dcmqi: an open source library for standardized communication of quantitative image analysis results using DICOM

Christian Herz[1,2], Jean-Christophe Fillion-Robin[3], Michael Onken[4], Jörg Riesmeier[5], Andras Lasso[6], Csaba Pinter[6], Gabor Fichtinger[6], Steve Pieper[7], David Clunie[8], Ron Kikinis[1,2,9,10], Andriy Fedorov[1,2]

[1] Department of Radiology, Brigham and Women's Hospital, Boston, Massachusetts, United States of America, [2] Harvard Medical School, Harvard University, Boston, Massachusetts, United States of America, [3] Kitware Inc., Clifton Park, New York, United States of America, [4] Open Connections GmbH, Oldenburg, Germany, [5] Freelancer, Oldenburg, Germany, [6] Laboratory for Percutaneous Surgery, School of Computing, Queen's University, Kingston, Ontario, Canada, [7] Isomics, Inc., Cambridge, Massachusetts, United States of America, [8] PixelMed Publishing, LLC, Bangor, Pennsylvania, United States of America, [9] Department of Computer Science, University of Bremen, Bremen, Germany, [10] Fraunhofer MEVIS, Bremen, Germany

Contact: Andriy Fedorov, Brigham and Women's Hospital, Harvard Medical School, 75 Francis St, Boston, MA 02215. Office: 617-525-6258. Email: andrey.fedorov@gmail.com.

Abstract

Quantitative analysis of clinical image data is an active area of research that holds promise for precision medicine, early assessment of treatment response, and objective characterization of the disease. Interoperability, data sharing, and the ability to mine the resulting data are of increasing importance, given the explosive growth in the number of quantitative analysis methods being proposed. The Digital Imaging and Communications in Medicine (DICOM) standard is widely adopted for image and metadata in radiology. *dcmqi* (DICOM for Quantitative Imaging) is a free, open source library that implements conversion of the data stored in commonly used research formats into the standard DICOM representation. *dcmqi* source code is distributed under BSD-style license. It is freely available as a precompiled binary package for every major operating system, as a Docker image, and as an extension to *3D Slicer*. Installation and usage instructions are provided in the GitHub repository at https://github.com/qiicr/dcmqi.

Introduction

Quantitative analysis of imaging data in clinical setting holds the promise for objective, non-invasive and reproducible characterization of disease. Development of the algorithms to enable such analyses has been an active area of clinical research for decades, leading to a proliferation of the analysis tools and public datasets. In parallel, there is increasing push for improving reproducibility of scientific research and sharing of both the original image data and the results of analysis in a machine-readable and reusable format (1). Meaningful data sharing and interoperability are hampered by the use of ad-hoc data formats in the research community.

We introduce a free open source library, *dcmqi*, that implements conversion of the results commonly produced in quantitative imaging research into the standard Digital Imaging and Communications in Medicine (DICOM) (2) format .

DICOM (2) is the most widely used standard for clinical image data and metadata. DICOM has been adopted by most equipment manufacturers in the field of radiology. DICOM enables interoperability of the various processing, archival, and visualization components, both clinical and research. DICOM provides a framework for communicating not only images, but also imaging-related information, accompanied by the structured and often coded "composite context" attributes such as those identifying the patient, and describing acquisition and provenance of the data. Composite context enables indexing and cross-referencing of the data. DICOM applications are not restricted to clinical use only. The standard defines objects for storing image and image-related data, which are also suitable to support imaging research, including that done using pre-clinical and animal imaging.

Adoption of DICOM for research applications has been limited for a number of reasons relating to the overall complexity of the standard, lack of understanding of its capabilities, and the significant effort required for the DICOM software maintenance. Existing DICOM toolkits reduce implementation effort, but often lack specific higher level semantic abstractions, structures and concepts relevant in quantitative imaging. *dcmqi* aims to lower some of the barriers towards adoption of the DICOM standard for the purposes of quantitative imaging research. This manuscript provides a high-level overview of *dcmqi*. The details are available in the *dcmqi* user guide.

Materials and Methods

dcmqi is a collection of software libraries and command line tools for converting data between the commonly used research formats and their DICOM representation counterparts. *dcmqi* is the result of generalization of the conversion tools we developed in an earlier project that demonstrated application of DICOM for imaging-based treatment response assessment in nuclear medicine oncological application (3).

We chose C++ as the primary language for implementing *dcmqi* in order to integrate with the existing open source libraries and tools. *dcmqi* builds upon DICOM Toolkit (DCMTK) (4) and Insight Toolkit (ITK) (5). DCMTK provides a general purpose application programming interface (API) for reading, creating and writing DICOM objects. ITK is commonly used for developing image segmentation and registration tools, and provides support for popular research data formats. The decision to create a new library instead of implementing the new functionality within DCMTK was motivated by the need of using ITK and the wish to keep this dependence compartmentalized and well-defined.

Conversion from research formats into DICOM relies on metadata necessary to create a valid and meaningful DICOM object. *dcmqi* aims to reduce the user burden in providing such metadata, e.g., by automatically propagating attributes of the composite context that are available in the reference DICOM data specified by the user (e.g., patient and study information), and initializing references to the source objects used in the derivation. Inevitably, some of the metadata is not present in the source data, and must be populated by the user. These attributes are specified using JavaScript Object Notation (JSON), formalized by a JSON Schema (6) defined for the individual data types we support. This approach results in documents that are both human- and machine-readable, and can be automatically validated. *dcmqi* uses *ajv* (7) for validating the conformance of JSON documents with the schema. This validation is done as part of the continuous integration (CI) testing, and can be conveniently accessed via the web-based validation tools that accompany *dcmqi*. Reverse conversion from DICOM into research formats is also supported to enable integration with the tools that are not DICOM-aware.

Results

This manuscript accompanies the initial release of *dcmqi*. The source code, available at https://github.com/qiicr/dcmqi, has been in development since January 2016, and is distributed under a non-restrictive BSD-style license. Installation and usage instructions are available within the repository. CI is implemented for Windows, Linux and macOS to exercise the build process, functionality of the individual converters, and to check the validity and compliance with the schema for the JSON components of the library. Compliance of the resulting objects with the standard is verified using *dciodvfy* from *dicom3tools* (8). *dcmqi* converters can be built from source, are available as a ready-to-use Docker image or via a visual user interface within 3D Slicer (9). OS-specific installers of *dcmqi* are provided for Windows, Linux and macOS. Introductory tutorial is available at http://qiicr.org/dcmqi-guide/tutorials/intro.html. The process of setting up the environment and installing *dcmqi* Docker image is explained in Video 1.
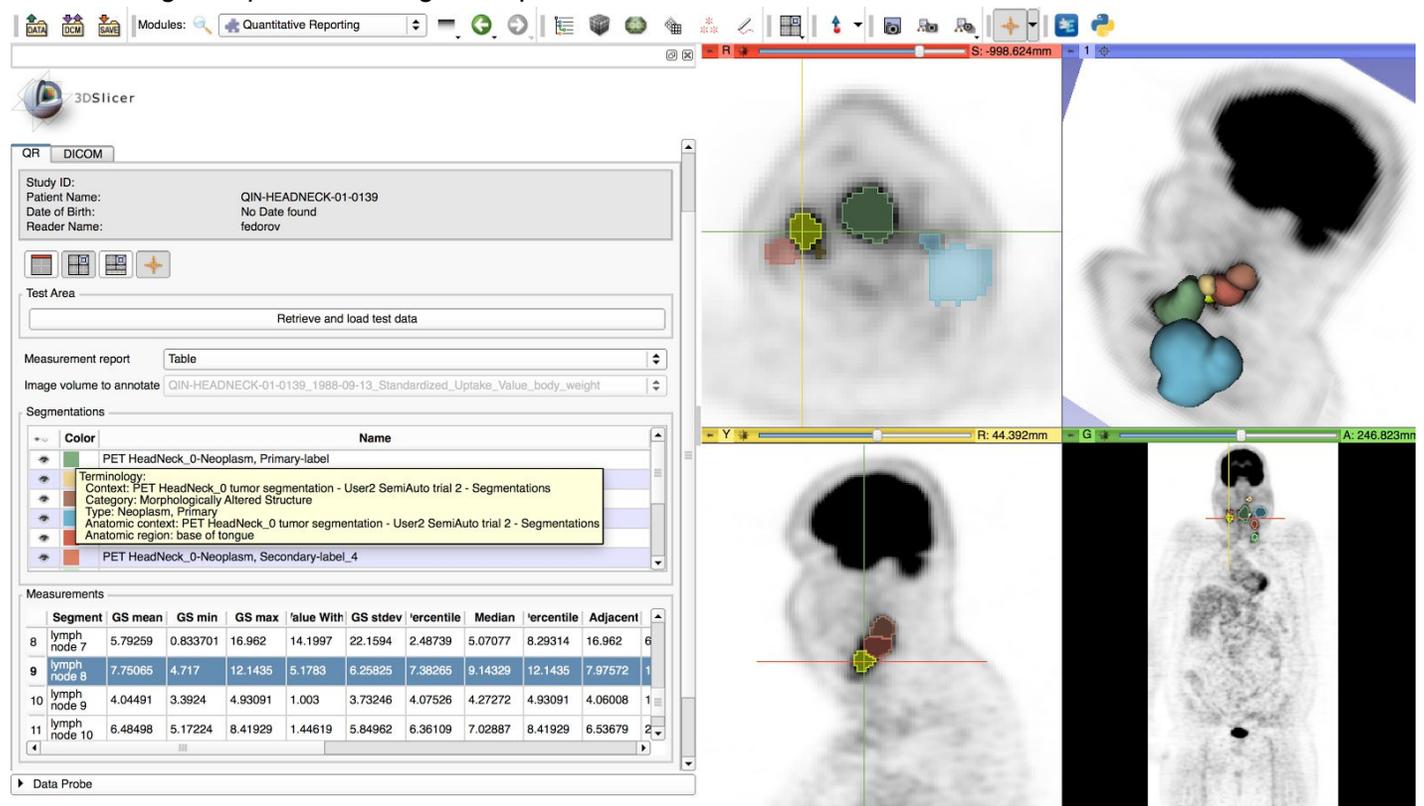
Figure 1: Result of loading post-processing results stored in DICOM for subject 139 from the TCIA QIN-HEADNECK collection (3) using QuantitativeReporting extension of 3D Slicer, which is using dcmqi for DICOM data conversion. The interactive visualization includes multiplanar reformats of the original PET image data with overlay of segmented areas, 3-dimensional rendering (upper right), coded terminology of the annotated regions (yellow tooltip rectangle), and a table of the image-based measurements over the segmented areas. All of these data components were restored from DICOM objects retrieved from TCIA.

As of Spring 2017, *dcmqi* supports the following types of quantitative image analysis results:

1. **Rasterized image segmentations** defining image regions of interest as DICOM Segmentation image objects. Examples of segmentation tasks include generation of atlases of normal anatomy, or definition of the neoplastic areas for subsequent volumetric evaluation and other quantitative characterization tasks, including radiomics analysis (10). Metadata includes coded semantics of the segmented region, anatomical location of the finding, type of the algorithm used, recommended color for the visualization of the segmentation. Video 2 demonstrates the process of converting segmentation result generated using *ITK-Snap* (11), a popular image segmentation tool, into a DICOM Segmentation image object.

2. **Parametric maps** as DICOM Parametric Map objects: result of post-processing of image data. As an example, Apparent Diffusion Coefficient (ADC) maps derived by fitting various models to Diffusion-Weighted Magnetic Resonance Imaging (DW-MRI) data have shown promise in characterizing aggressiveness of prostate cancer (12). Resulting objects contain coded quantity and units for the pixel-wise values, and the method used to derive those values.

3. **Measurements** calculated over the volumetric regions defined by rasterized image segmentations, as DICOM Structured Reporting (SR) Template ID 1500 (SR TID 1500) objects. Measurements can range from simple calculation of region volume to more complex quantitative measures, such as image textures (13). The resulting SR objects contain coded quantity, units, and reference to the segmentation object used for measurements calculation. The process of loading and visualizing segmentation objects and the associated measurements from a publicly available TCIA dataset (3) using 3D Slicer platform (9) is illustrated in Video 3.

Functionality of *dcmqi* is exposed via graphical user interface in the *QuantitativeReporting* extension of 3D Slicer. *QuantitativeReporting* supports the interactive image annotation and quantitative analysis workflow that allows an operator to load a DICOM image series, use automated segmentation tools of 3D Slicer to define the volume of interest, automatically calculate the volume of interest based measurements, and save the result of this workflow as a cross-referenced collection of DICOM objects. *QuantitativeReporting* can also be used to reload the volumetric measurements stored using SR TID 1500 together with the referenced image and segmentation data (see Fig.1 and Video 3).

We are participating in community projects to raise awareness of the importance of interoperable communication of derived imaging information, and the relevant capabilities of the DICOM standard and supporting tools. DICOM for Quantitative Imaging (DICOM4QI) is a "connectathon"-style initiative that we organized within Radiological Society of North America (RSNA) annual convention (14). DICOM4QI gathers developers of both academic and commercial platforms to test DICOM-based interoperability. In 2016, a number of both open source and commercial platforms chose *dcmqi* to implement DICOM research data storage support (14). Within the NCI Quantitative Imaging Network (QIN), we apply *dcmqi* for converting the analysis results produced by the QIN algorithm challenges and community projects into DICOM representation.

Discussion

The goal of *dcmqi* is to simplify the use of the DICOM standard for the purposes of quantitative imaging research in order to support standardized collection, archival and sharing of research results. Previous related efforts aimed at providing more general-purpose DICOM conversion and editing tools (e.g., *dccp* from *dicom3tools*, *dcmodify* from DCMTK, and RSNA Clinical Trials Processor). We limit the scope of *dcmqi* specifically to the needs and data types relevant in quantitative imaging research, and focus on the task of use case driven conversion of the data.

A key challenge in maintaining *dcmqi* converters is that the standard itself does not provide machine-readable, formalized descriptions of the DICOM objects. The metadata schemas we introduced formalize *dcmqi* conversion process, but do not aim to re-define the standard. While those schemas will evolve in the future, the output objects will comply with the same DICOM standard. Our future plans include improved flexibility and harmonization of *dcmqi* parameter schema with the standard and the standard-defined JSON model of DICOM (15). The latter was not used directly due to our prioritization for readability of the metadata JSON.

The DICOM standard is evolving alongside *dcmqi*. We work with the DICOM community to resolve gaps, limitations or inconsistencies that we identify in the process of implementation, as discussed in (3). We expect further refinements of both the library and the standard as we expand the collection of use cases and applications of *dcmqi* in quantitative imaging research.

References

1. Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding

Principles for scientific data management and stewardship. Sci Data [Internet]. 2016;3:160018. Available from: http://dx.doi.org/10.1038/sdata.2016.18

2. Bidgood WD Jr, Horii SC, Prior FW, Van Syckle DE. Understanding and using DICOM, the data interchange standard for biomedical imaging. J Am Med Inform Assoc [Internet]. 1997;4:199–212. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9147339

3. Fedorov A, Clunie D, Ulrich E, Bauer C, Wahle A, Brown B, et al. DICOM for quantitative imaging biomarker development: a standards based approach to sharing clinical data and structured PET/CT analysis results in head and neck cancer research. PeerJ [Internet]. 2016;4:e2057. Available from: http://dx.doi.org/10.7717/peerj.2057

4. Eichelberg M, Riesmeier J, Wilkens T, Hewett AJ, Barth A, Jensch P. Ten years of medical imaging standardization and prototypical implementation: the DICOM standard and the OFFIS DICOM toolkit (DCMTK). Proc SPIE 5371, Medical Imaging 2004: PACS and Imaging Informatics [Internet]. International Society for Optics and Photonics; 2004 [cited 2015 Nov 24]. page 57–68. Available from: http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=841367

5. Ibanez L, Schroeder WJ. The ITK Software Guide 2.4. Kitware, Inc.; 2005. page 804.

6. JSON Schema [Internet]. 2016 [cited 2017 Jan 26]. Available from: http://json-schema.org/documentation.html

7. Poberezkin E. Ajv - Another JSON Schema Validator [Internet]. 2017 [cited 2017 Jan 23]. Available from: http://epoberezkin.github.io/ajv/

8. Clunie D. Dicom3tools Software [Internet]. David Clunie's Medical Image Format Site. 2009 [cited 2015 Nov 22]. Available from: http://www.dclunie.com/dicom3tools.html

9. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-CC, Pujol S, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging [Internet]. 2012;30:1323–41. Available from: http://dx.doi.org/10.1016/j.mri.2012.05.001

10. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer [Internet]. 2012;48:441–6. Available from: http://dx.doi.org/10.1016/j.ejca.2011.11.036

11. Yushkevich P a., Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage [Internet]. 2006;31:1116–28. Available from: http://dx.doi.org/10.1016/j.neuroimage.2006.01.015

12. Hambrock T, Somford DM, Huisman HJ, van Oort IM, Witjes JA, Hulsbergen-van de Kaa CA, et al. Relationship between apparent diffusion coefficients at 3.0-T MR imaging and Gleason grade in peripheral zone prostate cancer. Radiology [Internet]. 2011;259:453–61. Available from: http://dx.doi.org/10.1148/radiol.11091409

13. Aerts HJW, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun [Internet]. Nature Publishing Group; 2014 [cited 2014 Sep 9];5. Available from: http://dx.doi.org/10.1038/ncomms5006

14. Fedorov A. DICOM4QI: DICOM for Quantitative Imaging connectathon [Internet]. gitbook. 2017 [cited 2017 May 19]. Available from: https://qiicr.gitbooks.io/dicom4qi

15. National Electrical Manufacturers Association (NEMA). DICOM JSON Model. DICOM PS318 - Web Services [Internet]. 2016 [cited 2016 Oct 29]. Available from: http://dicom.nema.org/medical/dicom/current/output/chtml/part18/chapter_F.html