

Scalable Quality Assurance for Neuroimaging (SQAN): modern web based quality control for medical imaging - project overview

Gopu, A., Young, M. D., Avena-Koenigsberger, A., Hayashi, S., West, J.D., Paramasivam, M., Perigo, R.

Abstract

Medical imaging, a key component in clinical diagnosis and research of numerous medical conditions, is very costly and can generate massive datasets. Many projects lack funds to reacquire images if data quality issues are detected later. Data quality assurance (QA) requires continuous involvement by all stakeholders and use of specific quality control (QC) methods to rapidly identify data issues likely to require to post-processing correction or real-time re-acquisition. While many useful QC methods exist, they are often designed for specific use-cases with limited scope and documentation, making integration with other setups difficult. Here, we describe the Scalable Quality Assurance for Neuroimaging (SQAN - pronounced “scan”), an open-source software suite for protocol quality control and instrumental validation on medical imaging data, developed by Indiana University. SQAN runs a comprehensive QC pipeline, ensuring adherence to a research study’s protocol. The project involved significant engagement with researchers, scanner technologists, and data scientists, each of whom approach QC with their own unique priorities, expertise, insights, and web portal expectations. Since Fall 2017, a fully operational production SQAN service instance has supported 50+ research projects having QC’d ~3 million images and over 600 million metadata tags. SQAN is designed to 1) scale to any imaging center’s QC needs and 2) extend beyond protocol QC to include image level QC and integration into image pipeline and non-imaging database systems. This paper focuses on a functional overview and description of the web portal; the second paper in this series focuses on the data management and QC components of SQAN.

Background and Motivation

Medical imaging is integral to the clinical diagnosis and research of neurodegenerative diseases like Alzheimer’s and Parkinson’s, as well as brain cancers [1,2,3]. Medical imaging techniques -- including Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and Positron Emission Tomography (PET) -- generate massive datasets in the order of tens to even hundreds of thousands of images for a single subject, and studies include hundreds or even thousands of subjects [4,5]. Imaging is costly too, and many projects lack funds to reacquire images if QC issues are detected after the acquisition period.

Quality assurance (QA) is a fundamental first step in guaranteeing reliable and reproducible medical imaging research [6,7,8]. Importantly, QA requires continuous involvement by all stakeholders, and the use of appropriate quality control (QC) methods that rapidly identify data in need of post-processing correction or reacquisition. Statistical inaccuracies from incorrect imaging parameters, low image quality, scanner software updates [9], and motion artifacts [10] can all contribute to noise in data, leading to

unreliable and irreproducible results [6,7]. Further, because imaging data are often made available to the general scientific community, results may also impact downstream research [5].

While many useful QC methods exist, they are often limited in scope or require time-consuming manual techniques [8,11,12]. Existing automated or semi-automated QC procedures are of complex and lacking in documentation; they are generally designed for specific use-cases (i.e. scanner-specific, image-specific), making integration with other setups difficult. Automated and flexible QC on imaging protocols would supplement existing imaging QA/QC methods, and are particularly critical for multi-center projects with heterogeneous data. [11,12,13].

Origins of SQAN: Project team background

In mid-2015, the Indiana University Scalable Compute Archive (IU/SCA) team and the Indianapolis based RADY Imaging Center began discussing RADY's quality control needs for its datasets collected during research studies and clinical trials. We identified having a modern web portal user interface as an important factor in making such a system viable for use by RADY as well as for potential future adoption by other medical centers, irrespective of the computer skills (or lack thereof) of the user base. We determined that designing and developing an automated-yet-flexible, managed QC system for imaging protocols required a wide array of skills and experience spanning imaging center operational expertise, computer science and software development skills, as well as neuroscience domain expertise.

The RADY Imaging Center has been offering imaging services for over 65 research studies (and several clinical trials) for more than 15 years. The Center is equipped with multiple Siemens scanners spanning MRI, PET, and CT imaging modalities, and a staff of 6 professional technologists. The SCA - both the name of a software suite and the team that manages it - builds, delivers, and operates customized web user interfaces, secure data management systems, and integrated scientific software application pipelines. SCA systems - supporting astronomy [23a-d], IU's High Performance Computing (HPC) community at large [24], neuroscience [25a-c], dynamic image visualization [26], electron microscopy (IU EMCenter portal operational 2013-18, retired in 2019) - are securely accessible from any web-enabled device at any time. Several SCA projects - some already open source - have openly-available, published papers [27]. The SCA team members have backgrounds spanning a broad spectrum (including computer science, neuroscience, astronomy, informatics, and Linux system administration), with 70+ person years of experience enabling scientific research at IU. It became clear that the two teams possessed the skills necessary to successfully collaborate on a project to design and develop an automated, flexible, and managed QC system for imaging protocols.

After having invested ~3000 person hours on software development and ~1500 person hours on collaborative discussions, we now present the open-source Scalable Quality Assurance for Neuroimaging (SQAN - pronounced "scan") software suite including a modern and user-friendly web portal for protocol quality control and instrumental validation on medical imaging data. In the rest of this paper we provide a high level summary of the SQAN service stack, followed by a detailed description of the web portal and how it serves and satisfies the usability needs of various types of users. We wrap up with a list of immediate and long-term plans.

SQAN overview

SQAN runs a comprehensive QC pipeline, comparing various facets of each scan for each subject to a template. It ensures that all scans required across each modality (e.g., MRI, PET, CT) by a study's pre-defined protocol are present with the expected image counts, and that values match exactly, or within a percentage threshold of the template value. SQAN safely ignores keywords expected to differ between the template and subject exam (e.g., timestamps, subject demographics). Determining which keywords should be checked, and the checks' types and thresholds, requires significant engagement from and feedback cycles with researchers, scanner technologists, and data scientists. We developed extensive measures to address the edge-cases and quirks of the DICOM imaging standard as implemented by multiple manufacturers and utilized by researchers with a variety of goals.

The SQAN software suite includes several components: 1) an *Orthanc* DICOM image receiver; 2) an *incoming* application that pulls metadata tag (key-value pairs) off the Orthanc receiver and stores them into the database; 3) a comprehensive *QC process* that compares every image in a subject's scan to a specified template or other predetermined QC criteria, and stores the results in the database; 4) a modern web portal that provides the single point of access to data, templates, and QC results for a wide user base spanning faculty members who lead research studies, imaging center staff, and research associates; 5) an Application Programming Interface (API) that liaises between the web portal and the other backend components including the database. SQAN, built on the foundation of open-source software and platforms (GitHub, node.js, AngularJS, MongoDB, docker), can benefit any medical imaging center. A more detailed description of SQAN's data management and QC processes are available in the second paper of this series [17].

Since Fall of 2017, a fully operational production SQAN service instance has supported our RADY colleagues, providing QC for 65+ research projects; it has QC'ed 3 million images and validated 600 million metadata tags spanning MRI, CT, and PET modalities. We have begun the process of transitioning the project from an IU-internal project to a self-sustaining, open-source project (available at <https://github.com/IUSCA/SQAN>) with the longer term intention of bringing together a vibrant community of developers and maintainers; we expect two more imaging centers - Dartmouth College and Harvard University - to adopt SQAN for their protocol compliance needs. A demo portal is available at <https://sqan.sca.iu.edu>.

Usability (UX) considerations

Many existing QC solutions (e.g., XNAT's protocol check pipeline [17, 3]) have QC algorithms that could be useful for other imaging centers. However, they typically require time-consuming manual steps to setup, and offer limited documentation; they are also generally designed for specific use-cases (i.e., scanner-specific, image-specific), making integration with other setups difficult. Even after they are set up, users must request changes from administrators for actions like adding or removing QC check fields; adding, removing, or updating of templates; and updating parameter thresholds on existing checks.

SQAN's intuitive web portal interface offers a significant usability improvement over many existing QC solutions, and empowers users—even those who are not very technically-minded—to customize QC criteria themselves through an intuitive graphical user interface.

Our efforts into the design and development of SQAN included significant engagement with researchers, scanner technologists, and data scientists at Indiana University, each of whom approach QC with their own unique priorities, expertise, insights, and usability expectations that best address their particular goals. One of the guiding principles behind SWAN's design is to be flexible enough to enable both researchers and technologists to perform adequate QA assessments.

General protocol adherence is essential to both groups; thus, it is the foremost goal of the system. SQAN allows technologists to track down issues with the scanner that might otherwise go unnoticed. In this way, technologists can be proactive in rescheduling scans for upcoming studies and avoid having to repeat data acquisitions; this can save the imaging center thousands of dollars. SQAN can also save researchers and staff hundreds of hours of data processing by identifying data that needs to be corrected or reacquired before it is included in any analysis; importantly, it ensures that the data being analyzed (and possibly published) is accurate and reliable. Researchers can also potentially use the data captured by the system to correct data during processing, and increase the power of their statistical analysis. For example, when motion parameters are recorded during a scan, these can be used to identify and correct motion artifacts that would not be flagged by the imaging protocol QC process. The system's research summary feature, important to both researchers and technologists, allows users to easily identify problematic series or subjects that are not conducive to imaging research. Identifying these situations can allow researchers to better allocate resources in the future.

Web portal design considerations

The SQAN portal is written using the AngularJS client-side user interface framework, and is coupled with an Application Programming Interface (API) also written in javascript as a node.js express app. We use the bootstrap and font-awesome javascript libraries to maintain look and feel across the portal.

Note: We expect to add more detail to this section in the final paper. We also expect to migrate the front-end portal code to vue.js before we publish the final paper at SPIE2020, pending funding availability and other constraints.

Web portal features

SQAN's intuitive web portal interface is designed to empower a wide spectrum of users. Our goal from the outset was to enable as many functions from within the portal as possible so users never had to contact an administrator for basic operations. Accurately and intuitively conveying the condition and status of large datasets is an ongoing challenge in the field of user interface (UI) and user experience (UX) design and development. In designing our web portal, we considered the hierarchical data structure across all medical imaging modalities. Typically, a Research (study) is defined by a group of subjects and an

imaging protocol. The imaging protocol mandates what acquisition series (i.e., a specific set of scanner parameter settings) are performed during an exam, in what order, and how many images are acquired in each of those series. A subject can undergo one or more exams within a given research; therefore, a subject's exam is expected to contain the acquisition series that are determined in the research imaging protocol.

Exams view

Given the hierarchical nature of our datasets, we took a similar approach in developing the web portal through which users (i.e., researchers, technologists, and administrators) would interact with the metadata and the QC statuses and warnings/errors. Determining at a fundamental level how our system would be used helped drive the design process. Researchers informed us that they were most concerned with the overall QC status of each subject's scan, so we designed the landing UI element to display the subject-level QC statuses, where subjects are grouped by research and the subject QC status are summarized by color-bars indicating the proportion of images identified with the different types of image-level QC states, such as *pass*, *fail*, *no template*, etc. (see Figure 1).

Additional UX considerations: A researcher can mouse-over an element to receive a more detailed statistical QC summary for that exam. Note here that with all statistics, there are many different ways of presenting the same data: consider a scan where each image contains 100 key-value pairs and a single value is found to differ from the corresponding template in every image. One can interpret this as a 100% error rate as all images contain an error; an alternative interpretation is that the overall error rate is 1%, as the error appears in one out of 100 key-value pairs. This simple example illustrates non-trivial decisions that must be made to display and summarize the data's QC status; importantly, these decisions have real consequences on how users respond to QC issues. We took a measured approach to displaying QC statistics with particularly careful consideration of word choice, coloration, and warning size and location. The exam-level QC listing also includes a responsive search bar and filtering options (date range, modalities to display, and sorting order for the research studies and the subject listing within each research study), which allows rapid drill-down to the desired dataset.

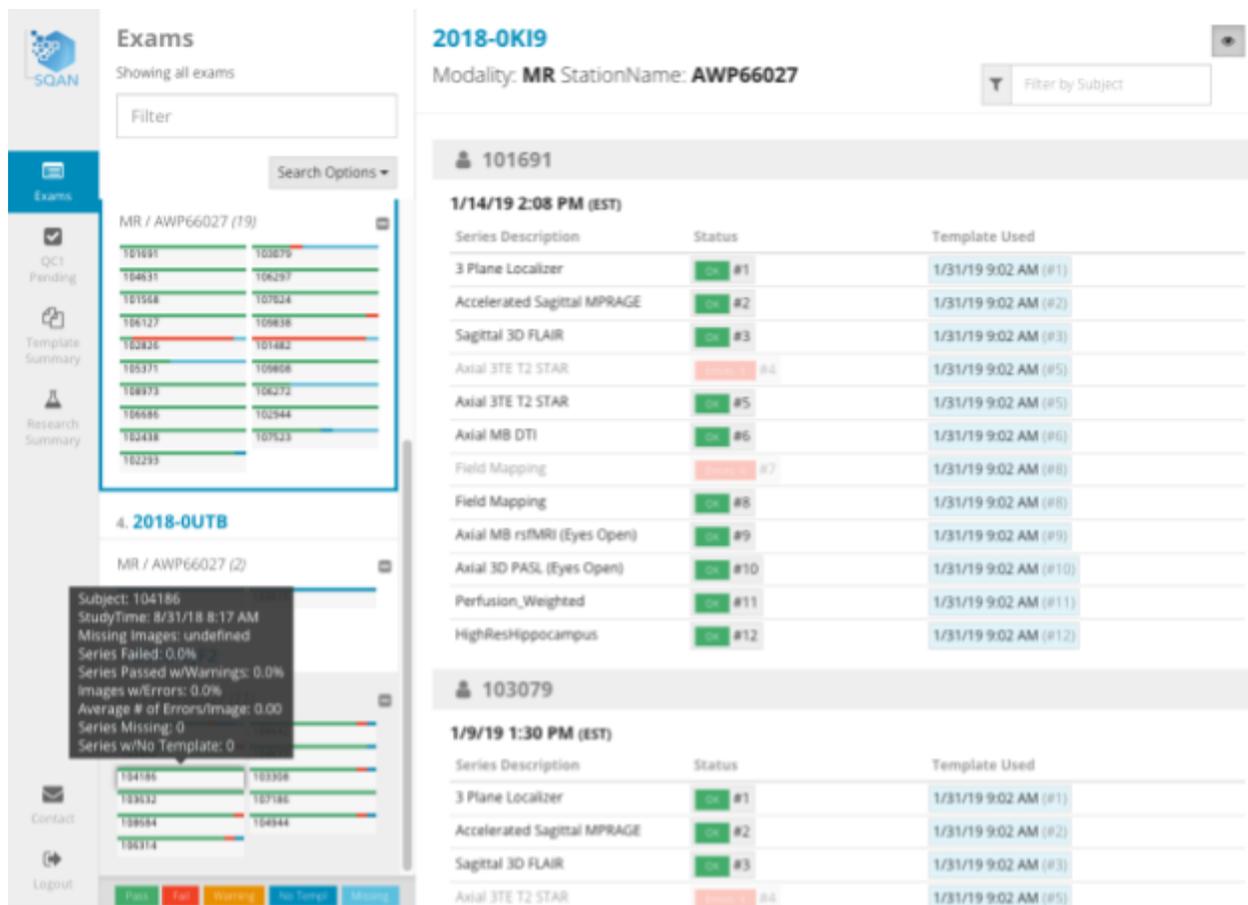


Figure 1. SQAN Exam View (left) and Exam Details (right)

QC pending view

The QC Pending view is a variant of the Exams view described above but only lists the research studies, and subjects/exams within them, that have QC issues reported and therefore require attention. While the Exams view is useful to review the QC status of an entire study or a series of subjects within one modality (e.g., MRI) within a study, the QC Pending view allows a technologist to review problematic datasets quickly and to contact responsible parties (e.g., the research study's PI) to solicit advice on how they should proceed.

Series detail view

In both the Exams and QC Pending views, users can click on an exam or research to see the list of series acquired in that exam, the series-level QC status, and the template that was used to perform QC. The details of each series can then be opened to see the image-level QC states for every image in that series

(Figure 2); finally, each image can be expanded to display the metadata key-value pairs and their QC states. At each QC-level, a summary of the QC status of the immediate lower level is clearly indicated.

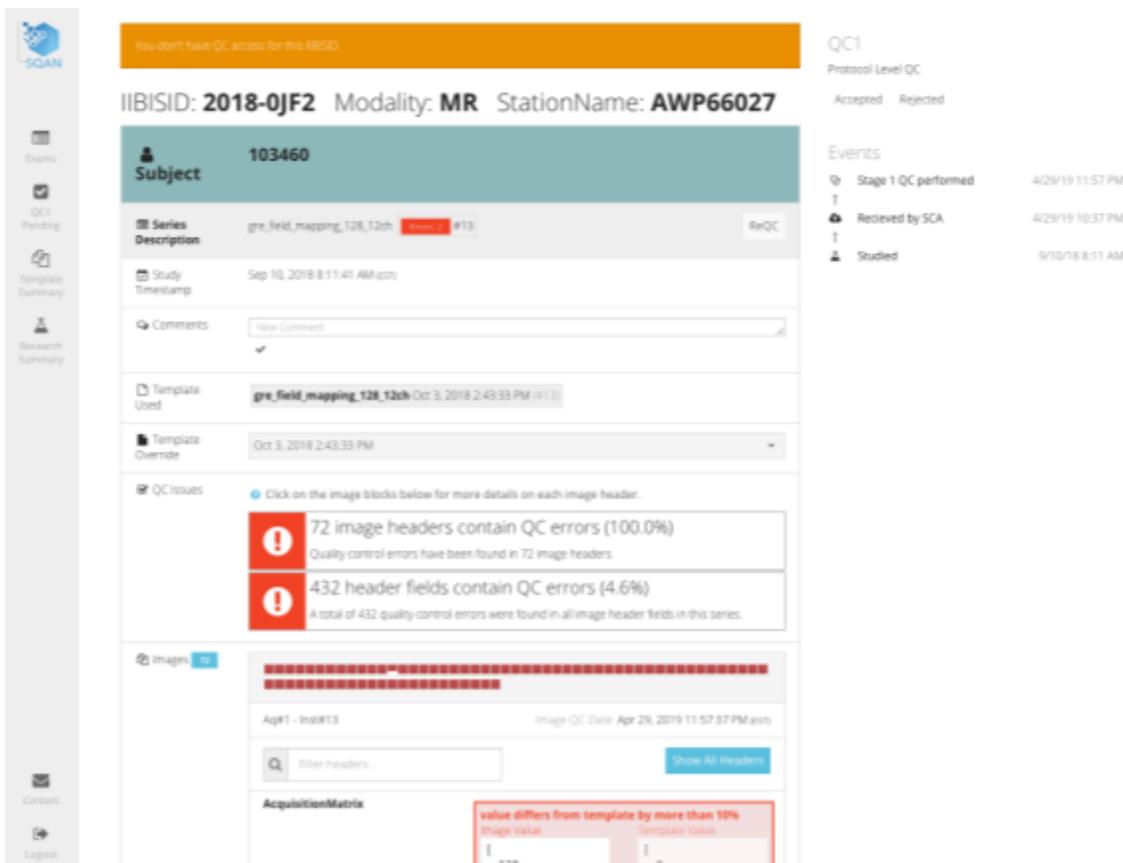


Figure 2. Series detail view, showing series and image QC errors

The automated QC state of an exam can be updated manually: a user may determine that a particular QC condition will not affect their research goals, and so a series may be marked as "passing" QC. Alternatively, if an updated set of templates has been received by SQAN, or if an older template is more appropriate, an exam may have its QC status recorded and then reset to undergo QC processing with another template. These procedures are all available to authorized users via the web interface. In addition, users are required to enter comments when manually updating a QC status; users may also alert/ask questions to responsible parties.

Additional UX considerations: Based on feedback from our users, we updated how the audit trail of events is displayed, especially if manual QC overrides are applied. We also updated the portal so QC issues are listed with color-coding and added QC-status-appropriate icons along with human-readable error messages, as shown in Figure 2. We use color-coding on the image listing and include a grey-colored box for missing images. The header listing (key-value pairs) only lists header tags with error or warning conditions by default, but can be expanded to show the entire header; a filter text box allows the user to navigate quickly to a particular header.

Research Summary view

The Exams and QC Pending views described above enable users to review the QC status of datasets in detail. We determined that researchers, in particular Principle Investigators (PIs) and their associates, often want to get a higher level summary of an entire research study in a small amount of user interface real estate. They want to be aware of one or more problematic subjects whose data they should exclude from their eventual analyses, or perhaps even of a particular acquisition series in which several subjects had trouble. In the Research Summary view, upon selection of a specific research study, we present a tabular listing of all subjects and acquisition series associated with that study, and a color coded status indicator for each subject and series. Clicking on a status indicator takes the user to the series detail view for that subject within that research.

Research Summary

Select Research

2018-0K19

Filter by

- subject...
- series description...

Download Summary

(.csv) (json)

2018-0K19 - A Phase II Study of Carboplatin and Patients with Advanced Recentralsity

PI: Alycia Casper (Alycia.Casper@iu.edu)

Dates: Apr 30, 2018 - Dec 30, 2021

Status: Anonymized

Study Type: Basic Science Human Research

Area of Study: AD

Modality/Lab: 3T MRI PRISMA (IIBIS at IUH Neuroscience), PET-CT (IIBIS at IUH Neuroscience Center)

Full Title: A Phase II Study of Carboplatin and Patients [...]

Transpose Table

Subjects	StudyTimestamp	MR												
		Accelerated Sagittal MPRAGE	Sagittal 3D FLAIR	Axial 3TE T2 STAR	Axial MB DTI	Field Mapping	Axial MB rsfMRI (Eyes Open)	Axial 3D PASL (Eyes Open)	Perfusion Weighted	HighResHippocampus	3 Plane Localizer	QC Phantom MPRAGE		
101691	1/14/19 2:08 PM	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
103079	1/9/19 1:30 PM	Green	Green	Red	Grey	Grey	Grey	Grey	Grey	Green	Green	Green	Green	Green
104631	12/5/18 7:53 AM	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
106297	11/30/18 1:30 PM	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
101568	11/24/18 2:02 PM	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
107024	10/28/18 9:01 AM	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
106127	5/29/18 12:46 PM	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
109838	5/26/18 8:55 AM	Green	Green	Green	Red	Green	Green	Green	Green	Green	Green	Green	Green	Green
102826	5/25/18 2:56 PM	Red	Red	Red	Red	Red	Red	Red	Red	Green	Red	Red	Red	Red
101482	5/4/18 2:01 PM	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red
105371	4/29/18 11:18 AM	Green	Green	Green	Grey	Grey	Grey	Grey	Grey	Green	Green	Green	Green	Green
109808	4/29/18 10:49 AM	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
108973	4/28/18 11:08 AM	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green

Figure 3: Research summary view. Consolidated information geared toward researchers/PIs.

Additional UX considerations: Users are able to filter by subject or the name of the acquisition series (referred to as “series description”). We also provide the ability to download a CSV or JSON file of the data presented in the view. We determined that some researchers want the subject listing to flow from top to bottom, while others want it to flow left to right, so we included a “transpose” button to change the row/column ordering of the view. This flexibility is useful in situations where a large number of subjects or series descriptions extends the table beyond the browser display window.

Optional feature: We collated the Research Summary view with information from a research study database that allows contextualization of a particular study. We list the title of the research study, the PI’s name, start and end dates for data acquisition for that study, study type (e.g., basic science human research vs. animal cognition), and modalities.

Template summary view

SQAN currently allows QC criteria to be set via template datasets sent specifically by the imaging center. There are occasions when a technician has sent a template over to SQAN and wants to verify it was received properly; there are other cases when an error condition is identified in a previously registered template within SQAN. Our portal includes a Template Summary view for this reason (Figure 4). Every template ever received is displayed with the corresponding research study name, modality, station (scanner) name, radio tracer, and the number of templates (referred to as “# Study Instances”) available in the system for that research study. Multiple templates are usually based on subject data or phantom data scanned on different dates, and are listed under different tabs as shown in Figure 4. The user can view the details of a particular template, including how many times that template has been used to QC imaging datasets, and the image count within that template. A technologist with sufficient authorization can delete a template or a specific acquisition series within a template, e.g., if they intend to resend that template (to correct errors) or provide an alternative template for that research study.

Additional UX considerations: Users can filter by research study ID (also referred to as IIBIS ID by our RADY colleagues). They are also able to sort by any of the fields on the template summary table header (e.g., modality, radio tracer). Entries are demarcated by background color differences to make them easy to parse for the user.

As part of our project roadmap, we intend to add more ways to define the criteria for quality control, by including custom QC criteria specific to a research study, or via replication of a subject’s dataset deemed to be an ideal dataset.

Template Summary Table

IIBISID ↑	Modality	Station Name	Radio Tracer	# Study Instances	View
0000-05FN	PT	CT71271	AV-1451	2	<
0000-067I	CT	CT71271		6	<
0000-0EYX	PT	CT71271	Florbetapir	1	<
0000-0PHO	PT	CT71271	NEURACEQ	2	<
0000-0RP4	PT	CT71271	Florbetaben	1	<
0000-0X5K	MR	AWP66027		1	<
2009-0Y8S	MR	AWP66027		1	<
2010-042O	MR	AWP66027		2	<
2010-04K8	PT	CT71271	Raclopride	3	∨

2013-10-17 12:00:04
2017-07-12 11:58:04
2014-10-17 13:53:00

Template Information

- Timestamp: 2017-07-12 11:58:04
- Number of Series: 4
- Series used for QC: 4

Action Links

- 🗑 Delete All Series
- 🗑 Delete 0 Selected Series
- 🔒 Hide Series

Series Number	Series Description	Times used for QC	# Images	
5	FBP_RAC_dynamic_50min_200	30	5995	🗑
6	FBP_RAC_static_50min_200	29	109	🗑
7	FBP_RAC_dynamic_50min_200_NAC	5	5995	🗑
8	FBP_RAC_static_50min_200_NAC	5	5995	🗑

2010-06P5
CT
CT71271

3
<

Figure 4: Template summary view

Administrative views: User and access control management

SQAN allows authorized administrative users to control other user accounts, including the data access groups each user is part of and the Access Control List (ACL) that maps research studies to one or more user groups, authorizing them to view/comment and/or perform QC operations on that study.

Accounts and Access Control

Users Groups ACL Create User

Name	Email	Primary Role	Roles	Created	Last Login
John West	jdwest@iupui.edu	admin	user admin	6/28/19 11:57 AM	7/26/19 4:22 PM
Sundar Paramasivam	mparamas@iu.edu	admin	user admin	6/28/19 11:57 AM	6/28/19 11:57 AM
Andrea Avena-Koenigsberger	alavenak@indiana.edu	admin	user admin	6/28/19 11:57 AM	7/31/19 2:17 PM
Arvind Gopu	agopu@iu.edu	admin	user admin	6/28/19 11:57 AM	8/14/19 1:48 PM
Michael D. Young	youngmd@iu.edu	admin	user admin	6/28/19 11:57 AM	8/2/19 11:01 AM
Raymond Perigo	rperigo@iu.edu	admin	user admin	6/28/19 11:57 AM	6/28/19 11:57 AM

Users Groups ACL Create Group

Name	Description	Members
RADY-QC Developer Team	RADY-QC Developer Team	Arvind Gopu Robert Henschel Sundar Paramasivam John West Michael D. Young Andrea Avena-Koenigsberger
RADY_Researchers_Testing	Group created to allow a small group of researchers access to all projects to allow testing and vetting of the system.	Mario Dzemidzic Shannon Leigh Risacher Karmen Yoder Yu-Chien Wu John West

Users Groups ACL

IIBISID	Can View/Comment	Can QC
2019-005P	RADY-QC Developer Team (6 users) RADY_Technologists (9 users)	RADY-QC Developer Team (6 users) RADY_Technologists (9 users)

Figure 5: Administrative controls: Users, groups, ACLs

Future plans

Our immediate plans for SQAN include:

- Extending and enhancing SQAN for wider adoption.
 - Enabling custom QC criteria definition within the portal.
- Developing documentation.
- Promoting wide-spread adoption.
- Building a vibrant community of developers/maintainers .
- Setting up an independent governing body.

In the longer-term future, we intend to expand the capabilities of the SQAN project to extend beyond metadata/protocol QC including but not limited to the areas described below.

Imaging Level QC

While the metadata QC that SQAN performs is an essential first step in data quality assurance, noise and image artifacts are not necessarily flagged by this stage of QC. Noise and artifacts are prevalent in medical imaging; they may originate from a subject's motion or from the instruments themselves. Many techniques have been developed over the years to detect and attempt to correct such artifacts [30, 31]. It is our goal to expand SQAN to be able to detect and catalogue these QC issues and to incorporate existing QC tools and methods developed by the research and clinical imaging community. The first step is to

connect SQAN to an imaging database to gain access to imaging. Rather than create our own system, we will draw on the very mature and robust XNAT system to accomplish this task. An XNAT system is already in production at IU, and we will work with the administrators of that system to connect the two. Once that is accomplished, we will draw on the already-in-use ImageX system or the XNAT viewer to allow users to look at the raw data and look for artifacts in the images. While this will initially be a largely manual task, we hope that in the future we will be able to draw on powerful machine learning tools to detect imaging artifacts automatically. We will also draw on existing QC tools for imaging types (fMRI, DWI, etc) to incorporate those into our system. Once processed data is available, we will also analyze it for artifacts using existing tools, and work with researchers and clinicians to identify areas where tools are in need of development. To this end, we hope to make SQAN a “one-stop shop” for all imaging QA needs. It is also our goal to continue to interface with the community to continually expand the capabilities of SQAN as new imaging techniques are developed.

Software Algorithms/Pipelining

SQAN’s QA needs to lead to better data outcomes and better processing. To that end, we propose to link SQAN with pipelining processes in order to automate processing of data as it is obtained, and to identify the correct parameters to use in those processes when possible. Pipeline systems exist currently (XNAT, brainlife.io, etc.), and our goal is to interface with those rather than build pipelines into SQAN itself. The ideal flow would be as follows: 1) raw data is ingested into SQAN and metadata is checked to make sure correct protocols have been followed; 2) once metadata has been verified, the raw imaging data will be made available for manual viewing or automated QC checks, if available; 3) with both QC steps passed, a call will be made to the appropriate processing pipeline to begin working on the data; 4) processed data will be subject to further QC checks; 5) final vetted data will be made available to researchers. To accomplish this ideal workflow, SQAN will need to be interconnected with imaging databases, pipelining systems, and computational infrastructure. Our plan is to develop these interconnections at IU where each of these pieces is available in some capacity. Once complete, we will have a fully automated imaging data system that will ingest raw data and disseminate final processed data that has been fully vetted by a robust QC system. Such a system would be able to make imaging centers more efficient, and the data more robust for critical analyses.

Integration with other non-imaging databases

One of the most critical aspects of imaging research is to be able to relate imaging metrics to other biomarkers. These can include demographics, genetic information, neuropsychiatric score, and medical records, just to name a few. As such, it is important that all of this data be available to researchers. To that end, we propose to connect SQAN with these various databases so that, once imaging QC has been completed, the user can easily connect the vetted imaging data with important metrics from other sources and more rapidly engage in the desired analyses.

Expanding SQAN to other imaging disciplines

Medical imaging goes beyond the scope of neuroimaging and the few scanners that SQAN currently supports. To that end, we will work with our partners and the wider community to identify instruments

and imaging types that should be integrated into our system. We will reach out to interested individuals to provide test data for these expansions, and to vet the additions for final implementation.

Conclusion

Quality assurance (QA) is a fundamental first step in guaranteeing reliable and reproducible scientific research, and requires continuous involvement by all stakeholders and the use of appropriate quality control (QC) methods to identify data quality issues. Medical imaging, fundamental to clinical research and diagnosis, is costly and generates large datasets, thereby making the need for quality assurance (QA) through adequate quality control (QC) methods even more critical. Existing medical imaging QC methods are designed for specific use-cases with limited scope. In this paper, we have presented Scalable Quality Assurance for Neuroimaging, an open source, automated, flexible QC software suite that has supported 50+ research projects QCing ~3 million images and ~600 million metadata tags. We have begun the process of transitioning the project from an IU-internal project to a self-sustaining, open-source project with its own community of developers and maintainers; as we get ready to deploy SQAN at two more imaging centers—Dartmouth College and Harvard University—we aspire to enable protocol QC at more such centers while also expanding on the levels of QC our system can provide by leveraging existing methods other centers may already have.

Acknowledgment: This work was supported in part by Shared University Research grants from IBM, Inc., to Indiana University.

References

1. Barthel, H., Schroeter, M. L., Hoffmann, K. T., & Sabri, O. (2015, May). PET/MR in dementia and other neurodegenerative diseases. In W.B. Saunders, Ed., *Seminars in nuclear medicine* (Vol. 45, No. 3, pp. 224-233).
2. Assaf, Y., Johansen-Berg, H., & Thiebaut de Schotten, M. (2019). The role of diffusion MRI in neuroscience. *NMR in Biomedicine*, 32(4), e3762.
3. Dumoulin, S. O., Fracasso, A., van der Zwaag, W., Siero, J. C., & Petridou, N. (2018). Ultra-high field MRI: Advancing systems neuroscience towards mesoscopic human brain function. *Neuroimage*, 168, 345-357.
4. Smith, S. M., & Nichols, T. E. (2018). Statistical challenges in “big data” human neuroimaging. *Neuron*, 97(2), 263-268.
5. Borghi, J. A., & Van Gulick, A. E. (2018). Data management and sharing in neuroimaging: Practices and perceptions of MRI researchers. *PloS one*, 13(7), e0200562.
6. Lu, W., Dong, K., Cui, D., Jiao, Q., & Qiu, J. (2019). Quality assurance of human functional magnetic resonance imaging: a literature review. *Quantitative Imaging in Medicine and Surgery*, 9(16).
7. Schwartz, D. L., Tagge, I., Powers, K., Ahn, S., Bakshi, R., Calabresi, P. A., ... & Papinutto, N. (2019). Multisite reliability and repeatability of an advanced brain MRI protocol. *Journal of Magnetic Resonance Imaging*. doi: 10.1002/jmri.26652.
8. Marcus, D.S., et al. (2013). Human Connectome Project informatics: Quality control, database services, and data visualization. *NeuroImage*, Vol. 80, pp. 202-219.
9. Greve, D. N., Mueller, B. A., Liu, T., Turner, J. A., Voyvodic, J., Yetter, E., ... & Ford, J. M. (2011). A novel method for quantifying scanner instability in fMRI. *Magnetic resonance in medicine*, 65(4), 1053-1061.
10. Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage*, 59(3), 2142-2154.
11. Wyman, B.T., et al. (2013). Standardization of analysis sets for reporting results from ADNI MRI data. *Alzheimer's & Dementia*, Vol. 9, pp. 332-337.
12. Alfaro-Almagro, F., et al. (2018). Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *NeuroImage*, Vol. 166, pp. 400-424.
13. Friedman, L., Stern, H., Brown, G. G., Mathalon, D. H., Turner, J., Glover, G. H., ... & Greve, D. N. (2008). Test–retest and between-site reliability in a multicenter fMRI study. *Human brain mapping*, 29(8), 958-972.
14. XNAT protocol check
<https://wiki.xnat.org/docs16/3-administrator-documentation/pipeline-engine/configuring-the-protocol-check-pipeline>
15. Thrall, J.H., et al. (2018) Artificial Intelligence and Machine Learning in Radiology: Opportunities, Challenges, Pitfalls, and Criteria for Success, *Journal of the American College of Radiology*, Vol. 15, Issue 3, Part B. pp. 504-508.
16. Gopu, A., Young, M. D., Hayashi, S., West, J., Paramasivam, M., Avena-Koenigsberger, A., Perigo, R. Scalable Quality Assurance for Neuroimaging (SQAN): Quality control for medical imaging – project overview. 2019. Preprint on IU Scholarworks. DOI: 10.5967/yk35-st11.

17. Young, M. D., Avena-Koenigsberger, A., Gopu, A., West, J., Paramasivam, M., Perigo, R. Scalable Quality Assurance for Neuroimaging (SQAN) - Data Management and Quality Control. 2019. Preprint on IU Scholarworks. DOI: 10.5967/481s-nt96.
18. IU SCA astronomy related portals
 - a. One Degree Imager - Portal, Pipeline and Archive (ODI-PPA): <https://portal.odi.iu.edu>
 - b. Globular Cluster Systems - SCA (GCS-SCA): <https://gcs.sca.iu.edu>
 - c. Spectral Archive (SpArc): <https://sparc.sca.iu.edu>
 - d. Blanco DECam Bulge Survey - SCA (BDBS-SCA): <https://bdb.sca.iu.edu/about>
19. High Performance Computing everywhere (HPC everywhere): <https://hpceverywhere.iu.edu>
20. IU SCA neuroscience related portals/projects
 - a. SQAN project website including demo portal: <https://sqan.sca.iu.edu>
 - b. Brainlife.io: <https://brainlife.io>
 - c. Connectivity pipeline project: <https://github.com/IUSCA/IUSM-connectivity-pipeline>
21. ImageX dynamic web based image visualization software stack: <https://imagex.sca.iu.edu>
22. IU SCA project publication listing: <https://sca.iu.edu/publications>.
23. IU SCA hands-on MEAN Stack Tutorial related links
 - a. Tutorial at ADASS 2019 conference: <http://adass2018.astro.umd.edu/tutorial.html>
 - b. Tutorial within IU <https://pti.iu.edu/newsletter/archive/20181029-Tutorial.html>
 - c. IU SCA MEAN Stack Tutorial on github: <https://github.com/IUSCA/mean-stack>
24. AngularJS <https://angularjs.org/>
25. VueJS <https://vuejs.org/>
26. GitHub Open Source Guides: <https://opensource.guide/>
27. The Linux Foundation Open Source Guides for Enterprise: <https://www.linuxfoundation.org/resources/open-source-guides/>
28. IU SCA staff give a talk titled “Modern Web Technologies Enabling Science & Research” at the Center of Excellence for Women in Technology (CEWiT) Summit. <https://bit.ly/2GRaAG3>
29. Slipping into the Jetstream: Summer internship program brings students to IU for high performance computing experience, conference participation. <https://itconnections.iu.edu/2018-august/rt-summer-interns.php>
30. Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage*, 84, 320-341.
31. Power, J. D., Plitt, M., Gotts, S. J., Kundu, P., Voon, V., Bandettini, P. A., & Martin, A. (2018). Ridding fMRI data of motion-related influences: Removal of signals with distinct spatial and physical bases in multiecho data. *Proceedings of the National Academy of Sciences*, 115(9), E2105-E2114.