimino JJ. Desiderata for controlled medical vocabularies in the twenty-first century. Meth Inform Med. 1998;37(4-5):394-403. Epub 1998/12/29.
2. Elkin PL, Brown SH, Carter J, Bauer BA, Wahner-Roedler D, Bergstrom L, et al. Guideline and quality indicators for development, purchase and use of controlled health vocabularies. Int J Med Inform. 2002;68(1-3):175-86. Epub 2002/12/07.
3. Greenberg J. Metadata Generation: Processes, People and Tools. Bulletin of the American Society for Information Science and Technology. 2003;29(2):18-21.
4. Currier S, Barton J, O'Beirne R, Ryan B. Quality assurance for digital learning object repositories: issues for the metadata creation process. Alt-J Research in Learning Technology. 2004;12(1):5-20.
5. Sicilia M. Metadata, semantics, and ontology: providing meaning to information resources. Int J Metadata, Semantics and Ontologies. 2006;1(1):83-6.
6. Malaxa V, Douglas I. A framework for metadata creation tools. Interdisciplinary Journal of Knowledge and Learning Objects. 2005;1:151-62.
7. Oracle. An Oracle White Paper: Oracle Database 11g DICOM Medical Image Support 2009.
8. OECD. Health at a Glance 2011: OECD Indicators. OECD Publishing: 2011.
9. Muller H, Kenny E, Sternberg P. Textpresso: An Ontology-Based Information retrieval and Extraction System for Biological Literature. PLoS Biol. 2004;2(11):e309.
10. Del Bimbo A. Visual Information Retrieval. San Francisco, California: Morgan Kauffman Publishers, Inc. ; 1999.
11. Health Information Policy Council. Background paper: Uniform minimum health data sets In: Services. UDoHaH, editor. Washington, DC1983.
12. Werley H, Devine E, Zorn C, Ryan P, Westra B. The Nursing Minimum Data Set: abstraction tool for standardized, comparable, essential data. Am J Public Health. 1991;81(4):421-6.
13. Werley H, Lang N, Westlake S. Brief summary of the Nursing Minimum Data Set Conference. . Nurs Manage. 1986;17(7):42-5.
14. Meaney F, Cunningham G, Riggle S. Development of a national genetic services database Proc Annu Symp Comput Appl Med Care. 1991;1991:424-8.
15. Porock DO, DP; Zweig, S; Rantz, M; Mehr, D; Madsen, R; Petrosk, G. Predicting Death in the Nursing Home: Development and Validation of the 6-Month Minimum Data Set Mortality Risk Index. Journal of Gerontology: Medical Sciences. 2005;6A (4):491-8.
16. Office of Rare Disease Research. Global Rare Disease Patient Registry and Data Repository Common Data Elements. 2013 [cited 2013 March 1]; Available from: http://www.grdr.info/index.php?option=com_content&view=article&id=3&Itemid=5.
17. Daneshvari S, Youssof S, Kroth P. The NIH Office of the Rare Disease research Patient Registry Standard: A Report from the University of New Mexico's Oculopharyngeal Muscular Dystrophy Patient Registry. JAMIA. 2013;in press.
18. Jenders RA, McDonald CJ, Rubinstein Y, Groft SC. Applying Standards to Public Health: An Information Model for a Global Rare-Diseases Registry. AMIA 2011 Annual Symposium; Washington, D.C.2011. p. 1819.
19. Rubinstein YR, Groft SC, Bartek R, Brown K, Christensen RA, Collier E, et al. Creating a global rare disease patient registry linked to a rare diseases biorepository database: Rare Disease-HUB (RD-HUB). Contemp Clin Trials. 2010;31(5):394-404. Epub 2010/07/09.
20. IEEE LTSC. Draft standard for learning object metadata. 2002.

21. Greenberg JR, DW editor. Semantic Web construction: an inquiry of authors' views on collaborative metadata generation. Proceedings of the International Conference on Dublin Core and Metadata for e-Communities: Supporting Diversity and Convergence; 2002; Florence, Italy: Firenze University Press.
22. Thompson CS, J. Minimum data set development: Air transport time-related terms. International Journal of Medical Informatics. 2002;65:121-33.
23. Hillman D. Metadata Quality: From Evaluation to Augmentation Cataloging and Classification Quarterly. 2008.
24. Dushay NH, D. Analyzing Metadata for Effective Use and Re-use. DC-2003: 2003 Dublin Core Conference; Seattle, Washington2003.
25. Goossen TE, PJMM; Feuth, T; Dassen, TWN; Hasman, A; van den Heuvel, WJA. A comparison of nursing minimal data sets. J Amer Med Assoc. 1998;5(152-163).
26. Laws PS, EA. Report on the evaluation of the Perinatal National Minimum Data Set In: Unit ANPS, editor. Sydney, Australia2004.
27. OMI. Office of the Medical Investigator 2010 Annual Report. University of New Mexico: 2010.
28. US Census Bureau. 2010 US Census. 2010 [cited 2011 January 19, 2011]; Available from: http://2010.census.gov/2010census/pdf/2010_Questionnaire_Info.pdf.
29. Yousuf M. Using Experts' Opinions through Delphi Technique. Practical Assessment Research & Evaluation 2007;12(4).
30. Hsu CS, BA. The Delphi technique: Making sense of consensus. Practical Assessment Research & Evaluation. 2007;12(10).
31. Van De Ven AD, AL. The Effectiveness of Nominal, Delphi, and Interacting Group Decision Making Processes. The Academy of Management Journal. 1974;17(4):605-21.
32. Cuhls K. Delphi method. 2003 [cited 2013 December 10, 2013]; Available from: http://www.unido.org/fileadmin/import/16959_DelphiMethod.pdf.
33. Harris P, Taylor R, Thielke R, Payne J, Gonzalez N, Conde J. Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009;42(2):377-81.
34. Goodman LA. Snowball Sampling. Ann Math Statist. 1961;32(1):148-70.
35. Hou JC, Z; Qin, X; Zhang,D Automatic image search based on improved feature descriptors and decision tree. ICAE. 2011;18(2):167-80.
36. Robards J, Evandrou M, Falkingham J, Vlachantoni A. Marital status, health and mortality. Maturitas. 2012;73(4):295-9.
37. Verbrugge LM. Marital Status and Health. Journal of Marriage and Family. 1979;41(2):267-85.
38. Umberson D. Gender, marital status and the social control of health behavior. Soc Sci Med. 1992;34(8):907-17.