

**Quality Assurance Project Plan for
Environmental Justice in Hydraulic Fracturing Analysis**

Prepared for:

U.S. Environmental Protection Agency
Office of Science Policy
Office of Research and Development

Prepared by:

Jessica Agatstein
MIT Washington Summer Intern
Office of Science Policy
Office of Research and Development

June 13, 2011

Distribution (Element A3)

This Quality Assurance Project Plan (QAPP) will be distributed to the staff members of the U.S. EPA as listed in *Table 1*.

Table 1: QAPP Distribution List.

Name Title	Contact Information
Jessica Agatstein Data Analyst and Project Lead	(202) 564-6658 agatstein.jessica@epa.gov
Jeanne Briskin Hydraulic Fracturing Study Plan Coordinator	(202) 564-4568 briskin.jeanne@epa.gov
Susan Burden Hydraulic Fracturing Data Analysis Technical Research Lead	(202) 564-6308 burden.susan@epa.gov
Steve Watkins OSP Quality Assurance Manager	(202) 564-3744 watkins.stephen@epa.gov
Steve Vandegrift Program Quality Assurance Manager	(580) 436-8684 vandegrift.steve@epa.gov

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

1. Project Management	5
1.1 Element A4: Project Organization	5
1.2 Element A5: Problem Definition and Background.....	6
1.2.1 Background	6
1.2.2 Statement of Key Questions and Project Objectives	6
1.3 Element A6: Project Description.....	7
1.3.1 Data Obtainment and Sources.....	7
1.3.1.1. Demographic Data	7
1.3.1.2. Hydraulic Fracturing Data	7
1.3.1.3. GIS Mapping Data	8
1.3.2 Data Analysis.....	8
1.4 Element A7: Quality Objectives and Criteria	9
1.4.1. Demographic Data Acceptance Criteria.....	9
1.4.2. Hydraulic Fracturing Data Acceptance Criteria.....	10
1.4.3. GIS Mapping Data Acceptance Criteria.....	11
1.5 Element A8: Special Training/Certification	11
1.6 Element A9: Documentation and Records.....	12
2. Data Generation and Acquisition	12
2.1 Element B5: Quality Control	12
2.2 Element B9: Non-direct Measurements	12
2.3 Element B10: Data Management.....	13
3. Assessment and Oversight	13
3.1 Element C1: Assessments and Response Actions	13
3.1.1. Data Quality Audit.....	13
3.1.2. Technical Systems Audit	13
3.2 Element C2: Reports to Management	14
4. Data Validation and Usability	14
4.1 Elements D1 and D2: Data Review, Verification, Validation, and Validation Methods.....	14
4.2 Element D3: Reconciliation with User Requirements.....	14

1. Project Management

This section addresses project management, including project background and purpose, roles and responsibilities, and key research questions and objectives. In its *Requirements of Quality Assurance Project Plans*,¹ EPA identifies the following nine project management elements:

- A1: Title and Approval Sheet,
- A2: Table of Contents,
- A3: Distribution List,
- A4: Project Organization,
- A5: Problem Definition and Background,
- A6: Project Description,
- A7: Quality Objectives and Criteria,
- A8: Special Training/Certifications, and
- A9: Documentation and Records.

Elements A1, A2, and A3 have already been provided in earlier portions of this document. The remaining elements are presented below.

1.1 Element A4: Project Organization

Project organization for this environmental justice analysis is depicted below in *Figure 1*. Jessica Agatstein will be responsible for the secondary data collection, analysis, and presentation, and will thus be responsible for ensuring that the quality of work meets the requirements of the EPA Hydraulic Fracturing Study. She will also keep the Project QA Officer, Steve Watkins, advised of any quality problems that arise in this study. The Project QA Officer will be responsible for maintaining QA activities and the official, approved QA Project Plan throughout the course of the project.

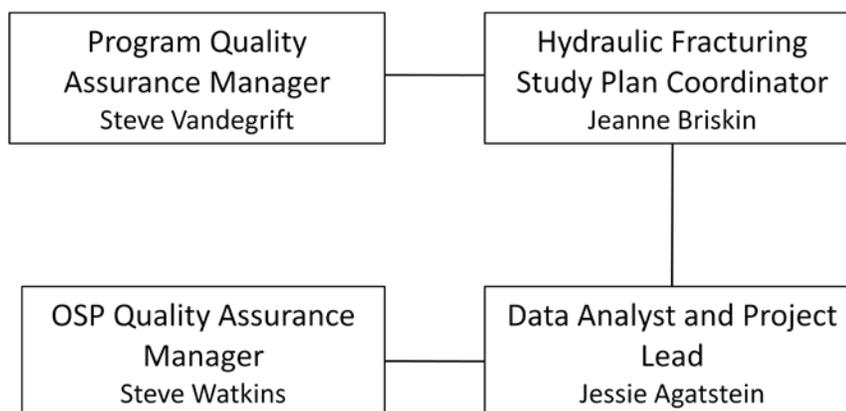


Figure 1: Environmental Justice in Hydraulic Fracturing: Project Organization Chart.

¹ EPA Requirements for QA Project Plans (QA/R-5) (EPA, 2001a).

1.2 Element A5: Problem Definition and Background

1.2.1 Background

In recent decades hydraulic fracturing has been increasingly used to access one of the nation's key energy resources, natural gas. In the process of hydraulic fracturing, natural gas or oil is extracted from so-called "unconventional" reservoirs via high-pressure injection of water, chemical additives, and proppants. In response to the growing use of fracturing in the United States, the U.S. Congress' Appropriation Conference Committee directed the EPA to research the potential impacts of hydraulic fracturing on drinking water resources in Fiscal Year 2010. In response to this direction, in February 2011 the EPA produced a *Draft Plan to Study the Potential Impacts of Hydraulic Fracturing on Drinking Water Resources* to be reviewed and commented on by the EPA's Science Advisory Board (SAB),² and the EPA expects to release research results as described in this study plan in 2012 and 2014. In addition to specifying many distinct research questions in this study plan, the EPA identified potential environmental justice concerns related to hydraulic fracturing as worthy of further research.

The EPA defines *environmental justice (EJ)* as "the fair treatment and meaningful involvement of all people regardless of race, color, national origin, or income with respect to the development, implementation, and enforcement of environmental laws, regulations, and policies."³ Since expanding the conversation on environmentalism and achieving environmental justice has been named one of the EPA's seven Agency-wide priorities, environmental justice research was included in the EPA's draft study plan.

1.2.2. Statement of Key Questions and Project Objectives

This project will satisfy the environmental justice research goals as outlined in the EPA's draft study plan. As described in this Quality Assurance Project Plan (QAPP), this project will use an index- and map-based analysis to screen for locations where hydraulic fracturing may be disproportionately co-located with environmental justice communities and then potentially identify locations for further study.

This project will answer the following key questions:

- Are particular populations, such as low-income, minority, indigenous, young, or elderly populations, disproportionately co-located with a high intensity of hydraulic fracturing activities?
- Are currently operating well sites disproportionately co-located with particular communities, such as minority, low-income, indigenous, young, or elderly communities?
- Is hydraulic fracturing wastewater disproportionately treated and disposed of (via publicly-owned treatment works (POTWs) or commercial treatment plants (CWTs)) in or near particular communities, such as minority, low-income, indigenous, young, or elderly populations?

² U.S. Environmental Protection Agency. *Draft Plan to Study the Potential Impacts of Hydraulic Fracturing on Drinking Water Resources*. EPA/600/D-11/001. Office of Research and Development, February 2011.

³ EPA's definition of EJ can be found at <http://www.epa.gov/environmentaljustice/basics/index.html>. EPA's definition of EJ was informed by E.O. 12898.

1.3 Element A6: Project Description

This project will use an index-based approach to compare a nationally representative set of well sites fractured between 2009 and 2010 (collected and analyzed by the EPA), wastewater treatment plants accepting hydraulic fracturing wastewater (collected by the EPA and PA DEP), and key demographic information (from the U.S. Census Bureau) on a county by county basis. The results of these comparisons will then be visualized and analyzed using Geographic Information System (GIS) mapping with ArcGIS software.

This project will be completed by August 2011. At the conclusion of this analysis, a series of maps will be produced in which hydraulic fracturing data and demographic data are compared at the regional and national scales. The key maps, and environmental justice analysis of those maps, will be incorporated into the EPA Hydraulic Fracturing Study Report in 2012.

1.3.1 Data Obtainment, Sources, and Rationale for Data Selection

1.3.1.1. Demographic Data

Demographic information, including information on race, income, and age, will be obtained from the U.S. Census Bureau. Race (and Hispanic or Latino) will be obtained from the 2010 Census by county. This information will be used to determine the percent of minority and indigenous residents present in each given county.

Income levels will be taken from the Census Bureau's 2009 Small Area Income and Poverty Estimates to determine average income levels of each county and the percent of residents below the poverty line in each county.

Percent of each county's population aged 5 or under or 65 and above will be obtained from the 2005-2009 American Community Survey's 5-Year Estimate.

1.3.1.2. Hydraulic Fracturing Data

Hydraulic fracturing data will be acquired from the EPA's Office of Research and Development. The EPA collected information on the location of well sites fractured by nine hydraulic fracturing service companies between September 2009 and September 2010 and then assessed the representativeness of that data. Though this data is measured in well sites per county, and thus is not at a particularly high resolution, it is the most representative data available on the location of hydraulically fractured wells. Although it will not represent the character of communities surrounding any one given well site precisely, this data will be useful for identifying potential "problem areas," and is thus consistent with this project's role as a first assessment.

A list of publicly owned treatment works (POTWs) accepting hydraulic fracturing wastewater will be obtained from the EPA's Office of Wastewater Management. The location of these POTWs will be obtained from state-produced GIS files.

1.3.1.3. GIS Mapping Data

The demographic and hydraulic fracturing data described above will be mapped with GIS using a base map obtained from the most recent 2010 Census TIGER/Line Data. All produced maps will be projected in the Global Coordinate System North American Datum of 1983 (GCS NAD83).

1.3.2. Data Analysis

The existing demographic and hydraulic fracturing data described above will be combined into a suite of indices to determine if any correlations exist between race, income, or age and a high number of hydraulic fracturing well sites within a given county. These indices will include the following:

- A *percent-based index* in which the number of well sites within a given county will be multiplied by that county's vulnerable population as a percent of total population (see *Equation 1*). This will result in a group of percent-based indices specific to one demographic risk-factor (such as non-white minority, low-income, or below 5 and above 65) that is *independent* of the total size of a given county's population. This index highlights co-localization of environmental justice communities and well sites, regardless of the size of those EJ communities.

$$(1) \quad \text{Percent-based index} = (\text{number of well sites}) \times (\% \text{ "vulnerable"}),$$

where "vulnerable" means non-white minority, low-income or below 5 and above 65.

- A *burden-based index* in which the number of well sites within a given county will be multiplied by that county's number of people within a vulnerable population (see *Equation 2*). This will result in a set of percent-based indices specific to one demographic risk factor (such as non-white minority, low-income, or below 5 and above 65) that is *dependent on* the total size of a given county's population. This index highlights the co-localization of highly populated environmental justice communities and well sites, focusing attention on locations in which a large number of people are co-located with hydraulic fracturing activity.

$$(2) \quad \text{Burden-based index} = (\text{number of well sites}) \times (\% \text{ "vulnerable"}) \times (\text{county population}),$$

where "vulnerable" means non-white minority, low-income, or below 5 and above 65.

- A *disparity index* in which the number of well sites within a given county will be multiplied by the difference between that county's vulnerable population and the average vulnerable population in the United States. This will result in a set of percent-based indices specific to one demographic risk factor (such as non-white minority, low-income, or below 5 and above 65) that directly identifies the disproportionate co-localization of well sites with vulnerable populations.

$$(3) \quad \text{Disparity index} = (\text{number of well sites}) \times (\% \text{ "vulnerable"} - \text{U.S. average of } \% \text{ "vulnerable"}),$$

where "vulnerable" means: non-white minority, low-income, or below 5 and above 65.

Once these indices are calculated, they will be mapped using ArcGIS software to determine if hydraulic fracturing well sites are co-located with environmental justice communities. A similar analysis will be performed using the location of POTWs and CWTs (instead of the location of hydraulic fracturing well sites), subject to data availability.

If the resulting maps do not show a great deal of variation in indices within a given region, then this analysis will indicate that there is not a disproportionate co-localization of well sites with EJ communities. However, if the resulting maps do show a great deal of variation in EJ indices within a given region, this analysis *may* indicate that there is a disproportionate co-localization of well sites with EJ communities. Unfortunately, there is no clear statutory line that separates an acceptable level of difference from disproportionate co-localization. Therefore, this analysis will either indicate no observed differences in well site co-localization between demographic groups or some observed differences that suggest the *possibility* of a disproportionate co-localization of well sites with specific demographic groups.

1.4 Element A7: Quality Objectives and Criteria

This QAPP aims to ensure that secondary data used for this evaluation are of sufficient quality necessary to achieve the Agency's priority of achieving environmental justice. This section addresses the quality criteria used to assess the adequacy of secondary data used in this project, as well as the uncertainty in the results derived from the use of these data sources.

All project results will include documentation of data sources and the assumptions and uncertainties inherent within that data, as well as computations and calculations made with secondary data. Data assessment and acceptance criteria for this project are outlined in the three sections below.

1.4.1. Demographic Data Acceptance Criteria

All demographic data will be obtained from the 2010 Census when possible to minimize uncertainties and error. If data sets are not available from the 2010 Census, demographic data will be obtained from the American Community Survey. See *Table 2* for data acceptance criteria that will be used to evaluate all demographic data.

Table 2. Demographic data acceptance criteria.

Acceptance Criterion	Description/Definition	Specification
Timeliness	Are the data current?	What time period is covered by data? Does this time period correspond to time period of the well locations data?
Internal consistency	For an individual data source, do national numbers equal the sum of all sub-national numbers?	Does it appear that parts of the data are missing or inconsistent?
Completeness	For a given data source, does it cover the entire population of a given geographic unit of analysis?	There should be documentation of whether or not the reported data cover the entire geographic unit of analysis or are restricted to specific locations.
Representativeness	For a given data source, does it represent the distribution of ethnic, income, or age minorities with a sufficient level of confidence?	There should be documentation of level of confidence and uncertainty from the American Community Survey..

1.4.2. Hydraulic Fracturing Data Acceptance Criteria

All hydraulic fracturing data will be obtained from the EPA, as described in section 1.3.1.2. *Table 3* further describes the data acceptance criteria that will be used to evaluate all hydraulic fracturing data.

Table 3. Hydraulic fracturing data acceptance criteria.

Acceptance Criterion	Description/Definition	Specification
Timeliness	Are the data current?	What time period is covered by data?
Internal consistency	For an individual data source, do national numbers equal the sum of all sub-national numbers?	Does it appear that parts of the data are missing or inconsistent?
Completeness	For a given data source, does it cover the entire country?	There should be documentation of whether or not the reported data cover the entire country or are restricted to specific locations.
Representativeness	Do the EPA/Westat data represent the distribution of well sites within a given county or region with a sufficient level of confidence?	There should be documentation of whether the reported data are representative of the distribution of wells in a given county.

1.4.3. GIS Mapping Data Acceptance Criteria

All county and state boundaries will be obtained from the 2010 Census TIGER/Line Shapefiles. *Table 4* further explains the data acceptance criteria that will be used to evaluate GIS mapping data.

Table 4. GIS maps acceptance criteria.

Acceptance Criterion	Description/Definition	Specification
Timeliness	Are the data current?	What time period is covered by data?
Internal consistency	For an individual data source, do national numbers equal the sum of all sub-national numbers? Within a given map, are all shapefiles projected in the same way?	Does it appear that parts of the data are missing or inconsistent?
Completeness	Do all data sources cover the entire country?	There should be documentation of whether or not the reported data cover the entire country, or if some are restricted to specific locations.
Comparability	Are the data from different data sources consistent?	Inspect the units of analysis and projections of data sources for consistency.

1.5 Element A8: Special Training/Certification

During the course of this project, Ms. Agatstein will access and analyze TSCA CBI data. Throughout this project Ms. Agatstein will adhere to CBI procedures when handling confidential information, and will manage all reports, documents, and other materials developed in accordance with the procedures set forth in EPA's TSCA CBI Protection Manual.⁴ Ms. Agatstein will maintain active TSCA CBI clearance, and all work involving TSCA CBI will be completed on the approved TSCA CBI computer assigned to her.

1.6 Element A9: Documentation and Records

All project documents will be submitted to the Hydraulic Fracturing Study Coordinator, Jeanne Briskin. The final report will include a detailed description of the analytical methods used to produce all map products, as well as any assumptions or uncertainties inherent in those methods.

The final report will contain an appendix listing the specific Census databases used and the locations of those databases online. This appendix will also specify any modifications that were made to the original

⁴ U.S. Environmental Protection Agency, *TSCA CBI Protection Manual*, EPA Office of Pollution Prevention and Toxics (7407 M), October 20, 2003.

databases obtained from the Census Bureau and EPA. All errors and uncertainties associated with each data set will be documented and included in this section.

The report will contain a second technical appendix that clearly documents how all indices were calculated and how information was entered into and manipulated in ArcGIS. All original GIS Shapefiles, demographic databases, and index data sets will also be attached to the final report on a CD disk and retained by the study coordinator.

2. Data Generation and Acquisition

This section addresses data acquisition and management activities, including the following elements identified by EPA:

- Element B5: Quality Control
- Element B9: Non-direct Measurements
- Element B10: Data Management

2.1 Element B5: Quality Control

All of the data used in this project will meet the criteria listed in *Tables 2, 3, and 4*, and will be examined to ensure that they meet these criteria at each stage (data acquisition, manipulation, mapping, and analysis).

2.2 Element B9: Non-direct Measurements

All data used in this project will be obtained from existing databases managed by the U.S. Census Bureau and the EPA. The specific sources and their intended uses are listed in detail in Section 1.3.1 and summarized below.

- Demographic data
 - Race and ethnicity will be obtained from the 2010 Census.
 - Income levels will be obtained from the Census Bureau's 2009 Small Area Income and Poverty Estimates.
 - Age levels will be obtained from the 2005-2009 American Community Survey's 5-Year Estimate.
- Hydraulic fracturing data
 - Number of well sites per county will be obtained from the EPA's 2010 information requests to nine hydraulic fracturing service companies.
 - POTWs accepting hydraulic fracturing wastewaters will be obtained from the EPA's Office of Wastewater Management.

- Location of POTWs accepting hydraulic fracturing wastewaters will be obtained from the Pennsylvania Department of Environmental Protection via the Pennsylvania Geospatial Technologies Office.
- GIS mapping data
 - County lines and projections will be obtained from the Census Bureau's 2010 TIGER/Line Shapefiles.

This data will be used to calculate county-level EJ indices (as described in Section 1.3.2) and then mapped using ArcGIS to screen for locations where hydraulic fracturing may be disproportionately co-located with environmental justice communities.

2.3 Element B10: Data Management

The data under this task will be maintained both in Excel files and Shapefiles to allow for ease of analysis in both programs. Variable names will be created and managed to match standard U.S. Census variable names, when applicable.

Some of the data used for this project will be TSCA CBI. All such data and products utilizing this data will be managed following the procedures set forth in EPA's TSCA CBI Protection Manual. The project lead, Ms. Agatstein, will maintain active TSCA CBI clearance, and will use a TSCA-compliant computer whenever working with TSCA CBI data.

3. Assessment and Oversight

This section describes the audits and other assessments needed to determine whether this QAPP is being implemented as approved and to increase confidence in the information obtained and produced as a result of this project.

3.1 Element C1: Assessments and Response Actions

All work conducted for the environmental justice project will be subject to technical review by EPA Hydraulic Fracturing Study Coordinators Jeanne Briskin and Susan Burden. Steve Watkins will serve as the QA Officer for this project and will review this QAPP for completeness and applicability. He will be available to assist Ms. Agatstein with QA issues as they arise and will periodically review compliance with this QAPP. This project will also undergo periodic data quality audits and technical systems audits, as described below.

3.1.1. Data Quality Audit

Data quality audits will evaluate whether the data used are the most appropriate and of the highest level of specificity available for the environmental justice analysis. Additionally, the data quality audits will ensure that data used are from valid sources.

3.1.2. Technical Systems Audit

A technical systems audit will occur toward the beginning of the environmental justice project to ensure that the most appropriate methods and models are employed in the analysis, and that the data are being handled in a manner consistent with TSCA CBI requirements.

3.2 Element C2: Reports to Management

Ms. Agatstein will supply a draft report on the map products and findings of this analysis to the hydraulic fracturing study team members Jeanne Briskin and Susan Burden for comment by July 15, 2011, and will then incorporate their comments into a final report that will be given to the study team by August 1, 2011. Ms. Agatstein will keep the study team involved through weekly technical progress updates in which she will describe any problems encountered and solicit feedback as necessary to ensure quality of the finished product.

4. Data Validation and Usability

This section addresses the quality of the completed final report to see if this product will conform to the objectives outlined in this QAPP, especially given this project's use of existing data sets.

4.1 Elements D1 and D2: Data Review, Verification, Validation, and Validation Methods

All data will be reviewed for completeness, representativeness, and statistical certainty. They will be examined to determine whether the data cover a representative portion of the counties in which hydraulic fracturing activity was reported by nine hydraulic fracturing companies between 2009 and 2010. She will also examine the uncertainties of existing data sets to ensure that all data is of sufficient quality to adhere to the criteria outlined in this QAPP.

All final map products will be examined to ensure that all data sets are correctly and clearly displayed on regional or national maps. In addition, all data sets and produced maps will be reviewed for apparent outlier values, which will be examined to determine whether these are indeed true values, the result of data entry errors, or have some other explanation. These outliers will be discussed in the final report, as they may either indicate data errors or may indicate the key regions in which environmental injustices may be occurring.

4.2 Element D3: Reconciliation with User Requirements

The checks that will be used to determine the timeliness, consistency, completeness, and comparability of all existing data and final products are described in Section 1.4. These quality measures will be reported in all project deliverables, which will allow the hydraulic fracturing study team and later data users to determine if the data are of sufficient quality for other uses.

Ms. Agatstein will work with the study team and the QA staff to determine to what extent the data that do not meet the specified data acceptance criteria may be used to support further study and how this determination will be documented. In addition to an evaluation of data quality, Ms. Agatstein will identify data sources, assumptions made, changes or modifications to data, and calculations used in their development in the draft and final environmental justice reports. These identifications will be sufficiently detailed and transparent to ensure the reproducibility of the work by third parties.