Published in final edited form as:

Magn Reson Imaging. 2012 November ; 30(9): 1249–1256. doi:10.1016/j.mri.2012.04.007.

Informatics Methods to Enable Sharing of Quantitative Imaging Research Data

Mia A. Levy, MD, PhD¹, John B. Freymann², Justin S. Kirby², Andriy Fedorov, PhD³, Fiona M. Fennessy, MD, PhD³, Steven A. Eschrich, PhD⁴, Anders E. Berglund, PhD⁴, David A. Fenstermacher, PhD⁴, Yongqiang Tan, PhD⁵, Xiaotao Guo, PhD⁵, Thomas L. Casavant⁶, Bartley J. Brown⁶, Terry A. Braun⁶, Andre Dekker⁷, Erik Roelofs⁷, James M. Mountz, MD, PhD⁸, Fernando Boada⁸, Charles Laymon⁸, Matt Oborski⁸, and Daniel L Rubin, MD, MS⁹

John B. Freymann: freymannj@mail.nih.gov; Justin S. Kirby: kirbyju@mail.nih.gov; Andriy Fedorov: fedorov@bwh.harvard.edu; Fiona M. Fennessy: ffennessy@partners.org; Steven A. Eschrich: steven.eschrich@moffitt.org; Anders E. Berglund: anders.berglund@moffitt.org; David A. Fenstermacher: david.fenstermacher@moffitt.org; Yongqiang Tan: yt2314@columbia.edu; Xiaotao Guo: xg2145@columbia.edu; Thomas L. Casavant: tomc@eng.uiowa.edu; Bartley J. Brown: bbrown@eng.uiowa.edu; Terry A. Braun: tabraun@eng.uiowa.edu; Andre Dekker: andre.dekker@maastro.nl; Erik Roelofs: erik.roelofs@maastro.nl; James M. Mountz: mountzjm@upmc.edu; Fernando Boada: boadfe@upmc.edu; Charles Laymon: cml14@pitt.edu; Matt Oborski: oborskimj@upmc.edu; Daniel L Rubin: dlrubin@stanford.edu

¹Vanderbilt University School of Medicine, Department of Biomedical Informatics and Medicine, Division of Hematology and Oncology

²Clinical Research Directorate/CMRP, SAIC-Frederick, Inc., NCI-Frederick, Frederick, Maryland 21702

³Brigham and Women's Hospital, Harvard Medical School

⁴H. Lee Moffitt Cancer Center and Research Institute

© 2012 Elsevier Inc. All rights reserved.

Corresponding Author: Mia A. Levy, MD, PhD, Vanderbilt Ingram Cancer Center, 2220 Pierce Avenue, 691 Preston Research Building, Nashville TN 37232-6838, (O) 615-936-0809, (F) 615-936-1782, mia.levy@vanderbilt.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Disclosures of conflict of interest

Mia Levy - none
John Freymann- none
Justin Kirby - none
AndriyFedorov, PhD - none
Fiona Fennessy, MD, PhD - none
Steven Eschrich - none
Anders Berglund - none
David Fenstermacher - none
Yongqiang Tan- none
XiaotaoGuo- none
Thomas L. Casavant - none
Bartley Brown - none
Terry Braun - none
Andre Dekker
Erik Roelofs
James Mountz- none
Fernando Boada - none
Charles Laymon - none
Matt Oborski - none
Daniel Rubin - none

⁵Columbia University

⁶Center for Bioinformatics and Computational Biology, Holden Comprehensive Cancer Center, University of Iowa

⁷Maastricht University Medical Centre, Maastricht, The Netherlands. Department of Radiation Oncology (MAASTRO), GROW-School for Oncology and Developmental Biology

⁸University of Pittsburgh

⁹Stanford University School of Medicine, Department of Radiology and Medicine (Biomedical Informatics Research)

Abstract

Introduction—The National Cancer Institute (NCI) Quantitative Research Network (QIN) is a collaborative research network whose goal is to share data, algorithms and research tools to accelerate quantitative imaging research. A challenge is the variability in tools and analysis platforms used in quantitative imaging. Our goal was to understand the extent of this variation and to develop an approach to enable sharing data and to promote reuse of quantitative imaging data in the community.

Methods—We performed a survey of the current tools in use by the QIN member sites for representation and storage of their QIN research data including images, image meta-data and clinical data. We identified existing systems and standards for data sharing and their gaps for the QIN use case. We then proposed a system architecture to enable data sharing and collaborative experimentation within the QIN.

Results—There area variety of tools currently used by each QIN institution. We developed a general information system architecture to support the QIN goals. We also describe the remaining architecture gaps we are developing to enable members to share research images and image meta-data across the network.

Conclusions—As a research network, the QIN will stimulate quantitative imaging research by pooling data, algorithms and research tools. However, there are gaps in current functional requirements that will need to be met by future informatics development. Special attention must be given to the technical requirements needed to translate these methods into the clinical research workflow to enable validation and qualification of these novel imaging biomarkers.

Keywords

Quantitative Imaging Network; Data sharing; Imaging Informatics; Research Informatics; Image repository; Image meta-data repository; Clinical data repository; System architecture

Introduction

Over the past decade, there has been increasing emphasis on the power of sharing research data to improve public health[1,2]. Yet few publically available clinical image repositories exist that meet the requirements for the development of novel quantitative imaging biomarkers to assess treatment response in cancer. The Cancer Imaging Program (CIP) branch of the National Cancer Institute (NCI) established the Quantitative Imaging Network (QIN) to promote research and development of quantitative imaging methods for the measurement of tumor response to therapies in clinical trial settings, with the overall goal of facilitating clinical decision-making[3]. Using the U01 grant mechanism, the QIN is designed to support both individual research programs at each of the ten currently funded

institutions as well as the development of a research network that includes sharing of expertise, data and technologies.

As a research network, one of the goals of the QIN is to promote sharing of research data, imaging algorithms, and informatics tool across member sites. One of the goals of data sharing in this context is to enable secondary reuse of research data for validation of imaging algorithms and qualification of quantitative imaging biomarkers[4]. Algorithm validation sufficient to introduce into general clinical practice requires testing image-processing algorithms on multiple imaging data sets collected on multiple devices at multiple institutions, possibly using multiple imaging modalities (e.g. generalizability of image processing algorithm across Computed Tomography (CT), Positron Emission Tomography (PET), and Magnetic Resonance Imaging (MRI)). Image biomarker qualification sufficient to change clinical practice likewise requires validation across multiple patient data sets, possibly for multiple cancer diagnoses (e.g. breast cancer and lung cancer), multiple imaging modalities (e.g CT, PET, MRI), multiple classes of therapeutic interventions (e.g. multiple classes of systemic drug therapy, radiation therapy), and multiple lines of therapy (e.g. neoadjuvant versus metastatic therapy).

For the recent update of the Response Evaluation Criteria in Solid Tumors (RECIST) criteria, from the original version published in 2000 to the RECIST 1.1 version published in early 2009, a large retrospective database of target lesions was developed to test the impact of modifications to the criteria[5,6]. Metadata on 18,000 potential target lesions were obtained from 6512 patients in 16 metastatic cancer clinical trials. These trials represented multiple types of solid tumor malignancies. The database was used to evaluate the impact of changes to RECIST on the classification of patient response to treatment. The RECIST 1.1 criterion was, thus, validated only by comparing it to the previous standard approach and not by evaluating its correlation with survival end points. These databases, however, are not publicly available to enable the further development and validation of image processing approaches to estimating tumor burden for quantitative response criteria. Many other large image databases associated with clinical trials are likewise not publically available, such as those maintained by the American College of Radiology Imaging Network (ACRIN)[7] and Quality Assurance Review Center (QARC)[8].

Unlike the RECIST 1.1 biomarker qualification study where predominantly CT data was utilized, the QIN member institutions are developing quantitative imaging biomarkers for novel and emerging imaging modalities. This presents an opportunity to collect the primary research data generated from each QIN member site for use in secondary studies of novel imaging modalities and novel imaging biomarkers. In order to facilitate development and validation of novel response criteria, large datasets are needed that contain baseline and follow-up imaging studies, imaging feature of the cancer lesions, along with the corresponding diagnoses, therapies and clinical outcomes.

The initial focus of the QIN Informatics Workgroup has been to develop a practical and sustainable strategy for data sharing across the QIN member sites. The workgroup pools expertise in biomedical and imaging informatics from across the QIN network and the cancer Biomedical Informatics Grid (caBIG) imaging workspace [9]. This paper describes the activities of the QIN Informatics Workgroup over the first year. Due to the large number of acronyms in this space, we provide Table 1 as a summary of the commonly used acronyms used in this paper.

Methods

In November 2010 the QIN informatics workspace convened its first face-to-face meeting. At that meeting, each of the QIN sites described the types of data being collected and the informatics infrastructure in use to support their individual projects. Presentations were also given of open source informatics applications and data standards that could support data acquisition, storage, and sharing. At the end of the meeting, each site agreed to complete an online survey describing in more detail their data and tools.

Each QIN site participates in at least two therapeutic cancer clinical trials where imaging studies are collected as part of treatment response assessment. In order to better understand the range of data being collected as part of the QIN projects, we asked each site to provide information regarding the data collected to support their studies. This included information regarding the cancer diagnosis and treatment, imaging modalities, tumor imaging features, and supporting clinical data. Table 2 shows the specific survey questions. In particular, with respect to treatment, we inquired regarding the type of cancer diagnosis (e.g. breast cancer, lung cancer), type of cancer treatment (e.g. drug therapy, radiation therapy), and line of treatment (e.g. metastatic therapy, neoadjuvant therapy). With respect to image data, we inquired regarding the types of imaging modalities being used for the study (e.g. CT, PET, MRI), the image file formats used for image transfer, storage, and analysis, the local systems used to store these images, and any tools in use for image de-identification (if required).

Each QIN site is developing novel approaches for the collection and evaluation of quantitative imaging biomarkers for treatment response assessment. We thus inquired regarding the types of image features of the tumor that were being collected and the methods for their collection and storage. We inquired specifically regarding the use of manual, semi-automated, and automated methods for image feature extraction. Finally, as the primary QIN research question is distinct from the primary research question of the respective clinical trials, we inquired specifically regarding the classes of non-imaging clinical data and the local systems in use to collect and store that data for the QIN research question.

The QIN survey formed the basis for the development of a set of requirements for data sharing and a proposed system architecture. To do this, we reviewed the survey results, summarized the current state of the tools available to QIN sites, identified the remaining gaps, and evaluated alternative informatics methods that could fill those gaps. The results were presented to the larger QIN network at the first face-to-face meeting in March 2011. In this open forum, we discussed a strategy for first steps towards achieving the goals of data sharing among the QIN sites. We have spent the last year implementing these first steps.

Results

Survey Results

Ten QIN sites have completed the survey and we summarize the community responses. Table 3 demonstrates a high degree of variance between sites with respect to disease and treatment, covering 9 different cancer diagnoses and multiple combinations of systemic drug therapy or radiation therapy given either in the neoadjuvant or advanced disease setting. This represents an exciting potential for collection of research data for use in biomarker validation studies across multiple cancer diagnoses and treatment types.

In contrast, there is a high degree of overlap in the imaging modalities in use among member sites with a predominant use of CT, PET, and MRI (Table 4). Likewise, the two dominant image file formats are Digital Imaging and Communications in Medicine (DICOM) and the Philips Research File Format (PAR/REC). However, there is a great deal of variance in how

each local institution is handling their research image storage. Some institutions are simply using local electronic filing systems that enable limited query or sharing of image data within the institutions. Some are using local clinical or research Picture Archiving and Communication System (PACS) that enable either temporary or permanent storage of their images and enable enhanced local image query and retrieval functionality. Others are using more sophisticated image repository systems such as the National Biomedical Image Archive (NBIA)[10] an open source application developed by the caBIG program[11], and the Extensible Neuroimaging Archive Toolkit (XNAT)[12], an open-source software tool developed through the Biomedical Informatics Research Initiative (BIRN)[13]. Both enable management of imaging projects, user permissions, and sharing of image data both within and across institutions.

Within the context of multiple standard and novel imaging modalities, there are likewise numerous tumor image features that are being extracted and evaluated including anatomic, functional, and dynamic features (Table 5). In general, these image features are being extracted through manual, semi-automated, or automated approaches. Manual approaches include the use of image annotation tools[23] or image mark-up tools as part of vendor or open source DICOM viewers. Semi-automated and automated approaches include use of vendor, open source, and custom developed image-processing algorithms. Data regarding these image features, the *image meta-data*, is locally stored in a variety of ways ranging from simple spread sheets, to local database models with proprietary schemas, to open source image annotation databases such as the Annotation Image Markup Data Service (AIME) [24] and the Biomedical Image Metadata Manager (BIMM)[25] that utilize the Annotation Image Markup (AIM)[26] information model to inform their schema.

Most QIN sites plan to record similar classes of clinical data with variations for specific cancer diagnosis (e.g. Prostate Specific Antigen (PSA)) for prostate cancer studies) (Table 6). The clinical data being captured is highly simplified compared to the details required for the clinical trials themselves and specifically does not cover treatment toxicity related information. Most studies are recording and collecting this clinical data separately from the clinical trial data collection process and using simple spreadsheet or local database solutions to store the information. Some studies are also collecting tumor gene mutation data.

Requirements for Architecture for Data Sharing

The survey of QIN member sites shows a high degree of variability in the disease and type of treatment, and a high degree of overlap in both the imaging modalities and imaging features under investigation. However, many QIN sites lack informatics infrastructures for the storage and sharing of images, image meta-data and clinical data across the network. This is in part due to the lack of mature information models and tools to achieve this goal. We thus developed a set of functional requirements for an infrastructure that would enable the QIN to share research data.

1. In order to be most useful for the goals of algorithm validation and biomarker qualification, three classes of data need to be shared:
 - a. Images (including primary clinical images and derived images)
 - b. Image meta-data
 - c. Clinical data
2. Data repositories may be centralized or federated, but must allow investigators to control permissions for data distribution. If a centralized data model was selected, a commitment was needed from the NCI to support the ongoing maintenance of the repository.

3. Systems should be integrated to enable query and retrieval of data across data repositories.
4. Systems should be integrated to support research workflow including enabling research methods to access data for analysis and visualization.
5. Only properly de-identified data should be shared. The QIN should establish consensus in adopting standards for de-identification procedures.
6. Open source tools should be used when available.
7. Standard information models should be used when available.
8. Standard terminologies should be used when available.

Proposed System Architecture

From these functional requirements, the experience of the QIN sites in using various informatics tools, and the expertise of the QIN Informatics Workgroup, we propose a systems architecture to enable data sharing, collaborative experiments, and translation of these novel response assessment methods into clinical practice (Figure 1). The systems architecture includes informatics tools to support data repositories (Figure 1 blue boxes) based on standard information models and shared semantics (Figure 1 red boxes). The data repositories can be queried and the research data retrieved for use by research methods for additional analysis (Figure 1 purple boxes). Ultimately, translation in to clinical practice of these novel quantitative imaging biomarkers for cancer treatment response assessment would require integration with clinical systems (Figure 1 green boxes). Some components of this architecture could build upon existing tools and tools under development at each institution as part of their QIN proposal. Other components represent gaps that remain. We discuss below the details of this architecture, existing systems that may be candidates to fulfill this architecture, and remaining gaps.

The data storage layer (Figure 1 blue boxes) is the core feature of the proposed system architecture and integrates with the remaining layers. The QIN has three classes of data that require acquisition, transmission, storage and retrieval including: image data, image meta-data, and non-imaging clinical data. Each class of data has unique requirements for data representation, semantics, and storage. Each class of data has information models and existing tools at varying levels of sophistication and development.

The most advanced set of existing systems are those for image repositories. DICOM[14] is a well-established standard for clinical image storage and transmission. It is used by most commercial imaging modalities to store and transmit clinical images at the time of image acquisition. Some research imaging modalities however, generate image data in other file formats that may need to be converted to DICOM in the cases when this is required by the image repository. Likewise, many sites are creating derived images from their image analysis that may like need to be converted to DICOM if required by the image repository. Several QIN sites are using NBIA[10] and XNAT[12] to store research images. Both systems are open source and enable centralized or federated storage and retrieval models. They also enable management of image collections with control over user permissions for access to and distribution of images within the collection. The NCI CIP program funds an instance of the NBIA called The Cancer Imaging Archive (TCIA)[9] that is maintained by Washington University. TCIA currently contains 18 image collections for over 2600 research subjects that are either publically available or with limited access. It is however, limited to storage of DICOM files.

Compared to image repositories, image meta-data repositories are relatively early in their development and adoption. Unlike the wide adoption of the DICOM standard, most commercial DICOM viewers maintain proprietary information models to represent, transmit, and store image mark-up. DICOM Structured Reporting (DICOM-SR) [30] provides support for standardized template-based markup and annotation and is supported by several commercial products, such as Siemens syngo.via. Open source toolkits are available for interfacing DICOM-SR reports such as the DICOM toolkit (DCMTK) [31]. The Annotation Image Mark-up (AIM) standard [26] developed as part of the caBIG imaging workspace [9], is based on formal semantics, and has open source tools to support its implementation including image annotation template builders that utilize controlled terminologies [32] (Figure 1 red boxes), image annotation tools integrated into open source DICOM viewers (Figure 1 green boxes) for Osirix (Electronic Physician Annotation Device (ePad) [23]) and Clear Canvas [33], and image meta-data repositories (Figure 1 blue boxes) based on the AIM schema in both relational (BIMM [25]) and XML (AIME [34]) implementations. There are even tools developed to visualize trends in quantitative lesion features for use in clinical decision support [23,25] (Figure 1 green boxes).

While some QIN sites are utilizing these open source AIM based tools to collect and store their image meta-data, most are not. The primary tools for collecting image features of cancer lesions are manual tools integrated into DICOM viewers or image processing algorithms (Figure 1 purple box). The most commonly used manual tools are simple mark-up tools such as rulers or Standardized Uptake Value (SUV) measurements. The reviewer measures the image findings and typically records the values in a separate system such as a spreadsheet or local database. Image annotation tools are likewise manual but maintain the direct link to the voxels on the image to which the observations refer. Most QIN member sites are performing some form of manual measurements since this is the current standard of care for image based treatment response assessment and forms the basis for a comparison to the novel quantitative image features under development. However, most QIN sites are also utilizing or developing image processing algorithms to extract novel quantitative image features. To date, none of these methods have utilized the AIM information model to represent the output of the image processing algorithm. Furthermore, while the AIM model is theoretically expressive enough to represent complex dynamic image features being collected by QIN sites (Table 5), such implementations have not yet been demonstrated. These issues remain a barrier to QIN adoption of either the DICOM-SR or AIM information model and associated tools.

However, the clinical data repository remains the most challenging with respect to identifying a common set of tools and semantics to represent this heterogeneous set of data. Several candidate terminologies could be used to represent the clinical data including the common data elements from the Cancer Data Standards Registry and Repository caDSR [35] (Figure 1 red box). These common data elements could be used as the basis for a common set of electronic case report forms for the QIN clinical data collection, in systems like the Cancer Central Clinical Database (C3D [36]) hosted by the NCI. However, realistically this would likely require manual re-entry to the clinical data already collected by the QIN sites. Furthermore, these clinical data repositories are not currently integrated with the image and image meta-data repositories to enable query and retrieval of linked data across all classes of data. This would be required to enable statistical methods to perform analysis of the QIN data for algorithm validation and biomarker qualification (Figure 1 purple box).

Initial strategy for implementation of QIN data sharing

While several systems exist to support components of the proposed QIN systems architecture, many gaps remain. As an initial step towards achieving the goal of data sharing, the QIN sites agreed to utilize TCIA as a central repository to upload their research

images, initially on a limited access basis. This was in large part because CIP funds the ongoing maintenance of TCIA and the system supports an appropriate level of user permissions for early stage research projects. The sites also agreed on the use of an RSNA Clinical Trial Processor(CTP) image de-identification protocol to be used prior to uploading images. To date, eight QIN sites have begun the process of creating their respective image collections and uploading images to TCIA.

Having agreed on an approach for sharing research images, the QIN informatics workgroup then turned to the issue of sharing image meta-data. They agreed to investigate the use of AIM and AIM-based tools to represent, transmit and store image meta-data. In order to move this forward, CIP funded the development of an AIM application programming interface (API) that developers can use to facilitate incorporating AIM into their applications. The java-based AIM API permits them to save image features in the AIM XML format. Version 1 of the AIM API has recently been released[37]. To facilitate the adoption of AIM by the end users and fill the gap in the availability of software tools that support image annotation and serialization into AIM format, CIP has also funded the effort to support AIM in 3D Slicer software.

Discussion

The proposed system architecture for QIN data sharing and collaborative research is very ambitious and not likely to be achieved in the first 5 years of the program. In particular, the greatest challenge will likely remain system integration. It is expected however, that at a minimum, image sharing should be possible for the QIN sites using TCIA in the first 5 years. Furthermore, the QIN sites should at the very least assemble a concrete set of use cases to improve our understanding in the applicability of the existing formats for sharing image meta-data and identify possible limitations of the available information models for representing and storing image meta-data.

Ideally, these systems should not simply support the requirement of data sharing as an end by itself. They should support the QIN sites in the workflow of conducting their research. In preparation for the first set of QIN grant renewals in three years, important work is needed to clearly identify a path forward for each site to share all of their research data.

Conclusions

The first two years of the QIN have seen great progress in developing and implementing a strategy for data sharing across the network. Our initial efforts have resulted in a catalogue of the types of experiments and data that each member of the network is undertaking, a proposed systems architecture to support the research mission of the QIN, and an initial strategy and implementation of image data sharing. Work is ongoing to clarify the requirements and identify tools that can be integrated to support the QIN research mission.

Acknowledgments

Grant funding:

NIH/NCI PAR-11-150

NCI Contract No. HHSN261200800001E

caBIG Imaging Workspace contract No. 98411XSB2

REFERENCES

1. Boulton G, Rawlins M, Vallance P, Walport M. Science as a public enterprise: the case for open data. *Lancet*. 2011 May 14; 377(9778):1633–1635. [PubMed: 21571134]
2. Walport M, Brest P. Sharing research data to improve public health. *Lancet*. 2011 Feb 12; 377(9765):537–539. [PubMed: 21216456]
3. Formerly PAR-08-225, now PAR-11-150. Available from: <http://grants.nih.gov/grants/guide/pa-files/PAR-11-150.html>
4. Buckler AJ, Bresolin L, Dunnick NR, Sullivan DC, Aerts HJWL, Bendriem B, Bendtsen C, Boellaard R, Boone JM, Cole PE, et al. Quantitative imaging test approval and biomarker qualification: interrelated but distinct activities. *Radiology*. 2011 Jun; 259(3):875–884. [PubMed: 21325035]
5. Bogaerts J, Ford R, Sargent D, Schwartz LH, Rubinstein L, Lacombe D, Eisenhauer E, Verweij J, Therasse P. Individual patient data analysis to assess modifications to the RECIST criteria. *European Journal of Cancer (Oxford, England: 1990)*. 2009 Jan; 45(2):248–260.
6. Moskowitz CS, Jia X, Schwartz LH, Gönen M. A simulation study to evaluate the impact of the number of lesions measured on response assessment. *European Journal of Cancer (Oxford, England: 1990)*. 2009 Jan; 45(2):300–310.
7. ACRIN > HOME. [cited 2012 January 29] Available from: <http://www.acrin.org/>
8. Quality Assurance Review Center. [cited 2012 January 29] Available from: <http://www.qarc.org/>
9. In Vivo Imaging (IMAG) Workspace. [cited 2012 January 29] Available from: <https://cabig.nci.nih.gov/workspaces/Imaging>
10. National Biomedical Imaging Archive. [cited 2012 January 29] Available from: <https://cabig.nci.nih.gov/tools/NCIA>
11. Welcome to the caBIG® Community Website. [cited 2012 January 29] Available from: <https://cabig.nci.nih.gov/>
12. XNAT: Open source informatics for biomedical imaging research | XNAT.org. [cited 2012 January 29] Available from: <http://www.xnat.org/>
13. BIRN. [cited 2012 January 29] Biomedical Informatics Research Network | The Conduit for Biomedical Research. Available from: <http://www.birncommunity.org/>
14. DICOM Homepage. [cited 2012 January 29] Available from: <http://medical.nema.org/>
15. Teem: nrrd. [cited 2012 January 31] Available from: <http://teem.sourceforge.net/nrrd/>
16. RSNA.org: Clinical Trials Processor. [cited 2012 January 29] Available from: http://www.rsna.org/Informatics/Clinical_Trials_Processor.cfm
17. OsiriX - DICOM Viewer. [cited 2012 January 29] Available from: <http://www.osirix-viewer.com/>
18. 3D Slicer. [cited 2012 January 29] Available from: <http://www.slicer.org/>
19. PMOD Home Page. [cited 2012 January 29] Available from: <http://www.pmod.com/technologies/index.html>
20. TRIAD - Transfer of Images and Data. [cited 2012 January 29] Available from: <https://triad.acr.org/>
21. Dropbox - Simplify your life. [cited 2012 January 29] Available from: <http://www.dropbox.com/>
22. Open Source Clinical Image and Object Management. [cited 2012 January 29] Available from: <http://www.dcm4che.org/>
23. Rubin DL, Rodriguez C, Shah P, Beaulieu C. iPad: Semantic annotation and markup of radiological images. *AMIA ... Annual Symposium Proceedings / AMIA Symposium*. AMIA Symposium. 2008:626–630. [PubMed: 18999144]
24. Wang F, Pan T, Sharma A, Saltz J. Managing and Querying Image Annotation and Markup in XML. *Proceedings of SPIE*. 2010 Jan 1.7628(2010):762805. [PubMed: 21218167]
25. Korenblum D, Rubin D, Napel S, Rodriguez C, Beaulieu C. Managing biomedical image metadata for search and retrieval of similar images. *Journal of Digital Imaging: The Official Journal of the Society for Computer Applications in Radiology*. 2011 Aug; 24(4):739–748. [PubMed: 20844917]
26. Channin DS, Mongkolwat P, Kleper V, Rubin DL. The Annotation and Image Mark-up project. *Radiology*. 2009 Dec; 253(3):590–592. [PubMed: 19952021]

27. Definiens Health Image Intelligence - From Images to Clinical Decisions - Home. [cited 2012 January 29] Available from: <http://www.definiens.com/>
28. MIM Software - Home. [cited 2012 January 29] Available from: <http://www.mimsoftware.com/>
29. The Cancer Imaging Archive - Home. Available from: <http://cancerimagingarchive.net/>.
30. Clunie DA. DICOM structured reporting and cancer clinical trials results. *Cancer Informatics*. 2007; 4:33–56. [PubMed: 19390663]
31. OFFIS DCMTK: dcmsr: a structured report library and utility apps. Available from: http://support.dcmthk.org/docs/mod_dcmsr.html.
32. Annotation and Image Markup - Enterprise Service: AIM Template Builder (ATB) - NCI Tracker. [cited 2012 January 29] Available from: <https://tracker.nci.nih.gov/browse/AIMES/component/10392>
33. ClearCanvas: Community Plugins. [cited 2012 January 29] Available from: <http://www.clearcanvas.ca/dnn/Community/Plugins/tabid/73/Default.aspx>
34. Wang F, Pan T, Sharma A, Saltz J. Managing and Querying Image Annotation and Markup in XML. *Proceedings of SPIE*. 2010 Jan 1.7628(2010):762805. [PubMed: 21218167]
35. Cancer Data Standards Registry and Repository (caDSR). [cited 2012 January 30] Available from: <https://cabig.nci.nih.gov/concepts/caDSR/>
36. Cancer Central Clinical Database (C3D). [cited 2012 January 30] Available from: <https://cabig.nci.nih.gov/tools/c3d>
37. AIM API - QILWiki. [cited 2012 January 30] Available from: http://www.stanford.edu/group/qil/cgi-bin/mediawiki/index.php/AIM_API
38. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, Zegers CM, Gillies R, Boellard R, Dekker A, Aerts HJ. Radiomics: extracting more information from medical images using advanced feature analysis.

Figure 1.
QIN Proposed System Architecture to enable data sharing, collaborative experimentation and translation of methods to clinical practice. eCRF = electronic case report form.

Table 1

Alphabetical listing of commonly used acronyms

Acronym	Acronym Meaning
ACR	America College of Radiology
ACRIN	American College of Radiology Imaging Network
AIM	Annotation Image Markup
AIME	AIM data service
API	Application Programming Interface
BIMM	Biomedical Image Metadata Manager
BIRN	Biomedical Informatics Research Initiative
C3D	Cancer Central Clinical Database
caBIG	cancer Biomedical Informatics Grid
caDSR	Cancer Data Standards Registry and Repository
CIP	Cancer Imaging Program
CT	Computed Tomography
CTP	Clinical Trial Processor
DCE MRI	Dynamic Contrast Enhanced Magnetic Resonance Imaging
DCMTK	DICOM Tool Kit
DICOM	Digital Imaging and Communications in Medicine
DICOM-SR	DICOM Structured Reporting
DICOM-RT	DICOM Radiotherapy
eCRF	Electronic Case Report Form
ePAD	Electronic Physician Annotation Device
FDG PET	Fluorodeoxyglucose Positron Emission Tomography
FLT PET	Fluoro-L-thymidinePositron Emission Tomography
Fmiso PET	18F-fluoromisonidazole positron emission tomography
FES PET	18F-fluoroestradiol PET
IDL	Interactive Data Language
MRI	Magnetic Resonance Imaging
NBIA	National Biomedical Image Archive
NCI	National Cancer Institute
NRRD	Nearly Raw Raster Data
PACS	Picture Archiving and Communication System
PAR/REC	Philips Research File Format
PET	Positron Emission Tomography
PSA	Prostate Specific Antigen
RECIST	Response Evaluation Criteria in Solid Tumors
QARC	Quality Assurance Review Center
QIN	Quantitative Imaging Network
SUV	Standardized Uptake Value

Acronym	Acronym Meaning
TCIA	The Cancer Imaging Archive
TRIAD	Transfer of Images and Data System
XNAT	Extensible Neuroimaging Archive Toolkit

Table 2

Survey Questions

Category of Question	Questions
Disease and treatment	What type of cancer diagnosis? What type of cancer treatments? What line of cancer treatment?
Image data	What imaging modalities? What image file formats? What systems in use for storing images locally? What tools for image de-identification?
Image meta-data	What tumor image features? What image meta-data acquisition methods and tools? What local image meta-data database (and respective image) meta-data information model)?
Clinical data	What classes of clinical data? What local systems to store clinical data?

Table 3

QIN Survey Results: Disease and treatment

QIN matrix survey question	Summarized results across the network
Cancer diagnosis	<ul style="list-style-type: none"> • Breast Cancer • Colon Cancer • GlioblastomaMultiforme • Head and Neck Cancer • Hepatocellular Carcinoma • Lung Cancer • Lymphoma • Prostate Cancer • Sarcoma
Treatment type	<ul style="list-style-type: none"> • Systemic drug therapy • Radiation therapy
Line of therapy	<ul style="list-style-type: none"> • Primary • Neoadjuvant • Metastatic

Table 4

QIN survey results: image data

QIN matrix survey question	Summarized results across the network
Image file formats	<ul style="list-style-type: none"> Digital Imaging and Communications in Medicine (DICOM) [14] Philips Research File Format (PAR/REC) Nearly Raw Raster Data (NRRD) [15]
Image Modalities	<ul style="list-style-type: none"> CT – Perfusion CT MRI – Dynamic Contrast Enhanced (DCE) MRI, Perfusion MRI, MR Spectroscopy, MR Sodium Imaging PET – fluorodeoxyglucose (FDG) (static and dynamic), fluoro-L-thymidine (FLT) (static and dynamic), 18F-fluoromisonidazole (Fmiso) (static), other dynamic (Fluoride, Acetate, thymidine, verapamil, CO₂, Water, Ammonia, 18F-fluoroestradiol (FES))
Image de-identification	<ul style="list-style-type: none"> RSNA Clinical Trial Processor (CTP) [16] XNAT remap[12] Osirix[17] 3D Slicer[18] PMOD[19] American College of Radiology (ACR) Transfer of Images and Data TRIAD[20] System In-house
Local image file storage	<ul style="list-style-type: none"> Local file system Dropbox[21] Manufacturer PACS DCM4CHEE[22] NBIA[10] XNAT[12]

Table 5

QIN survey results: image meta-data and image meta-data acquisition tools

QIN matrix survey question	Summarized results across the network
Tumor image features	<ul style="list-style-type: none"> • Linear dimension • Volume • Standardized Uptake Value (SUV) • Perfusion kinetic parameters • Transport • Metabolic flux • Hypoxic volume • Total metabolic index • Chemical shift imaging • Radiomics[38]
Image meta-data acquisition tools	<ul style="list-style-type: none"> • Electronic Physician Annotation Device (ePAD)[23] • 3DSlicer[18] • PMOD[19] • Matlab • Interactive Data Language (IDL) • Definiens[27] • MimVista[28] • GE Cinetool • Locally developed • Radiotherapy Treatment Planning System (various vendors)
Local image meta-data database (<i>and respective image meta-data information model</i>)	<ul style="list-style-type: none"> • Spreadsheet (<i>none</i>) • AIM Data Service (AIME) [24](<i>AIM</i>)[26] • Biomedical Image Metadata Manager (BIMM)[25](<i>AIM</i>)[26] • DICOM Radiotherapy (DICOM-RT) • 3D Slicer (<i>proprietary</i>) [18] • Filemaker Pro (<i>none</i>) • Internal database (<i>none</i>)

Table 6

QIN survey results: clinical data

QIN matrix survey question	Summarized results across the network
Classes of clinical data	<ul style="list-style-type: none"> • Basic patient demographics (age, gender) • Tumor tissue features (histology, tissue biomarkers including genomics) • Serum biochemical tumor markers • Cancer Stage • Social history (smoking, alcohol) • Treatment history (type of treatment, drugs, start and stop dates) • Disease response to treatment • Vital status (living, dead)
Local systems to store clinical data	<ul style="list-style-type: none"> • Spreadsheets • RedCap • Filemaker Pro • Internal Databases