Quality Assurance Project Plan

Health and Toxicity Theme
Hydraulic Fracturing Study
Immediate Office National Center for Environmental Assessment
Office of Research and Development
United States Environmental Protection Agency Research Triangle Park, NC 27711

Document Control Number: QAPP-NCEA-IO-HFS-HTT/2012/02-r00

February 2012

PREPARED BY

Name: Lyle D. Burgoon, Ph.D. Phone: (919) 541-7808
Title: Leader, Systems Biology and Bioinformatics

Signature: /s/ Date: 2/23/2012

QUALITY ASSURANCE APPROVAL

Name: Cheryl Itkin Phone: (703) 347-8557
Title: Director of Quality Assurance & Peer Review Coordinator

Signature: /s/ Date: 2/23/2012

MANAGEMENT APPROVAL

Name: Darrell Winner Phone (703) 347-0210
Title: Acting Deputy Director, NCEA

Signature: /s/ Date: 3/2/2012

EPA does not consider this internal planning document an official Agency dissemination of information under the Agency's Information Quality Guidelines, because it is not being used to formulate or support a regulation or guidance; or to represent a final Agency decision or position. This planning document describes the quality assurance/quality control activities and technical requirements that will be used during the research study. EPA plans to publish the research study results in a draft report, which will be reviewed by the EPA Science Advisory Board. The final research report would be considered the official Agency dissemination. Mention of trade names or commercial products in this planning document does not constitute endorsement or recommendation for use.

TABLE OF CONTENTS

TABLE OF CONTENTS II

LIST OF ACRONYMS III

DISTRIBUTION LIST 1

ORGANIZATION AND RESPONSIBILITIES 2

ROLES AND RESPONSIBILITIES 2

NCEA H&T Lead (Lyle D. Burgoon) 3

NCEA H&T Scientists (Ila Cote, Nina Wang, Lyle Burgoon, ORISE Fellows) 3

NCEA Director of Quality Assurance (Cheryl Itkin) and Estimated QA/QC Resources 3

PROJECT DESCRIPTION AND OBJECTIVES 4

PROJECT TIMELINE 5

QUALITY OBJECTIVES AND CRITERIA 5

SPECIAL TRAINING/CERTIFICATION 5

REPORTS TO MANAGEMENT 5

DOCUMENTATION, RECORDS, AND DATA MANAGEMENT 6

DATA REVIEW, VERIFICATION AND VALIDATION 6

ASSESSMENT AND RESPONSE 6

REFERENCE DOCUMENTS 7

QAPP CHANGE PROCEDURE AND HISTORY 7

APPENDIX A: SOFTWARE QUALITY ASSURANCE PLAN (SQAP) 8

SQAP/QAPP CHANGE PROCEDURE AND HISTORY 32

LIST OF ACRONYMS

ADE	Associate Director for Ecology
ADH	Associate Director for Health
ATSDR	Agency for Toxic Substances and Disease Registry
CalEPA	California Environmental Protection Agency
CBI	Confidential Business Information
DCO	Document Control Officer
DQA	Director of Quality Assurance
EPA	Environmental Protection Agency
HEAST	Health Effects Assessment Summary Tables
HFS	Hydraulic Fracturing Study
H&T	Health and Toxicity
HTML	Hypertext Markup Language
HTS-V	High Throughput Screening Value
iNPD	Interim National Program Director
IRIS	Integrated Risk Information System
MRL	Minimal Risk Levels for Hazardous Substances Database
NCEA	National Center for Environmental Assessment
NCCT	National Center for Computational Toxicology
NIST	National Institute of Standards and Technology
OP	Operating Procedure
ORD	Office of Research and Development
OSIM	Office of Science Information Management
PI	Principal Investigator
PSS	Program Support Staff
PPM	Policy and Procedures Manual
PPRTV	Provisional Peer-Reviewed Toxicity Value
QA	Quality Assurance
QAM	Quality Assurance Manager
QAPP	Quality Assurance Project Plan
QAARWP	Quality Assurance Annual Report and Work Plan
QC	Quality Control
QMP	Quality Management Plan
QMS	Quality Management System
QMSR	Quality Management System Review
QSA	Quality Systems Audit
QSAR	Quantitative Structure Activity Relationship
OSWER	Office of Solid Waste and Emergency Response
RCU	Research Cores Unit
RR	Readiness Review
RTP	Research Triangle Park
SSWR	Safe and Sustainable Water Resources
TSA	Technical Systems Audit
USEPA	United States Environmental Protection Agency

LEFT BLANK INTENTIONALLY

DISTRIBUTION LIST

This Quality Assurance Project Plan (QAPP) will be distributed to all US Environmental Protection Agency Health and Toxicity Theme staff and their management:

Dr. Lyle Burgoon, Health and Toxicity Co-Lead, National Center for Environmental Assessment (NCEA)

Dr. Keith Houck, Health and Toxicity Co-Lead, National Center for Computational Toxicology (NCCT)

Dr. Darrell Winner, Acting Deputy Center Director, NCEA

Dr. David Dix, Deputy Center Director, NCCT

Dr. Michael Troyer, Division Director, NCEA

Cheryl Itkin, Director of Quality Assurance and Peer Review Coordinator, NCEA

Stephen Little, Quality Assurance Manager, NCCT

Dr. Ila Cote, Senior Science Advisor, NCEA

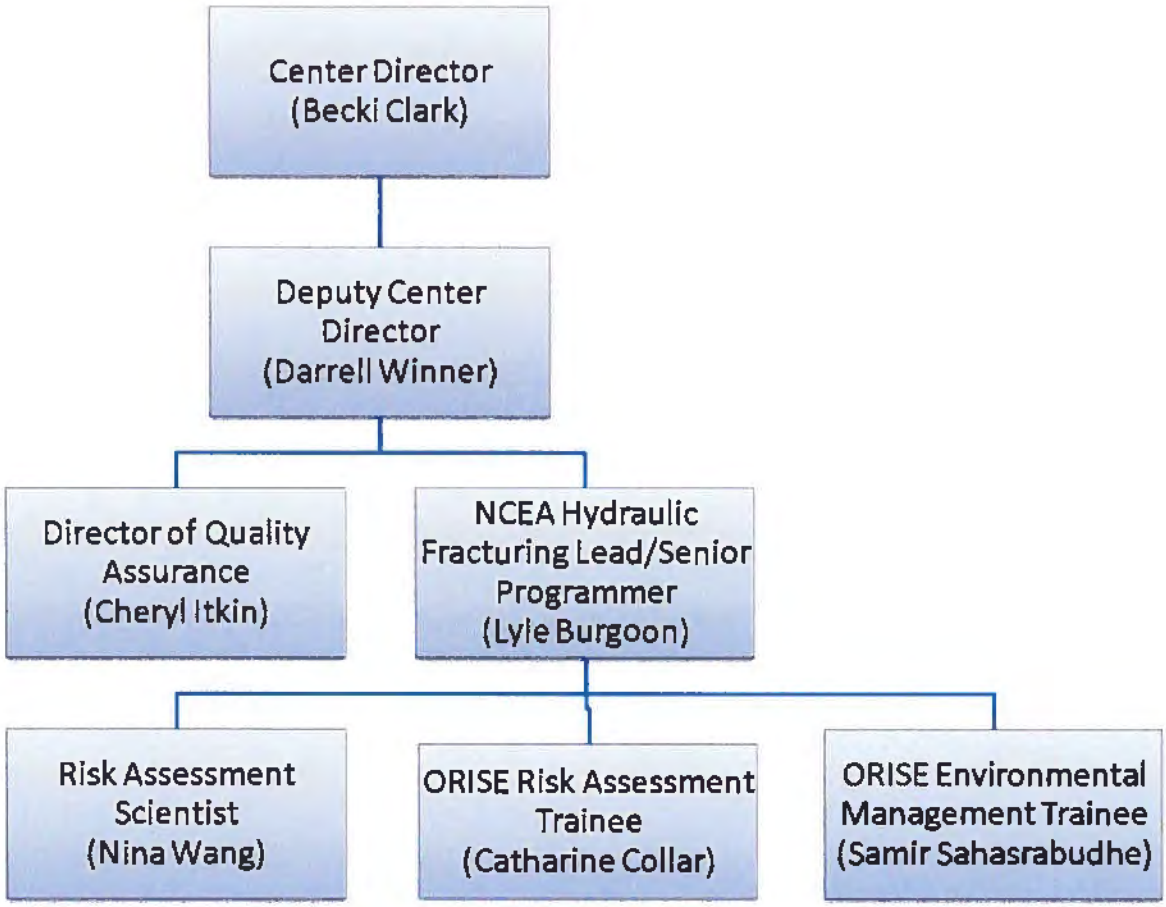
Dr. Nina Wang, Biologist, NCEA

Dr. Ann Richard, Research Chemist, NCCT

Dr. Catharine Collar, ORISE Fellow, NCEA

Samir Sahasrabudhe, ORISE Fellow, NCEA

ORGANIZATION AND RESPONSIBILITIES



The Office of Research and Development Hydraulic Fracturing Study Research Team is managed by the Interim National Program Director (iNPD) for the Safe and Sustainable Water Resources Research Program (SSWR). The work products covered by this QAPP will be generated by the NCEA team in the H&T Team. The H&T Team is co-led by Lyle Burgoon (NCEA) and Keith Houck (NCCT). The H&T Co-Leads report to the Technical Research Lead on research items, and to the Study Coordinator for all other items.

Quality assurance activities for the NCEA H&T Team are managed by the NCEA Director of Quality Assurance (DQA). The NCEA DQA is responsible to the Program Quality Assurance Manager (PQAM) for overall Study QA. The PQAM is responsible and accountable to the ORD Director of Quality Assurance and accountable to the Study Coordinator.

Roles and Responsibilities

This section shall describe each of the roles and delegated responsibilities in the NCEA H&T Team.

NCEA H&T Lead (Lyle D. Burgoon)

The NCEA H&T Lead shares accountability and responsibility for the overall performance of the H&T Team with the NCCT H&T Lead. The H&T Co-Leads shall share responsibility for the creation of any and all H&T Quality Assurance documents. The NCEA H&T Lead shall specifically be responsible for the creation of any and all NCEA-specific QAPPs and Software Quality Assurance Plans (SQAPs), and shall have the authority to delegate responsibility for preparing sections of the NCEA-specific QAPP/SQAP to NCEA members of the H&T Team. The NCEA H&T Lead will advise the NCEA Director on the most appropriate scientific and analytical strategies proposed by the H&T Team for final decision.

NCEA H&T Scientists (Ila Cote, Nina Wang, Lyle Burgoon, ORISE Fellows)

NCEA H&T Scientists are experts in the field of chemical risk assessment. They will be responsible for determining the most scientifically acceptable method for interpreting and using the health and toxicity data. The NCEA H&T Scientists will also be responsible for developing High Throughput Screening Values (HTS-Vs). HTS-Vs development will be covered in a separate QAPP under separate cover. NCEA H&T Scientists will also provide independent verification and validation that the software developed for this project is fit for purpose, and that the data being parsed and returned is correct. The specific methods for doing these verifications and validations will be outlined later in this QAPP/SQAP.

NCEA Director of Quality Assurance (Cheryl Itkin) and Estimated QA/QC Resources

NCEA's Director of Quality Assurance (DQA) also serves as NCEA's Quality Assurance Manager (QAM) and will perform the responsibilities outlined in NCEA's QMP. This role is responsible for the review and approval of all HF QA/QC documents generated by or for NCEA. The DQA will submit NCEA HF Quality Assurance Project Plans (QAPPs) to the HF PQAM for concurrence that they meet HF Research Program requirements and will be responsible for the review and approval of NCEA HF QAPPs. An essential part of the QA system is an assessment/audit and the NCEA DQA or designee, will perform QA Technical System Audits (TSAs), as required by the HF QMP and NCEA HF QAPPs. It is the responsibility of the NCEA DQA to ensure that audits are conducted without conflict of interest. The NCEA DQA will also review NCEA H&T Quarterly Reports of problems and corrective actions, and shall audit these corrections. The DQA will participate in meetings (e.g., teleconferences) organized by the HF NCEA Team and the HF PQAM.

PROJECT DESCRIPTION AND OBJECTIVES

This Quality Assurance Project Plan (QAPP)/Software shall serve as the primary quality assurance plan for the National Center for Environmental Assessment (NCEA) portion of the Health and Toxicity (H&T) Theme's work in the Hydraulic Fracturing Study (HFS).

The goals of this project are specifically to:

1) Identify toxicity values for chemicals identified as being part of, or produced as a result of, hydraulic fracturing operations
2) Screen, categorize, and prioritize these chemicals
    a. Screening: Identify chemicals where toxicity values are known and not known
    b. Categorize: Bin chemicals into low, medium, high, and uncertain concern
    c. Prioritize: Prioritize chemicals of uncertain concern based on Quantitative Structure Activity Relationship (QSAR) analysis and known toxicity values of low certainty.

Six of the top priority chemicals of uncertain concern will be communicated to the NCCT for consideration in ToxCast analyses.

This QAPP will be used by NCEA scientists on the H&T team when identifying previously established toxicity values from existing sources for chemicals identified as part of or produced as a result of hydraulic fracturing operations. These existing data sources for toxicity values shall include:

- (Tier 1) Integrated Risk Information System Database (IRIS)
- (Tier 2) Provisional Peer-Reviewed Toxicity Value Database (PPRTV)
- (Tier 3) Agency for Toxic Substances and Disease Registry (ATSDR) Minimal Risk Levels for Hazardous Substances Database (MRL)
- (Tier 3) California Environmental Protection Agency (CalEPA) peer reviewed toxicity values
- (Tier 3) EPA Health Effects Assessment Summary Tables (HEAST) database

These data sources were chosen as they represent the Office of Solid Waste and Emergency Response (OSWER) "Recommended Human Health Toxicity Value Hierarchy." Specifically, OSWER recommends that Regional risk assessors consult these tiered data sources in order, from Tier 1 through 3 (e.g., if a toxicity value exists in IRIS, it controls over one from PPRTV). These sources all contain peer-reviewed toxicity assessments and toxicity values.

A master spreadsheet of toxicity values will be generated for every chemical in each of these data sources by consulting the specific database. In the case of the CalEPA toxicity values, the database file will be downloaded and converted into a spreadsheet using a single SQL query. This master spreadsheet (where each data source's values will be entered onto its own worksheet) will be generated by an NCEA H&T Scientist. Once the master spreadsheet has been completed, each value will be verified against the original data source using a physical record. A second NCEA H&T Scientist will further verify and validate 10% of the records, chosen at random, within each spreadsheet.

Following verification and validation of all records, an NCEA H&T Scientist will cross-walk the list of chemicals identified as part of, or produced as a result of, hydraulic fracturing operations against the master spreadsheet. Specifically, the list of hydraulic fracturing chemicals will be considered confidential business information (CBI), will contain physico-chemical properties, and will be delivered from the National Center for Computational Toxicology (NCCT) to NCEA via the Research Triangle Park (RTP) CBI Document Control Officer (DCO). An NCEA CBI-cleared H&T Scientist will then cross-walk the CBI hydraulic fracturing chemical list against the master spreadsheet to identify toxicity values. Using the OSWER guidance, toxicity values from a Tier 3 source shall only be used when a Tier 2 value is not available. Further, a Tier 2 value shall only be used when a Tier 1 value is not available.

The decision framework for screening, categorization, and prioritization of chemicals is still being developed. This QAPP will be updated to reflect quality assurance activities associated with the decision framework once it has been approved.

Project Timeline

March 2012 Deliverable: The list of toxicity values will be appended to the list of physico-chemical properties compiled by NCCT for the chemicals identified as part of, or produced as a result of, hydraulic fracturing operations.

March 2012 Deliverable: The list of screened, categorized, and prioritized chemicals for ToxCast analysis consideration.

QUALITY OBJECTIVES AND CRITERIA

We will use the same quality criteria as set forth by OSWER in their guidance on "Recommended Human Health Toxicity Value Hierarchy." Specifically, toxicity values are obtained from the Tier 1 source first, if not available then a Tier 2 source is consulted, and finally a Tier 3 source is only consulted when a Tier 2 value is not available. We will not consider any Tier 3 sources other than those specifically listed within this QAPP. This is due to the fact that the sources listed are well known to us and their values are peer-reviewed.

SPECIAL TRAINING/CERTIFICATION

All NCEA H&T Scientists who are Federal employees must have obtained a CBI clearance to be part of the project team. Any NCEA H&T Scientists who join the project must obtain their CBI clearance prior to becoming actively involved in the project.

REPORTS TO MANAGEMENT

The NCEA H&T Co-Lead will provide updates and reports to NCEA management as requested, and will provide a final report at the conclusion of the study. The final report will detail any problems encountered, quality assurance activities performed, deviation from the QAPP, and corrective actions.

DOCUMENTATION, RECORDS, AND DATA MANAGEMENT

Existing data, including the master spreadsheet, will be saved on the EPA intranet, on ORD managed hardware. This hardware is secured from improper release, and appropriate disaster mitigation methods are in place to prevent data loss.

Confidential Business Information will be handled and managed as required by CBI rules. Safeguards will be in place to prevent inadvertent and accidental release of CBI. In Cincinnati this includes the use of a CBI safe-room. In RTP this includes checking all CBI in with the DCO at the end of the day, or the use of a TSCA CBI approved safe. All CBI and working documents will be closed out according to CBI rules once they are no longer needed.

DATA REVIEW, VERIFICATION AND VALIDATION

A master spreadsheet of toxicity values will be generated for every chemical in each of these data sources by consulting the specific database. In the case of the CalEPA toxicity values, the database file will be downloaded and converted into a spreadsheet using a single SQL query. This master spreadsheet (where each data source's values will be entered onto its own worksheet) will be generated by an NCEA H&T Scientist. Once the master spreadsheet has been completed, each value will be verified against the original data source using a physical record. A second NCEA H&T Scientist will further verify and validate 10% of the records, chosen at random, within each spreadsheet.

To ensure the cross-walk function, which will use functions built into the spreadsheet software, works properly, an NCEA H&T Scientist will validate that the matching fields actually match properly (e.g., CASRN). This is accomplished by using a simple validation function where the spreadsheet returns a 0 into an adjacent cell if the values from the CBI listing and the master spreadsheet match. If they do not match, the spreadsheet enters a 1 into an adjacent cell. The NCEA H&T Scientist will then sum the column with the 0/1 values, and if the sum is 0, then it means that the CASRNs match appropriately. In the instance where the sum is greater than 1, then the NCEA H&T Scientist will debug the problem. This process will be repeated until the sum returned is 0.

ASSESSMENT AND RESPONSE

A Technical Assessment of project operation will be conducted under this HF Research Program. The type of assessment that will be conducted for this work will be a technical systems audit (TSA) scheduled by the NCEA H&T Lead in coordination with the NCEA DQA, when it is most appropriate. The TSA focuses on the final assessment performed to evaluate whether the end product meets the desired performance criteria and outcomes meet the original objectives of the project. Inspection of results, problem resolution and corrective action reports, and interim progress reports will be reviewed during the audit. The TSA will qualitatively document the degree to which QC procedures and processes specified in this approved QAPP are being implemented and will identify problems that are not resolved.

REFERENCE DOCUMENTS

1. EPA Order C10 2106.0, 2008, EPA Office of Environmental Information (OEI), Washington, D.C.

2. ANSI/ASQC E4-2004, *Specifications and Guidelines for Quality Systems for Environmental Data Collection and Environmental Technology Programs*, American National Standards Institute/American Society for Quality Control, 2004.

3. U.S. Environmental Protection Agency, *Guidance for Quality Assurance Project Plans*, EPA QA/G-5, Office of Environmental Information, Washington D.C., December, 2002.

4. U.S. Environmental Protection Agency, EPA Records Schedules, Office of Environmental Information, National Records Management Program, Washington, D.C.
http://www.epa.gov/records/policy/2155/rm_policy_cio_2155_1_2.pdf

5. U.S. Environmental Protection Agency, *TSCA CBI Protection Manual*, 7700A1, Office of Pollution Prevention and Toxics, Washington D.C., October 2003.
http://www.epa.gov/oppt/pubs/tsca-cbi-protection-manual.pdf

QAPP Change Procedure and History

The QAPP is a living document. NCEA Team Members may request changes be made to the QAPP by contacting the NCEA H&T Lead. The NCEA H&T Lead will have the final decision as to whether to change the document, and how. The NCEA H&T Lead may delegate this responsibility as necessary to others. Any changes to the document will be noted in the following table:

Date of Change	Name of Personnel Editing Document	Nature of Change (include description and current page number(s) if applicable)
2-13-12	Lyle D. Burgoon	Initial Write.
2-15-12	Cheryl Itkin	Edits to Acronyms, references, Assessment & Response, Appendix
2-17-12	Vicki Soto	Edits to TOC, Appendix

Software Quality Assurance Plan

Health and Toxicity Theme
Hydraulic Fracturing Study
Immediate Office
National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Document Control Number: QAPP-NCEA-IO-HFS-HTT/2012/01-r00

January 2012

PREPARED BY

Name: Lyle D. Burgoon, Ph.D. Phone: (919) 541-7808
Title: Leader, Systems Biology and Bioinformatics

Signature: /s/ Date: 02/07/2012

QUALITY ASSURANCE APPROVAL

Name: Cheryl Itkin Phone: (703) 347-8557
Title: Director of Quality Assurance & Peer Review Coordinator

Signature: /s/ Date: 02/07/2012

MANAGEMENT APPROVAL

Name: Darrell Winner Phone (703) 347-0210
Title: Acting Deputy Director, NCEA

Signature: /s/ Date: 02/07/2012

TABLE OF CONTENTS

LIST OF ACRONYMS 11

PURPOSE 13
- *PURPOSE AND SCOPE* 13
- *NAMES OF SOFTWARE ITEMS COVERED* 13
- *SOFTWARE LIFECYCLE COVERED* 13
- *SOFTWARE LIFETIME* 14

DEVELOPMENT OF HTS-V DOCUMENTS 14

REFERENCE DOCUMENTS 14

MANAGEMENT 15
- *ORGANIZATION: NCEA HEALTH AND TOXICITY HYDRAULIC FRACTURING STUDY TEAM* 15
- *TASKS* 16
- *ROLES AND RESPONSIBILITIES* 16
  - *NCEA H&T Lead (Lyle D. Burgoon)* 17
  - *NCEA H&T Lead Software Developer (Lyle D. Burgoon)* 17
  - *NCEA H&T Software Developer (To Be Determined)* 17
  - *NCEA H&T Scientists (Ila Cote, Nina Wang, Lyle Burgoon, ORISE Fellows)* 17
  - *NCEA Director of Quality Assurance (Cheryl Itkin) and Estimated QA/QC Resources* 17

DOCUMENTATION 19
- *Software Requirements Description (SRD)* 19
- *Software Design Description (SDD)* 19
- *Verification and Validation Plans* 20
- *Verification and Validation Results Reports* 21
- *User Documentation* 21
- *Software Configuration Management Plan* 21

STANDARDS, PRACTICES, CONVENTIONS AND METRICS 22
- *Content* 22

SOFTWARE REVIEWS 25

TESTS 25

PROBLEM REPORTING AND CORRECTIVE ACTIONS 25

TOOLS, TECHNIQUES AND METHODOLOGIES 26

MEDIA CONTROL 26

SUPPLIER CONTROL 26

RECORDS COLLECTION, MAINTENANCE, AND RETENTION 26

TRAINING 26

ASSESSMENT AND RESPONSE 27

RISK MANAGEMENT 27

GLOSSARY 28

CROSSWALK 29

SQAP/QAPP CHANGE PROCEDURE AND HISTORY 32

LIST OF ACRONYMS

ADE	Associate Director for Ecology
ADH	Associate Director for Health
ANSI	American National Standards Institute
ASQC	American Society for Quality Control
BDD	Behavior Driven Development
DQA	Director of Quality Assurance
EPA	Environmental Protection Agency
GLPS	Good Laboratory Practice Standards
HFS	Hydraulic Fracturing Study
H&T	Health and Toxicity
HTML	Hypertext Markup Language
HTS-V	High Throughput Screening Value
IEEE	Institute of Electrical and Electronics Engineers
iNPD	Interim National Program Director
NCEA	National Center for Environmental Assessment
NCCT	National Center for Computational Toxicology
NIST	National Institute of Standards and Technology
OP	Operating Procedure
ORD	Office of Research and Development
OSIM	Office of Science Information Management
PI	Principal Investigator
PPM	Policy and Procedures Manual
PPRTV	Provisional Peer-Reviewed Toxicity Value
PQAM	Program Quality Assurance Manager
QA	Quality Assurance
QAM	Quality Assurance Manager
QAPP	Quality Assurance Project Plan
QAARWP	Quality Assurance Annual Report and Work Plan
QC	Quality Control
QMP	Quality Management Plan
QMS	Quality Management System
QMSR	Quality Management System Review
QSA	Quality Systems Audit
RCU	Research Cores Unit
RR	Readiness Review
RTP	Research Triangle Park
SDR	Software Design Description
SQAP	Software Quality Assurance Plan
SRD	Software Requirements Description
SSWR	Safe and Sustainable Water Resources
TDD	Test Driven Development
TSA	Technical Systems Audit
USEPA	United States Environmental Protection Agency

LEFT BLANK INTENTIONALLY

PURPOSE

Purpose and Scope

This Quality Assurance Project Plan (QAPP)/Software Quality Assurance Plan (SQAP) shall serve as the primary quality assurance plan for the National Center for Environmental Assessment (NCEA) portion of the Health and Toxicity (H&T) work in the Hydraulic Fracturing Study (HFS). The SQAP portion is compliant with the Institute of Electrical and Electronics Engineers (IEEE) Standard 730-2002 (IEEE Standard for Software Quality Assurance Plans). See Appendix A for a crosswalk between IEEE and EPA R5 Quality Assurance (QA) requirements.

This QAPP/SQAP will be used by NCEA scientists on the H&T team when developing hypertext markup language (HTML) parsers to obtain toxicity values from web-based resources, building a master database of toxicity values, and querying the master database of toxicity values. Development of High Throughput Screening Values (HTS-Vs) for use in the HFS will be covered by a separate QAPP under separate cover.

This QAPP/SQAP will not cover work or activities that are not part of the HFS. Furthermore, this QAPP/SQAP will not cover work being performed by the National Center for Computational Toxicology (NCCT). In instances where there is a discrepancy between this QAPP/SQAP and other previously approved QAPP/SQAPs that are currently in use at NCEA, the default is that the currently approved conflicting QAPP/SQAP shall have force over and above this QAPP/SQAP, unless the NCEA Director of Quality Assurance makes a determination to the contrary in writing.

Names of Software Items Covered

The software applications and systems covered by this QAPP/SQAP are currently under development and currently known as:

- IRISParser
- PPRTVParser
- MRLParser
- HEASTParser
- ToxValues Database (TV-DB)

The software applications and systems covered may include multiple software classes, including those classes which were written by outside parties and are used by this software. Thus, this QAPP/SQAP only covers those parts of the software applications and systems that are "original works", as that term is defined in Title 17 of the United States Code, of this project.

Software Lifecycle Covered

This QAPP/SQAP shall cover the entire software lifecycle.

Software Lifetime

The software development activity covered under this QAPP/SQAP will have a calendar lifetime not to exceed 1 year from the date coding starts. Should development or refinement be necessary beyond the 1 year anniversary, a new QAPP/SQAP will need to be approved and executed.

DEVELOPMENT OF HTS-V DOCUMENTS

The HTS-V documents will follow the similar procedures as already outlined for PPRTVs. The QAPP for HTS-V development will be a separate document from this one.

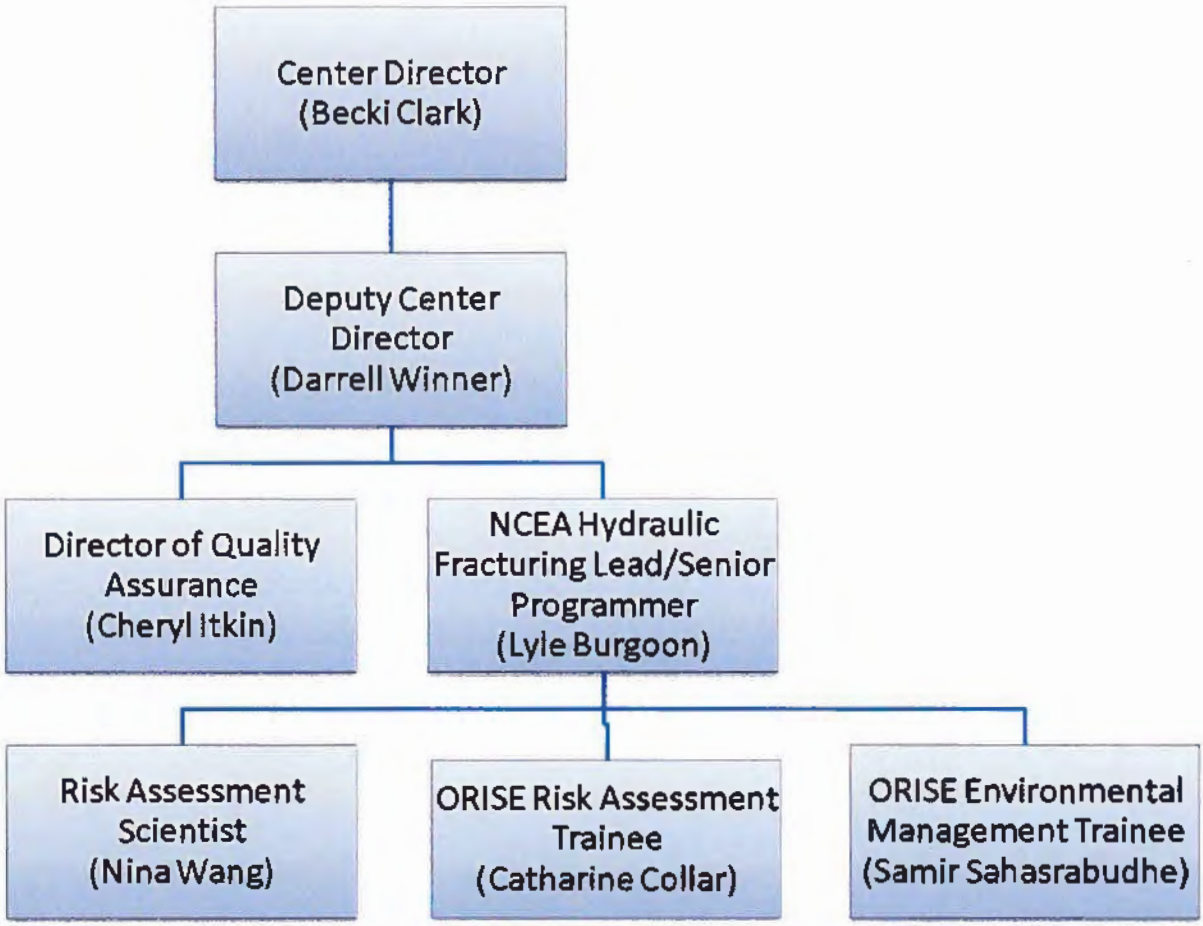
REFERENCE DOCUMENTS

1. IEEE Std 730 - 2002, IEEE Standard for Software Quality Assurance Plans

2. U.S. Environmental Protection Agency, *Guidance for Quality Assurance Project Plans*, EPA QA/G-5M, Office of Environmental Information, Washington D.C., December, 2002.

MANAGEMENT

Organization: NCEA Health and Toxicity Hydraulic Fracturing Study Team



The Office of Research and Development Hydraulic Fracturing Study Research Team is managed by the Interim National Program Director (iNPD) for the Safe and Sustainable Water Resources Research Program (SSWR). The work products covered by this QAPP/SQAP will be generated by the NCEA team in the H&T Team. The H&T Team is co-lead by Lyle Burgoon (NCEA) and Keith Houck (NCCT). The H&T Co-Leads report to the Technical Research Lead on research items, and to the Study Coordinator for all other items.

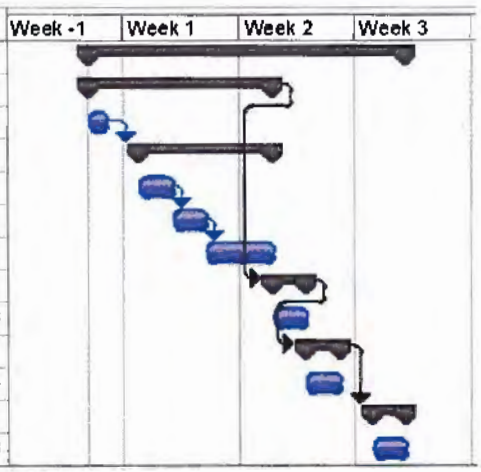
Quality assurance activities for the NCEA H&T Team are managed by the NCEA Director of Quality Assurance (DQA). The NCEA DQA is responsible to the Program Quality Assurance Manager (PQAM) for overall Study QA. The PQAM is responsible and accountable to the ORD Director of Quality Assurance and accountable to the Study Coordinator.

Tasks

This QAPP/SQAP shall cover the entire software lifecycle.

The following tasks will be undertaken at the specified time intervals.

ID	Task Name	Duration
1	**Master Database Development**	13 days
2	**IRIS Parser**	7 days
3	Generate Model Classes	1 day
4	**Generate Parsers**	6 days
5	Generate OralRFD Parser	2 days
6	Generate Oral Slope Parser	2 days
7	Generate Inhalation RFC Parser	2 days
8	**Heast Parser**	2 days
9	Generalized HEAST Parser	2 days
10	**PPRTV Parser**	2 days
11	Generalized PPRTV Parser	2 days
12	**ATSDR MRL Parser**	2 days
13	Generalized Parser	2 days



The entry criteria shall be that the preceding task is completed to spec, and has been tested to be fit for purpose as per the specification or design documentation. The exit criteria for a task shall be that the task is completed to spec, and has been tested to be fit for purpose as per the specification or design documentation. Since this is a relatively small software project, with a user pool limited to software developers, there will not be a formal design/requirements document developed. Instead, the specs for each task will be maintained in the class code associated with each task. The specs and fitness for purpose will be evaluated using either standard Test Driven or Behavior Driven Development (TDD and BDD, respectively) practices. Specifications of these methods are discussed later in this SQAP/QAPP. The completion of this entire coding process is a major milestone for the H&T Team.

Roles and Responsibilities

This section shall describe each of the roles and delegated responsibilities in the NCEA H&T Team.

NCEA H&T Lead (Lyle D. Burgoon)

The NCEA H&T Lead shares accountability and responsibility for the overall performance of the H&T Team with the NCCT H&T Lead. The H&T Co-Leads shall share responsibility for the creation of any and all H&T Quality Assurance documents. The NCEA H&T Lead shall specifically be responsible for the creation of any and all NCEA-specific QAPPs/SQAPs, and shall have the authority to delegate responsibility for preparing sections of the NCEA-specific QAPP/SQAP to NCEA members of the H&T Team. The NCEA H&T Lead will advise the NCEA Director on the most appropriate scientific and analytical strategies proposed by the H&T Team for final decision.

NCEA H&T Lead Software Developer (Lyle D. Burgoon)

The NCEA H&T Lead Software Developer shall be responsible for the development of all NCEA H&T software applications. This shall also include mathematical models and code that is specific to another application (e.g., code used in the R statistical language). The Lead Software Developer shall ensure that all quality assurance practices, policies, and plans are followed and implemented by NCEA H&T Software Development Staff. When necessary, the Lead Software Developer shall develop software as well as be responsible for the development of all software quality assurance documents. The NCEA H&T Lead Software Developer shall be responsible to the NCEA H&T Lead.

NCEA H&T Software Developer (To Be Determined)

The NCEA H&T Software Developer shall be responsible for the development of NCEA H&T software applications as directed by the NCEA H&T Lead Software Developer. The NCEA H&T Software Developer shall identify and implement all applicable quality assurance practices, policies, and plans. The NCEA H&T Software Developer may be delegated additional responsibilities from the NCEA H&T Lead Software Developer or NCEA H&T Lead. The NCEA H&T Software Developer shall be responsible to the NCEA H&T Lead Software Developer.

NCEA H&T Scientists (Ila Cote, Nina Wang, Lyle Burgoon, ORISE Fellows)

NCEA H&T Scientists are experts in the field of chemical risk assessment. They will be responsible for determining the most scientifically acceptable method for interpreting and using the health and toxicity data. The NCEA H&T Scientists will also be responsible for developing High Throughput Screening Values (HTS-Vs). HTS-Vs development will be covered in a separate QAPP under separate cover. NCEA H&T Scientists will also provide independent verification and validation that the software developed for this project is fit for purpose, and that the data being parsed and returned is correct. The specific methods for doing these verifications and validations will be outlined later in this QAPP/SQAP.

NCEA Director of Quality Assurance (Cheryl Itkin) and Estimated QA/QC Resources

NCEA's Director of Quality Assurance (DQA) also serves as NCEA's Quality Assurance Manager (QAM) and will perform the responsibilities outlined in NCEA's QMP. This role is responsible for the review and approval of all HF QA/QC documents generated by or for NCEA.

The DQA will submit NCEA HF Quality Assurance Project Plans (QAPPs) to the HF PQAM for concurrence that they meet HF Research Program requirements and will be responsible for the review and approval of NCEA HF QAPPs. An essential part of the QA system is an assessment/audit and the NCEA DQA or designee, will perform QA Technical System Audits (TSAs), as required by the HF QMP and NCEA HF QAPPs. It is the responsibility of the NCEA DQA to ensure that audits are conducted without conflict of interest. The NCEA DQA will also review NCEA H&T Quarterly Reports of problems and corrective actions, and shall audit these corrections. The DQA will participate in meetings (e.g., teleconferences) organized by the HF NCEA Team and the HF PQAM.

Our general and professional coding standards dictate that Quality Assurance and Control (QA/QC) is an essential part of the product development process. QA/QC is "baked-in" to every product we develop through our documentation and code development procedures. Thus, QA/QC is an essential component of the job of every member of the H&T team. With respect to code development, at least 50% of any Software Developer's code development FTE is spent on QA/QC activities. For NCEA H&T Scientists, they will spend approximately 5-10% of their FTE on software development QA/QC activities.

DOCUMENTATION

Software Requirements Description (SRD)

Due to the relatively low size and complexity of this software system, a separate SRD is not required. Instead, the SRD follows:

Requirement #1:
The System must connect to the IRIS, PPRTV, HEAST, or ATSDR MRL website.

Requirement #2:
The System, once connected, must identify all of the chemicals available for query from the website.

Requirement #3:
After Requirement #2 is satisfied, the System must query the website, one chemical at a time, capture the webpage as raw HTML in memory, and parse out the toxicity values.

Requirement #4:
After Requirement #3 is satisfied, the System must enter the parsed toxicity value into the appropriate table in the Master Toxicity Value Database.

Operating Environment:
The operating environment shall be any operating system with Ruby installed. Specific packages required for operating the software will be made known by the Developer at the time of product delivery.

Safety Requirements:
The System will only write to a SQLite or MySQL database. The System will not read any files from the hard disk. Ruby is a mature programming language with a mature interpreter, and has a low risk profile based on data available through the NIST Vulnerabilities Database. The risk posed by this System is low.

Software Design Description (SDD)

Due to the relatively low size and complexity of this software system, a separate SDD is not required. Instead, the SDD follows:

The System shall consist of two types of classes: Models and Controllers. Models are classes that implement the data model. The data model shall consist of those classes which model the data obtained by the System, and which will be read into the database. Controllers are classes that implement the business logic of the System.

The System shall consist of 4 separate applications, one for each separate data source and parser. Although the separate applications may share the same code base, they will not otherwise communicate or interact with each other. A master system application will be responsible for activating each separate application, and populating the Master Toxicity Database with the data from each separate application.

Verification and Validation Plans

This software system will be developed using Agile programming practices, including the software development best practice of either test-driven or behavior driven development (TDD and BDD, respectively). Under both practices, the software is verified as it is developed. Both TDD and BDD require that separate, very simple pieces of code, called the unit tests (TDD) or system behavior tests (BDD) are written prior to the application being developed. Application code is then written to implement the function covered by the unit or system behavior test. Once the application code's function passes (i.e., the result of the application code function is verified to be the expected value), the developer moves on to develop the next function. As each code function is written, all of the previous tests are re-run (this is called regression testing). The purpose of re-rerunning the previous tests is to ensure that new code do not interfere with the proper functioning of previous code. Any time a change or fix is made to any part of the code base, all of the tests must be re-run, and code fixed, until all of the tests that had previous passed, continue to pass. Application development is complete once all of the unit or system behavior tests pass.

The TDD/BDD tests will meet at least the following conditions:

- For each website/database, a test for 10 randomly chosen chemicals that are listed in the database.
- At least 2 of the 10 randomly chosen chemicals must have a blank value, or a "not determined" value, or similar, for at least one reported toxicity value class (e.g., RfD, RfC).
- The tests will determine if the application correctly identifies all known toxicity values for that chemical for the website/database being parsed.
- Tests will also be run to determine that data uploaded into the database are valid by comparing values in the database to values in application memory. This check will be run for 10 chemicals.

In addition to TDD or BDD, the data produced by the applications will be validated. Validation will occur through a blind process. One NCEA H&T Scientist will obtain all toxicity values from all websites/databases, and copy the toxicity values into a spreadsheet – this will be the validation pool. The validation pool will be sent to a second NCEA H&T Scientist. The second NCEA H&T Scientist will compare the validation pool results with the results from the Master Toxicity Value Database. The second NCEA H&T Scientist will check to ensure that the results from the validation pool match those values within the Master Toxicity Value Database. If differences exist, the second NCEA H&T Scientist will make a notation of the discrepancies and send a notice to the first NCEA H&T Scientist of which chemicals failed. The first NCEA H&T Scientist will re-validate their results, make corrections as necessary, and send their re-validated pool to the second NCEA H&T Scientist. The second NCEA H&T Scientist will re-compare the two lists. If a discrepancy still exists, the second NCEA H&T Scientist will notify the application developer of the discrepancy, and the developer will need to rectify the software. The process will then be repeated. If the discrepancy cannot be rectified, a conference will be held where both

NCEA H&T Scientists bring up the website and re-validate the chemical information. The entire process will then be repeated until there is validation of the results.

Verification and Validation Results Reports

A Verification Results Report shall be prepared by the software developer outlining the steps taken during verification, any difficulties with verification, and the end results of the verification.

A Validation Results Report shall be prepared by the two NCEA H&T Scientists validating the application. This report shall outline the steps taken during validation, any difficulties with validation, and the end results of the validation process.

User Documentation

Documentation will be produced that describes how to install, operate, manage, and maintain the software. The level of documentation will be appropriate for a software engineer, as that is the intended user. Documentation will be developed using community standards, including in-line documentation within the source code and README file(s). The GNU Software standard for documentation will be followed.

Software Configuration Management Plan

All software will be developed using the Git software version control system. All software will be checked out into a development branch. Once specific pieces of the code pass their specs and tests, the code will be checked back in. Change contention will be avoided by directing only specific people to work on specific parts of the code. Should change contention occur, or if a change in one piece alters the test results of another's code upon check-in, the coders will cooperate to fix the issues.

STANDARDS, PRACTICES, CONVENTIONS AND METRICS

Content

Documentation and Commentary Standards

All source code will be documented with a standard header. The header shall have the following form:

```
##########################################################################
# name_of_class_or_file.file_extension
#
# Name and Affiliation of Software Engineer/Coder
#
# Version Information
# Version      #.#      MM-DD-YY   Annotation (e.g., Initial Write, Updated section to reflect
# blah)
#
# Purpose
# State the purpose of this class or particular code snippet.
#
# Notes
# Note anything important or worthwhile for future coders here.
#
# Dependencies
# List what this code depends on.
#
# Requires the following to be installed:
# List any requirements for other external software packages that must be installed for this to
# work.
##########################################################################
```

An example header follows:

```
##############################################################################
#########
# iris_results_scraper.rb
#
# Lyle D. Burgoon
# Leader, Systems Biology and Bioinformatics
# Immediate Office
# National Center for Environmental Assessment (NCEA)
# Office of Research and Development
# US Environmental Protection Agency
#
# Version Information
# Version     1.0     9-9-11  Initial Write
#
```

```
# Purpose
# The purpose of this class is to parse the results page from the IRIS compare site. I've
# saved the results page for now on the hard drive.
# The parser will do the following:
#
# 1) Obtain the tox values from the website for each chemical
#
# Notes
# 1) The IRIS Compare Values HTML files do NOT use HTML best practice notations for their
#    tables -- they have nested tables when they are not used
#
# Dependencies
# This is part of the IRIScraper application; however, it will be migrated to the IRIS
# Ruby library after initial testing (as part of the IRIScraper application)
#
# Requires the following gems to be installed:
# nokogiri
##################################################################################
##############
```

The goal for in-line documentation of the source code is that anyone can pick up the code and understand what each line or a code block does. There is no such thing as "over documenting" the code.

Design Standards

Generally, if code involves a graphical user interface, then the Model-View-Controller design framework would be used. Since this code only involves a series of parsers, and inputting data into a database, the simpler Model-Controller framework will be used. Under this framework, an object model is generated that reflects specific data abstractions. Many of these data abstractions will also reflect database tables. Thus, objects derived from the Model classes will hold the data. The objects from the Controller classes will perform the business logic (e.g., parsing, document input-output (I/O)).

Coding Standards

Naming Conventions

Naming of variables, classes, and methods shall reflect the purpose of the variable, class, or method (i.e., they should be meaningful). Variables used for iteration through loops shall be named *i* or *j* out of standard practice, unless a meaningful name is used. A variable name may start with an underscore (e.g., _variable) when acceptable or expected under the community standards for that language. Compound names shall be written in the format that is standard for the language of choice. For instance, in Ruby, the standard is "snake_case", whereas in Java the standard is CamelCase.

Control Structure Syntax

All control loops shall have the enclosing brace follow on the same line as the control loop declaration (e.g., if (i < terminatorValue){ }. The ending brace will be at the same indentation level as the control command (e.g.

```
if(i < terminatorValue){

        ...

} //end if
```

All code contained within a loop will be indented 1 tab (should be 5 spaces) from the indentation level of the control command (if, else, etc.)

Testing Standards

Test or Behavior Driven Development (TDD and BDD, respectively) methods shall be used. Under the TDD framework, tests are written for the software package prior to any other code being written. The goal is to then write code such that the tests to pass. A regression testing framework shall be used, whereby all previous tests that pass must continue to pass as additional code is developed.

Under the BDD framework, user specs are developed, which in turn specify certain behaviors (which can be thought of as tests) that the software must exhibit in order to pass. Regression testing is still a component of the BDD framework, whereby behavior tests must continue to pass as additional code is developed. The advantages of the BDD framework are that they allow the end user to be an active participant in the development, as the behaviors are written either by the end user, or in such a way as the end user can understand what is supposed to happen, and can readily identify if the software is being tested appropriately.

SOFTWARE REVIEWS

Due to the relatively small size and limited scope of this software project, the following software reviews will be conducted:

- Verification and Validation Plan Review (VVPR)
- Functional Audit (FA)
- Physical Audit (PA)

The VVPR ensures the adequacy of the verification and validation plan. This will be conducted as part of the review process for the overall SQAP/QAPP.

The FA is held at the completion of the coding aspect of the project to ensure the software will deliver as promised. In this case, this will be an audit of the verification and validation plan results.

The PA will be held as the completion of the coding aspect of the project to ensure the software was adequately documented.

Other audits will be held and scheduled as the NCEA Director of Quality Assurance, the Program Quality Assurance Manager, or the ORD Director of Quality Assurance dictates or feels are necessary.

TESTS

This section generally details all tests not otherwise specified in the Verification and Validation Plan. Generally, these tests would be hardware integration or other similar tests. Given the simple nature of this project, there will be no tests performed outside of the Verification and Validation Plan.

PROBLEM REPORTING AND CORRECTIVE ACTIONS

Due to the short life of this project, all problem reporting and corrective actions will be channeled through the NCEA H&T Lead. The NCEA H&T Lead shall develop a Quarterly Report of all problems and corrective actions taken after product delivery. If no problems are noted, then no report will be issued. The Quarterly Report will be sent to the NCEA Director of Quality Assurance.

The NCEA Lead Software Developer and NCEA Software Developer will have the responsibility of ensuring the problem is corrected, and noting the corrective action taken. The NCEA Director of Quality Assurance shall audit these corrections at their leisure.

TOOLS, TECHNIQUES AND METHODOLOGIES

The Git version control system shall be used for all software/code version control. The Aptana Studio 3 software shall be used for all code development. All code will be developed in the Ruby Programming Language (v 1.9). A SQLite or MySQL database shall be used for data management.

The project methodology in use is the Scrum-Ban (or Scrumban) method. IceScrum software may be used to facilitate project management if the ORD Office of Science Information Management approves the installation of the tool. If IceScrum fails to obtain approval, standard Scrum-Ban practices will be performed using a white-board and sticky notes.

MEDIA CONTROL

No physical copies of the software will be developed.

SUPPLIER CONTROL

No suppliers are being used in this process. All outside software being used for this project are available as open source software, and have a long life and history. Although there are known bugs, these bugs have been checked and verified that they will not hamper, hinder, or otherwise impact this project. All software development on our specific code will be performed in-house.

RECORDS COLLECTION, MAINTENANCE, AND RETENTION

Records pertaining to this project shall include all Agency records, including software source code, that are generated through the life of this project. All records will be collected and maintained in accordance with current Agency guidance, standards, and rules regarding records. The retention schedule for all records associated with this project shall be set by the NCEA Records Management staff.

Prior to the conclusion of this project, all study staff will send their records to the NCEA H&T Lead for compilation into the NCEA H&T Study File.

TRAINING

All NCEA employees have been trained to perform their duties. If it occurs that additional training is required, the NCEA H&T Lead shall communicate these additional needs to the appropriate Branch Chief or Division Director.

ASSESSMENT AND RESPONSE

A Technical Assessment of project operation will be conducted under this HF Research Program. The type of assessment that will be conducted for this work will be a technical systems audit (TSA) scheduled by the NCEA H&T Lead in coordination with the NCEA DQA, when it is most appropriate. The software reviews and configuration testing focus on interim assessments conducted iteratively throughout the software development process. In contrast, the TSA focuses on the final assessment performed in the final stages of development and after the software has been applied to evaluate whether the software meets the desired performance criteria and outcomes meet the original objectives of the project. Inspection of software documentation and test results, problem resolution and corrective action reports, and interim progress reports will be reviewed during the audit. The TSA will qualitatively document the degree to which software procedures and processes specified in this approved QAPP are being implemented and will identify problems that are not resolved.

RISK MANAGEMENT

A risk assessment for this project has already been performed. Specific risks fall into the category of catastrophic data loss, periodic data loss, and periodic inaccessibility of the Agency network. All of these risks are currently managed by the Office of Science Information Management (OSIM) in its role as the ORD IM/IT organization. We have assurances from OSIM that they have risk management protocols in place, that include network up-time guarantees and enterprise data back-up.

GLOSSARY

Behavior Driven Development: Software development framework that encompasses development of user stories that describe specific system behaviors in the user's domain-specific language, that can be used for direct testing and validation. Under the framework, the behavior tests are written prior to any code is written. Software is considered "complete" once all of the user-defined system behavior tests pass.

Regression Testing: Regression testing is a method where all prior tests must pass, in addition to the current test, for a particular code development change or addition to be considered complete. Regression testing is a fundamental aspect of the Behavior and Test Driven Development methods (BDD and TDD, respectively).

Scrum: Scrum is an agile project management method that allows software projects to adapt quickly to changing user requirements.

Scrum-ban (or Scrumban): A lean and agile project management method that allows software projects to adapt quickly to changing user requirements. The "ban" part comes from the Toyota Project Managment method called "Kanban." Kanban improves Scrum by recognizing that there is a limit to the amount of work that can get done, and to help teams prioritize work. Kanban also uses a public board where everyone on the team, and outside the team, can see the progress of projects.

Test Driven Development: Software development framework that encompasses development of tests that are written prior to any code. The software is considered "complete" once all of the tests pass.

CROSSWALK

CROSSWALK OF REQUIRED ELEMENTS BETWEEN
IEEE SQAP AND EPA QAPP
For Hardware/Software Development

IEEE Std. 730, Standard for Software Quality Assurance Plans (SQAPs) For Hardware/Software Development	EPA G-5M, Guidance for Quality Assurance Project Plans (QAPPs) for Model/Software Development
4.1 Purpose (of SQAP)	A1. Project management A1. Title and Approval Sheet A2. Table of Contents A3. Distribution List
4.2 Reference Documents	A9. Documentation and Records
4.3 Management 4.3.1 organization 4.3.2 tasks 4.3.3 roles and responsibilities 4.3.4 quality assurance estimated resources	A4. Project/Task Organization
4.4 Documentation: 4.4.1 purpose (of software) 4.4.2 minimum documentation 4.4.2.1 software requirements 4.4.2.2 software design description 4.4.2.3 verification and validation plans 4.4.2.5 user documentation 4.4.2.0 software configuration management plan 4.4.3 other documentation	A5. Problem definition/background A6. Project/Task description schedule A7. Quality objectives and criteria for hardware/software inputs/outputs A9. Documents and Records (Configuration documents, reports, and manuals) B5. Quality control B7. Calibration B9. Non-direct measurements (SOPs) B10. Data mgmt & HW/SW configuration (documentation)
4.5 Standards, practices, conventions, and metrics 4.5.1 purpose 4.5.2 content	A9. Documentation and Records (e.g., Configuration Management and Maintenance Manuals) B9. Non-direct measurements (SOPs) C2. Reports to management

4.6 Software Reviews 4.6.1 purpose 4.6.2 minimum requirements 4.6.2.1 software specifications review 4.6.2.2 architecture design review 4.6.2.3 detailed design review 4.6.2.4 verification and validation plan review 4.6.2.5 functional audit 4.6.2.6 physical audit 4.6.2.7 in-process audits 4.6.2.8 managerial reviews 4.6.2.9 software configuration management plan review 4.6.2.10 post-implementation review 4.6.3 other reviews and audits	B5. Quality control B10. HW/SW configuration (testing) C1. Assessment and Response Actions Oversight quantitative and qualitative assessments
4.7 Test	B10.b. Hardware and software Configuration (Testing) for programming error, (Software code development inspections; Software code verification and performance testing; Acceptance testing) C1. Assessment and Response Actions C1. Hardware /Software Assessments C1. Hardware /Software configuration tests C1. Plans for science and product peer review D1. Departures from Validation Criteria D2. Validation methods D3. Reconciliation with user requirements
4.8 Problem reporting, and corrective action	C1. Assessment and Response Actions (Performance Evaluations) C2. Reports to management
4.9 Tools, techniques, and methodologies	A9. Documentation and Records B9. Non-direct measurements (SOPs)
4.10 Media control	A9. Documentation and Records
4.11 Supplier Control	B8. Inspection/Acceptance requirements for Supplies and Consumables
4.12 Records collection, maintenance, and retention	A9. Documentation of records
4.13 Training	A8. Special training requirements/certification

4.14 Risk management	
Glossary	A1. Project management
SQAP change procedure & history	A1. Project management

The following EPA G-5 elements are not applicable to software development and IEEE -730:

B1. Sampling Process Design
B2. Sampling Methods
B3. Sample Handling and Custody
B4. Analytical Methods
B6. Instrument/Equipment Testing, Inspection, and Maintenance

References

Guidance for Quality Assurance Project Plans for Modeling (EPA QA/G-5M), 2002

IEE Std 730 - 2002, Standard for Software Quality Assurance Plans, (2002)

SQAP/QAPP CHANGE PROCEDURE AND HISTORY

The SQAP/QAPP is a living document. NCEA Team Members may request changes be made to the SQAP/QAPP by contacting the NCEA H&T Lead. The NCEA H&T Lead will have the final decision as to whether to change the document, and how. The NCEA H&T Lead may delegate this responsibility as necessary to others. Any changes to the document will be noted in the following table:

Date of Change	Name of Personnel Editing Document	Nature of Change (include description and current page number(s) if applicable)
12-13-11	Lyle D. Burgoon	Initial Write.
01-26-12	Lyle D. Burgoon, Cheryl Itkin	Page 5: Abbreviation "HFS" included. Page 8: URLs updated. Page 12: Clarified FTE commitments. Pages 13, 20: Corrected capitalization of "sqlite" and added MySQL as a database management option. Page 12, Role and Responsibilities, added the DQA/QA Manager role. Pages 14, 17, 19: Corrected various spelling/grammar issues. Page 21, Assessment and Response section and TSA, added. Page 23, Appendix A added. Mapped IEEE 730 Standard requirements to EPA R5 G5M requirements. Corrected various spelling/punctuation items before finalized.
01-27-12	Cheryl Itkin .	Edits