



HHS Public Access

Author manuscript

Sci Total Environ. Author manuscript; available in PMC 2018 January 15.

Published in final edited form as:

Sci Total Environ. 2017 January 15; 577: 195–201. doi:10.1016/j.scitotenv.2016.10.160.

Association of groundwater constituents with topography and distance to unconventional gas wells in NE Pennsylvania

Beizhan Yan^{1,*}, Martin Stute^{1,2}, Reynold A. Panettieri Jr.³, James Ross¹, Brian Mailloux², Matthew J. Neidell⁴, Lissa Soares², Marilyn Howarth⁵, Xinhua Liu⁶, Pouné Saberi⁵, and Steven N. Chillrud¹

¹Lamont-Doherty Earth Observatory of Columbia University, 61 Rt. 9W, Palisades NY 10964

²Department of Environmental Science, Barnard College, NY, NY

³Institute for Translational Medicine and Science, Rutgers Biomedical and Health Sciences, Rutgers University, New Brunswick, NJ

⁴Department of Health Policy and Management, Mailman School of Public Health, Columbia University, New York NY

⁵Center of Excellence in Environmental Toxicology (CEET), University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA

⁶Department of Biostatistics, Mailman School of Public Health, Columbia University, New York NY

Abstract

Recently we reported an association of certain diseases with unconventional gas development (UGD). The purpose of this study is to examine UGD's possible impacts on groundwater quality in northeastern Pennsylvania. In this study, we compared our groundwater data (Columbia 58 samples) with those published data from Cabot (1701 samples) and Duke University (150 samples). For each dataset, proportions of samples with elevated levels of dissolved constituents were compared among four groups, identified as upland far (i.e. 1km to the nearest UGD gas well), upland near (<1km), valley far (> 1 km), and valley near (< 1 km) groups. The Columbia data do not show statistically significant differences among the 4 groups, probably due to the limited number of samples. In Duke samples, Ca and Cl levels are significantly higher in the valley near group than in the valley far group. In the Cabot dataset, methane, Na, and Mn levels are significantly higher in valley far samples than in upland far samples. In valley samples, Ca, Cl, SO₄, and Fe are significantly higher in the near group (i.e. < 1 km) than in the far group. The association of these constituents in valley groundwater with distance is observed for the first time

Reprints and permissions information is available at <http://www.journals.elsevier.com/science-of-the-total-environment>.

*Correspondence and requests for materials should be addressed to yanbz@ldeo.columbia.edu; 845-365-8448 (office); 845-365-8155 (fax).

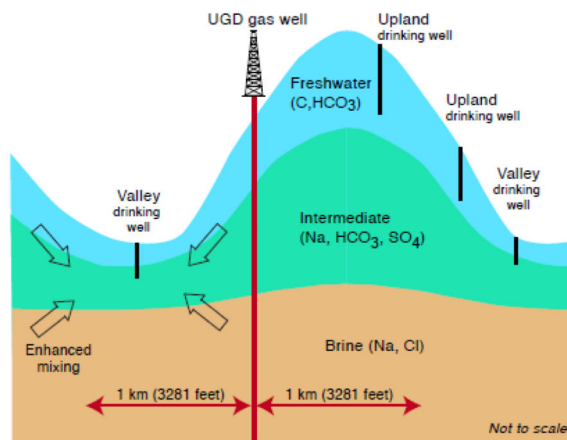
Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published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

Appendix is linked to the online version of the paper at <http://www.journals.elsevier.com/science-of-the-total-environment>

The authors declare no competing financial interests.

using a large industry dataset. The increase may be caused by enhanced mixing of shallow and deep groundwater in valley, possibly triggered by UGD process. If persistent, these changes indicate potential for further impact on groundwater quality. Therefore, there is an urgent need to conduct more studies to investigate effects of UGD on water quality and possible health outcomes.

Graphical abstract



Keywords

Unconventional gas development; hydraulic fracturing; Marcellus Shale; groundwater; environmental impacts; distance to gas wells

1. Introduction

The combined use of horizontal drilling and hydraulic fracturing (HF) techniques has led to rapid growth of unconventional gas development (UGD) throughout the world (Annevelink et al., 2016; McGlade et al., 2013). Globally there is about 193 trillion m^3 (Tcm) of shale gas resources (McGlade et al., 2013); Central and South America has 35.6 Tcm, followed by Africa (29.3), the U.S. (19.3), China (17.8), Canada (12.0), Mexico (11.4), Australia (11.2), etc. Despite available resources in many of these countries/regions, only the U.S., Canada, and China produced shale gas production at a commercial level as of 2015 data (EIA, 2016), with the U.S. producing 37 billion cubic feet per day (Bcf/d), Canada 4.1 Bcf/d, and China 0.5 Bcf/d. Based on China's Energy Development Strategic Action Plan (2014-2020), China seeks to largely increase its ability to produce shale gas, reaching about 3 Bcf/d in 2020 (Council, 2014). Many other countries appear to be holding off on UGD development, possibly given its impacts on environment and health remain unclear. Several environmental health studies have linked adverse health effects, such as asthma, increase of congenital heart defects of newborns, and hospitalization rates, to active UGD sites (Jemielita et al., 2015; McKenzie et al., 2015; Rasmussen et al., 2016; Stacy et al., 2015), although the cause of the association is not well understood. The exposure measures in these studies were based on the distance from the patient's home to the gas well or gas well density in the zip code of the patient's home.

The impact of UGD on water quality has been hotly debated for the past several years (Gorody, 2012; Hildenbrand et al., 2016; Jackson et al., 2013a; Jackson et al., 2013b; Meng, 2015; Molofsky et al., 2013; Olmstead et al., 2013; Siegel et al., 2015a; Siegel et al., 2015b; U.S.EPA, 2015; Vidic et al., 2013; Warner et al., 2012). Multiple reasons contribute to uncertainty, including the lack of systematic surveys of water quality pre- and post- drilling/production, substantial dilution and slow transport over months to hundreds of years along flow paths from contaminated sites to private wells or public water supplies, and natural fluctuations of groundwater chemistry. The groundwater quality in private wells can be impacted by myriad factors, including the geological source formation, topography (valley or upland) of the well site, groundwater flow path, recent amount of precipitation, productivity and the amount of daily water usage, other potential contamination sources (e.g., sewage, road spills of chemicals), wellbore/casing integrity, and well depth of private wells. For example, methane levels in groundwater (GW) have been found to be statistically higher in valleys and/or under confined groundwater conditions (Heisig and Scott, 2013). Uplands are groundwater recharge areas, while valleys tend to be discharge areas located at the end of flowlines with more evolved geochemistry. Because of these complexities, studies with a limited number of water samples tend to be inadequate to determine a well-defined impact. However, conducting a large study remains challenging (Burton et al., 2016; Drollette et al., 2015; Jackson et al., 2013a), as was illustrated in the recent EPA External Review Draft of the assessment of UGD on drinking water quality, which recognized the limitation due to lack of data (U.S.EPA, 2015).

Stimulated by Pennsylvania DEP regulations (Oil and Gas Act 13), industry began collecting and analyzing thousands of “pre-drilling or pre-alteration” water samples from private water wells if they are within 0.76 km (or beyond for certain operators) of proposed gas wells (Molofsky et al., 2013; Siegel et al., 2015b). Sharing such data will be a valuable endeavor to characterize environmental and health risks. Government institutions (e.g., the USGS and Pennsylvania Department of Conservation and Natural Resources, DCNR) have provided large datasets relative to water chemistry in their reports in northeastern PA (e.g., (Reese et al., 2014)), but their focus is not on water chemistry associated with UGD (Bowen et al., 2015).

Fortunately, publications require a release of the data used in articles. For example, Molofsky et al. (2013) released the data of methane and certain major and trace ions of 1701 water wells sampled in Susquehanna County collected by Cabot Oil & Gas Corporation. Although these 1701 samples are regarded as “pre-drilling” (i.e. prior to the drilling of specific proposed gas wells), about 20% (322 samples) were collected within 1 km of other gas wells drilled prior to the time of sampling, and therefore, based on the 1 km distance used by prior studies, can also be considered as potentially affected samples, i.e. “post-drilling or post-alteration” samples relative to the developed gas wells (Molofsky et al., 2013).

The ultimate goal of this study is to assess whether UGD possibly impacts groundwater quality. We hypothesized that levels of constituents in groundwater are associated with the distance to UGD gas wells. To test this hypothesis, over the past three years, we have collected 58 samples from private wells in NE Pennsylvania. Concurrently, realizing that

large datasets were needed for identifying the impacts, we started to analyze available data from other studies. In this study, we examined results from Duke University (150 samples) (Warner et al., 2012) and Cabot Oil & Gas (1701 samples) (Molofsky et al., 2013), which were both collected from private wells in the NE Pennsylvania. The Duke group has expanded their studies to include more samples (Darrach et al., 2014; Jackson et al., 2013a; Warner et al., 2013), but data of dissolved constituents of interest are not available online. Given the well-known difference in water quality between upland and valley groundwater, for each dataset, we grouped samples into 4 categories based on their topography (upland or valley) and the distance to the nearest gas well (< 1km and 1km) to investigate impacts of distance and/or topography.

2. Methods

Table S1 briefly compares sampling and analytical methods of the three studies (Columbia, Duke, and Cabot). Detailed methods used in the Duke and Cabot studies were described previously (Molofsky et al., 2013; Warner et al., 2012). The Columbia samples were collected through two related projects, a NY-PA comparison study (NPCS) and a 100-bottle by mail campaign. Participants were recruited by the Community Outreach & Engagement Cores (COEC) of the Center of Excellence in Environmental Toxicology (CEET) at the University of Pennsylvania (CEET, 2015). Participants were contacted by COEC first by email or phone and those interested were subsequently contacted by the Columbia field team.

For Columbia NPCS study (22 samples), a Teflon sampling line with pressure gauge was connected to the spigot at the pressure tank and the groundwater was run through the line until basic water parameters (pH, T, specific conductivity, etc.) reached stable conditions, normally taking about 10-20 minutes. Samples for elemental analysis by inductively coupled plasma mass spectrometry (ICP-MS) were collected into a 30 mL acid-prewashed HDPE bottles and acidified to pH 0.8 shortly after collection. For anion analysis by ion-chromatography (IC), samples were collected in a 125 mL HDPE bottle pre-washed with deionized water. The 100-bottle project samples (36 samples) were collected by residents from their own well. The purpose of this study is to evaluate the feasibility of the self-collection and mailing approach for future large studies. Sampling kits were mailed to participants, in which four pre-treated plastic bottles (two bottles for IC analysis and two for ICP-MS analysis), an explanation letter of the project, step-by-step hard copy instructions, and a prepaid shipping box with returning address were included. Participants were requested to collect samples from the spigot at the pressure tank and kitchen faucet, respectively. After logging sampling information, participants were asked to promptly return these samples to Columbia University. Upon receipt, samples for ICP-MS analysis were acidified. All samples including those for IC analysis were stored at 4 degree C until analysis, which was typically conducted within 1 month. We only included samples collected from the spigot at the pressure tank in this study in order to minimize the possible interference in results by water treatment. About 20% of homes investigated installed water treatment systems (e.g., the filtering system to remove Fe and Mn precipitation) between the pressure tank and kitchen faucet. Five wells were collected by both the NPCS and 100 Bottle Projects, allowing a comparison between different sampling approaches. In addition, three

wells were collected repeatedly in order to understand the seasonal change of groundwater. For those wells collected for multiply times, only one sample was included in this study. In total, 58 samples were included in this study and the majority of samples were collected in summer or fall.

Definition of valley and upland: In the Cabot study, samples within 1000 feet (~ 0.3 km) of a major National Hydrography Dataset (NHD) flowline or 0.15 km of minor tributaries to an NHD flowline were categorized as “valley.” Everything else was defined as “upland.” This definition used in the Cabot study was a practical method for a large dataset to quickly categorize topography. In order to compare the results among the three studies, we also adopt this definition for the Columbia and Duke samples.

Data analysis

Given the fact that some private wells (5 to 10) were sampled by both CU and Duke field teams, and possibly by Cabot as well, where the overlapped samples were not identifiable, we cannot pool the three datasets for analysis. The ion charge balance of major constituents was evaluated by calculating a percent charge error (Fritz, 1994; Siegel et al., 2015b). Also the distribution of elements was compared for the three data sets. For each of the three datasets, samples were categorized into four groups based on their topography and the distance to the nearest UGD gas well: upland far (i.e. > 1km), upland near (i.e. <1km), valley far, and valley near. Concentration threshold of dissolved constituents (both major and minor) were selected (criteria are detailed in next paragraphs) and the proportion of samples exceeding the specific threshold (e.g., > 20 mg/L for Na) in a group was computed, similar to the method used in Siegel et al. (2015b)

For each dataset, a Chi-square test was used to detect differences in the proportion of high levels of constituents among the four groups. The null hypothesis is that all four groups have the same proportion of high levels (i.e., no differences among groups), while the alternative hypothesis is that not all groups have the same proportion of high levels (some differences among groups). If the null hypothesis of no difference among all four groups was rejected, then the alternative hypothesis would be in favored, indicating effect of distance and/or topography.

To control for false positive error rate in multiple testing of null hypothesis (e.g., 14 constituents was tested in Cabot data), the false discovery rate (FDR) was calculated based on adjustment of p-values using an approach developed by Benjamini and Hochberg, 1995. Constituents with FDR < 0.05, which indicated the existence of group difference, were further compared using the ratio of proportions (RP). RP was computed between specific groups (e.g., valley near vs. upland far) and the generalized linear model was used to describe the pattern of RP and calculate the 95% confidence interval of the RP. Note that RP cannot be interpreted as relative risk because proportion of high level should not be interpreted as risk, most thresholds we used in this study were lower than EPA guidelines (Table 1), and certain major ions such as Ca are not regulated by the EPA and have not been found to be toxic.

The threshold for each constituent (Table 1) was selected based on its value at the 75th percentile (except methane) in the Cabot data, or the limit of detection (LOD) if more than 75% of samples were below the LOD. Taking Mn as an example, about 79% of samples were below 0.025 mg/L, its LOD value in the Cabot data, so the LOD was used as the threshold for Mn. We also tested 50th percentile and 90th percentile as thresholds, but neither were optimal for our research goal, i.e., investigating differences between the four groups to bring out associations with distance and topography. The 50th percentile was not sensitive enough to discern differences among groups, and the 90th percentile was set too high for certain constituents (e.g., Ca, and SO₄). Because levels of these constituents are relatively high in the Valley Near group, if 90th percentile was used for threshold, the majority of samples above the threshold would be in this group, leaving only a few or no samples in other groups and leading to findings with lesser statistical power. The 75th percentile setting was suitable for the threshold value for all constituents except methane after considering both sensitivity and statistical power issues. Methane levels in groundwater varied largely by season and sampling methods (e.g., pressure tank, kitchen tap, or garden hose) (Molofsky et al., 2016), in this study, we selected a relatively high threshold (e.g., 1 mg/L of methane) to minimize these artifacts. Based on our field experience, methane at this level typically can be measured even at different sampling times and by different sampling and lab analytical methods. We did not use the EPA Maximum Contaminant Level (MCL) or Secondary Maximum Contaminant Level (SMCL) as the thresholds, since the goal was to look for relative differences in water chemistry, not on immediate health concerns.

3. Results

In all three studies, major dissolved constituents (Ca, SO₄, Cl, Mg, and Na) had similar and skewed distributions and box plots showed similar concentration ranges of these major chemicals among the three studies (Fig. 1). Given that all samples were collected from the NE Pennsylvania region, this similarity lends confidence in the quality of the data sets. Charge balance of major anion and cations is another approach to examine data quality. Two samples in the Cabot data set with >30% charge-balance errors were not included in further analyses. After removing the 2 outliers in the Cabot data, the mean of the ion charge balance for the 1699 remaining samples was $-0.45 (\pm 10.64)\%$ and the mean of the absolute ion charge balance was $7.72 (\pm 6.43)\%$, comparable to numbers listed in (Fritz, 1994). Table 2 lists the proportion of samples with levels above the threshold for 14 constituents among the four groups in the Cabot dataset. Four constituents, magnesium (Mg), lead (Pb), Alkalinity (Alk), and Turbidity (Turb) had $FDR > 0.05$ (Table 2), indicating that they do not have statistically significant different among the four groups.

With the remaining 10 constituents in Cabot dataset, we use ratio of proportions (RP) and its 95% confidence interval (CI) for comparison between two groups (Fig. 2). An RP above 1 indicates that the comparison group (numerator) has a higher proportion of levels above the threshold than the reference group (denominator); if a 95% CI includes the value 1, then the proportion of levels above the threshold is not significantly different between the two comparison groups. Among the 10 constituents, the change in RP of barium (Ba) reverses the direction of sulfate (Fig. 2), consistent with Ba levels being affected by sulfate level in GW (more details in SI). Figure 2a shows the association of RP with elevation in samples

1 km; RP in valley samples is significantly above 1 for Na, Mn, specific conductivity (Cond), and methane. We also examined the effects of distance in both upland samples (Fig. 2b) and valley samples (Fig. 2c). For the upland samples, RP is significantly above 1 for Na and Cl, near vs. far, while for valley samples, constituents Ca, Cl, SO₄, Fe, TDS, and Cond significantly differ with distance, indicating higher levels in the valley near group than the valley far group. Fig. 2d shows effects of the combination of both the topography and distance; RP is significantly above 1 for SO₄, Cl, Fe, Methane, TDS, and Cond.

Table 3 shows the results of Columbia and Duke data. The proportions vary among groups; in Columbia set (58 samples), none of these differences are statistically significant, suggesting that the small sample size does not provide enough power to detect group differences. For the Duke data set (150 samples), Ca and Cl significantly differ among groups (Table 3), even after adjusting for multiple testing. Similar to Cabot results, in Duke valley samples, RP of Ca and Cl is significantly higher in the near group than in the far group (Table 4). Again the other comparisons seen in the Cabot data are not significant in the Duke data probably because the sample size is still too small.

4. Discussion

In nature, groundwater (GW) evolves along its pathway from upland to valley (Heisig and Scott, 2013); upland GW is relatively fresh with high concentration of Ca and carbonate, while valley GW typically contains more dissolved constituents (e.g., Cl, Na, and SO₄). Due to redox reactions along the pathway, Fe and Mn can be reduced from high valence to low valence states, the latter having increased solubility. The data showed in Fig. 2a are consistent with this general pattern— in the valley samples as compared to the upland groups, the RP of Na, Mn, methane and specific conductivity (Cond) are significantly higher and Cl, Fe, and total dissolved solids (TDS) are also higher but at borderline significance. The increase of Na, Cl, etc. in valley samples leads to the observed increase in Cond and TDS.

The association with distance to the nearest UGD well differs by topography. For upland samples, the majority of investigated constituents except Ba are not associated with distance (Fig. 2b), while for valley samples, many more chemicals, including Ca, Cl, SO₄, Ba, Fe, TDS, and Cond, show association with distance (Fig. 2c). This indicates that UGD can impact levels of dissolved constituents in valley groundwater, but not so much in upland groundwater. Comparison between Fig. 2a and Fig. 2d shows the combined effects of distance and topography. In the far group comparison (Fig. 2a), levels of Na, Mn, Cond, and methane are higher in valley samples than in upland samples (i.e., RP>1); levels of these constituents remain higher in group comparison between valley near vs. upland far (Fig. 2d), in which both distance (near vs. far) and topography (valley vs. upland) changed; however, levels of Ca and SO₄ were lower in valley than in upland in far group comparison (Fig. 2a) but higher in Fig. 2d. The level of Cl in the valley samples changes from insignificantly to significantly higher. Given that groundwater in valley settings can be affected by both shallow freshwater and underlying brine layers (Siegel et al., 2015a, Fig. 4), the increase of both freshwater Ca and SO₄ and Cl (from underlying brine layers) in the valley near group indicates the increase of mixing of freshwater or deep brine in these valley samples..

The reason for the possible enhanced mixing in valley <1km samples remains unclear. One possibility is that a reservoir-scale dilation triggered by the injection of HF fluid and contraction of reservoir by removal of shale gas and produced water may lead to some stress changes that propagate up to the surface and initiate mixing. It is well known that injection and extraction from oil and gas fields can promote differential subsidence, horizontal displacement and surface faulting, even for those fields with median depth of production >7000 ft (Nagel, 2001; Yerkes and Castle, 1969). This disruption to subsurface and surface layers near gas wells may lead to enhanced mixing in valleys. Compared to upland areas, valley zones typically have a greater fracture permeability of bedrock, possibly due to pre-existing deep-seated faults (Heisig and Scott, 2013), the denser networks of hydraulic fracturing horizontal lines close to the gas well may lead to dilation and contraction of disk or donut-shaped shale reservoirs (centered in the drilling borehole). We are currently exploring various remote sensing methods to examine the possible surface deformation in eastern PA. Plausibly, enhanced permeability by induced microseismicity may also explain the mixing phenomenon. Both an increase in permeability and change in water level have been observed after earthquakes (Rojstaczer and Wolf, 1992), but because of the very small magnitude of induced microseismicity associated with hydraulic fracturing, it is unlikely that any single microseismic event causes any noticeable impact (Abdulaziz, 2014). However, the cumulative effect of repeated high frequency, low magnitude microseismic waves over the relatively long time (months compared to seconds to hours of earthquake) related to UGD is unknown.

The dilation force or the microseismic vibration associated with UGD should reach both upland and valley environments, but the impact on valley appears to be more noticeable. Upland groundwater typically has a steep hydraulic gradient, preventing considerable mixing of brine from deep layers, while valley groundwaters have a much lower hydraulic gradient, allowing various water sources to mingle there, e.g., freshwater from upland, freshwater from the nearby river/creeks, saline water from deep layers, or a mixture of these sources (Heisig and Scott, 2013).

We cannot rule out other pathways that might possibly lead to observed differences as well. The enhanced mixing might happen in the first place while drilling through aquifers and other deep subsurface geological layers. Wellbore integrity has been long a concern, and the rate of the loss of structural integrity for unconventional wells in Marcellus shale inspected in Pennsylvania was estimated to be about 6.3% (Davies et al., 2014). Affected groundwater can be then transported down hydraulic gradient to nearby private wells. To investigate this potential mechanism, the locations and depths of these private wells would be needed, so that the specifics of their nearby gas wells and aquifer characteristics can be examined. Unfortunately this level of information is not available for us in the large Cabot data set.

Some constituents that have EPA guideline limits (e.g., SO₄, Cl, and Fe) show associations with distance in valley samples. Compared to the results in Siegel et al., 2015a, which is an industry funded project, the percentage of Cabot samples exceeding EPA guideline levels is relatively low, possibly because a fraction of samples from the Cabot study were filtered to remove particles (before 2011), while the Chesapeake samples were unfiltered (Siegel et al., 2015a). For example, in the Cabot data set, about 0.9% of samples exceed EPA Pb guideline

levels (0.015 mg/L), compared to 4.4% of the Chesapeake samples. Turbidity is also regulated by EPA guidelines, but its association with distance in valley samples is not significant (Fig. 2c). Another potential contaminant, arsenic (As), is not included in this statistical analysis because the Cabot data have only a small fraction (~7%) of samples with As levels above the detection limit (0.004 mg/L) and only 10 samples are above 0.01 mg/L, the EPA MCL of As. Cl, Fe, TDS, and SO₄ are included in the list of constituents measured that have Secondary Maximum Contaminant Limits (SMCL) and their levels are associated with distance (Table 1). EPA has a Health Advisory level of 20 mg/L for Na (for individuals on a 500 mg/day restricted sodium diet), which is an estimated level of a substance acceptable in drinking water based on health effect information and these are not legally enforceable (EPA 2009). In the valley samples from the Cabot dataset, Na levels are not significantly affected by distance from a UGD well. The Na levels in valley samples with less UGD effect (i.e., the valley far group) was already high, thus depending on the mixing ratio, mixing of freshwater (low Na level) and saline water (high Na level) does not necessarily lead to an increase of Na levels. In addition, Na can be involved in an ion-exchange reaction (Siegel et al., 2015a), further complicating the Na results.

Changes in levels of certain constituents (e.g., Fe and dissolved oxygen), if persistent, has the potential to mobilize other geogenic elements of concern such as As. In addition, if the observed associations for certain constituents are due to increased mixing, then these constituents may be early indicators of more changes to come. In the Barnett shale, which has been more densely drilled than the Marcellus Shale, the maximum As concentration from an active extraction area was substantially higher than both the maximum concentration among the nonactive/reference area samples and historical levels from this region (Fontenot et al., 2013). The authors suggested that small perturbations such as lowering of the water table either through groundwater withdrawals or drought conditions might explain these results.

Our work highlights the importance of sample size for teasing out conclusions, which is consistent that a large number of processes can impact groundwater chemistry thus requiring larger data sets to tease these impacts out. We also conducted statistical analyses in the combined data set from all three studies, neglecting the issue that a small number of water samples may be collected from the same wells by different teams. Major conclusions remain the same. The CU data alone (58 samples) with small group size (n=5, 10, 18, 25) did not show statistically significant differences among the 4 groups, compared to many significant associations observed in Cabot data. Similarly, in Duke dataset (150 samples), the differences in RPs of major constituents, except Ca and Cl, are insignificant. The high level of Cl in valley samples is consistent with observation in other studies (Heisig and Scott, 2013; Poth, 1973). Similar to Cabot results, proportion of elevated Ca and Cl levels in valley sample is significantly higher in near group than in far group; however, the Duke result would be more robust if there were more than 7 samples in this group (Table 4).

Using Cabot data we found that levels of certain constituents is associated with the distance to UGD gas wells and that groundwater in valley settings is more subject to change than in upland settings. These findings were not recognized by the original authors of the Cabot study and have not been reported in other studies. Compared to other studies (Molofsky et

al., 2013; Siegel et al., 2015a; Siegel et al., 2015b; Warner et al., 2013), we recognized that for some GW samples, pre-drilling data can be used as post-drilling data as well due to nearby UGD. We also minimized the topographic effects by separating samples into four groups based on topography and distance. In addition, we found that due to its complexity of groundwater, the sample size is important in order to obtain statistically powerful findings.

These observed associations cannot be deemed as causal relationships. To demonstrate causality, ideally, comparison of water quality pre- and post- UGD is needed. In an ongoing project, we are conducting a focused study to investigate the possible changes in water quality between pre- and post- UGD. In addition, industry has collected thousands of pre-drilling samples and many of these samples have a post-drilling comparison. Hopefully this study can encourage the industry to release these comparison data after de-identification. We also hope findings from this study can motivate funding agencies and industry to provide support for more studies to investigate whether the findings can be repeated in other areas, and if so, also fund studies focused on causes of these changes and whether they can affect long-term groundwater quality and health outcomes.

Shale gas is expected to be developed in many other countries/regions besides North America (Council, 2014; McGlade et al., 2013). To the best of our knowledge, the majority of groundwater studies have been conducted in the U.S., and results are often controversial. We hope this controversy can make other countries more prudent regarding to UGD and encourage them to conduct comprehensive pre-UGD baseline studies to be able to evaluate possible environmental and health impacts associated with UGD in their own countries. Furthermore, some countries (e.g., China) already have severe groundwater contamination issues from other sources, making it urgent to investigate whether the groundwater quality can get even worse by possible enhanced mixing associated with UGD.

We acknowledge limitations of this study as well. Since the Cabot data were generated by several different labs, inconsistency of data quality is possible. We do not have the distance to the National Hydrography Dataset streamlines and locations of each sample and the definition of valley and upland may not be accurate. We believe this definition underestimates the sample numbers from valleys (i.e. some valley samples may be classified, as upland samples since they are more than 1000 ft away from a stream). Not having well depth information limits our ability to link water samples to specific aquifers and limits our understanding of the variability in groundwater chemistry. In addition, the detailed information, such as location and well number of these gas wells, was not provided in Molofsky *et al.*, 2013; therefore, we are not sure whether they are in the valley or upland or the local characteristics (e.g., groundwater hydraulic head and gradients). Furthermore, the different sampling methods from different groups (e.g., filtering samples in a fraction of samples in Cabot vs. not filtering in other studies) complicate the result comparisons. Because of the absence of these important pieces of information, which may confound the associations we have observed, conclusions from this study are indicative and may be difficult to directly apply to other areas in the world. More studies with better defined topography, accurate well depths, and consistent sampling methods are needed to elucidate the impacts, including causal mechanisms.

Conclusions

Our analysis of the Cabot dataset found certain meaningful associations of elevated levels of certain dissolved constituents (e.g., Ca, SO₄, Cl, and Fe) with topography and the distance to the nearest gas well. Impacts in valley samples appear to be more noticeable. These associations were not mentioned in Molofsky *et al.*, 2013. The precise reasons inducing such associations with the distance to nearest UGD well are poorly understood. The associations indicate potential for further impact on groundwater quality. Though currently there is no consensus on the UGD's impact on groundwater quality, this study adds more evidence that UGD can impact groundwater, and it demonstrates a need to conduct more studies to investigate the impact of UGD on groundwater quality and monitor long-term variations (years to tens of years) in any changes in water chemistry after UGD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank two NIEHS P30 Environmental Health Centers at Columbia University (p30-ES009089) and the University of Pennsylvania (p30-ES013508) for providing pilot funding. Hughes Science Pipeline Project supported the Barnard Summer Research Institute with a grant. Paul Heisig is specially thanked for detailed discussion about the groundwater flow and chemistry in NY and PA. This is an LDEO contribution # ZZZ

References

- Abdulaziz AM. Evaluation of Microseismicity Related to Hydraulic Fracking Operations of Petroleum Reservoirs and Its Possible Environmental Repercussions. *Open Journal of Earthquake Research*. 2014:2014.
- Annelink MPJA, Meesters JAJ, Hendriks AJ. Environmental contamination due to shale gas development. *Science of The Total Environment*. 2016; 550:431–438. [PubMed: 26845179]
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1995:289–300.
- Bowen ZH, Oelsner GP, Cade BS, Gallegos TJ, Farag AM, Mott DN, et al. Assessment of surface water chloride and conductivity trends in areas of unconventional oil and gas development—why existing national data sets cannot tell us what we would like to know. *Water Resources Research*. 2015; 51:704–715.
- Burton TG, Rifai HS, Hildenbrand ZL, Carlton DD Jr, Fontenot BE, Schug KA. Elucidating hydraulic fracturing impacts on groundwater quality using a regional geospatial statistical modeling approach. *Science of The Total Environment*. 2016; 545–546:114–126.
- CEET. 2015. <http://ceet.upenn.edu/home/>
- Council CS. Energy Development Strategy Action Plan (2014-2020). 2014. http://www.gov.cn/zhengce/content/2014-11/19/content_9222.htm (in Chinese)
- Darrah TH, Vengosh A, Jackson RB, Warner NR, Poreda RJ. Noble gases identify the mechanisms of fugitive gas contamination in drinking-water wells overlying the Marcellus and Barnett Shales. *Proceedings of the National Academy of Sciences of the United States of America*. 2014; 111:14076–14081. [PubMed: 25225410]
- Davies RJ, Almond S, Ward RS, Jackson RB, Adams C, Worrall F, et al. Oil and gas wells and their integrity: Implications for shale and unconventional resource exploitation. *Marine and Petroleum Geology*. 2014; 56:239–254.
- Drollette BD, Hoelzer K, Warner NR, Darrah TH, Karatum O, O'Connor MP, et al. Elevated levels of diesel range organic compounds in groundwater near Marcellus gas operations are derived from surface activities. *Proceedings of the National Academy of Sciences*. 2015; 112:13184–13189.

- EIA. Shale gas production drives world natural gas production growth. 2016. <http://www.eia.gov/todayinenergy/detail.cfm?id=27512>
- Fontenot BE, Hunt LR, Hildenbrand ZL, Carlton DD Jr, Oka H, Walton JL, et al. An evaluation of water quality in private drinking water wells near natural gas extraction sites in the Barnett Shale Formation. *Environmental Science & Technology*. 2013; 47:10032–10040. [PubMed: 23885945]
- Fritz SJ. A SURVEY OF CHARGE-BALANCE ERRORS ON PUBLISHED ANALYSES OF POTABLE GROUND AND SURFACE WATERS. *Ground Water*. 1994; 32:539–546.
- Gorody AW. Factors affecting the variability of stray gas concentration and composition in groundwater. *Environmental Geosciences*. 2012; 19:17–31.
- Heisig, PM.; Scott, T-M. Occurrence of methane in groundwater of south-central New York State, 2012-systematic evaluation of a glaciated region by hydrogeologic setting: US Department of the Interior, US Geological Survey. 2013.
- Hildenbrand ZL, Carlton DD Jr, Fontenot BE, Meik JM, Walton JL, Thacker JB, et al. Temporal variation in groundwater quality in the Permian Basin of Texas, a region of increasing unconventional oil and gas development. *Science of The Total Environment*. 2016; 562:906–913. [PubMed: 27125684]
- Jackson RB, Vengosh A, Darrah TH, Warner NR, Down A, Poreda RJ, et al. Increased stray gas abundance in a subset of drinking water wells near Marcellus shale gas extraction. *Proceedings of the National Academy of Sciences of the United States of America*. 2013a; 110:11250–11255. [PubMed: 23798404]
- Jackson RE, Gorody AW, Mayer B, Roy JW, Ryan MC, Van Stempvoort DR. Groundwater Protection and Unconventional Gas Extraction: The Critical Need for Field-Based Hydrogeological Research. *Groundwater*. 2013b; 51:488–510.
- Jemielita T, Gerton GL, Neidell M, Chillrud S, Yan B, Stute M, et al. Unconventional Gas and Oil Drilling Is Associated with Increased Hospital Utilization Rates. *Plos One*. 2015:10.
- McGlade C, Speirs J, Sorrell S. Unconventional gas – A review of regional and global resource estimates. *Energy*. 2013; 55:571–584.
- McKenzie LM, Guo R, Witter RZ, Savitz DA, Newman LS, Adgate JL. Birth outcomes and maternal residential proximity to natural gas development in rural Colorado. *Everyday Environmental Toxins: Children’s Exposure Risks*. 2015:111.
- Meng Q. Spatial analysis of environment and population at risk of natural gas fracking in the state of Pennsylvania, USA. *Science of The Total Environment*. 2015; 515–516:198–206.
- Molofsky LJ, Connor JA, Wylie AS, Wagner T, Farhat SK. Evaluation of Methane Sources in Groundwater in Northeastern Pennsylvania. *Ground Water*. 2013; 51:333–349. [PubMed: 23560830]
- Molofsky LJ, Richardson SD, Gorody AW, Baldassare F, Black JA, McHugh TE, et al. Effect of Different Sampling Methodologies on Measured Methane Concentrations in Groundwater Samples. *Groundwater*. 2016
- Nagel NB. Compaction and subsidence issues within the petroleum industry: From wilmington to ekofisk and beyond. *Physics and Chemistry of the Earth, Part A: Solid Earth and Geodesy*. 2001; 26:3–14.
- Olmstead SM, Muehlenbachs LA, Shih JS, Chu ZY, Krupnick AJ. Shale gas development impacts on surface water quality in Pennsylvania. *Proceedings of the National Academy of Sciences of the United States of America*. 2013; 110:4962–4967. [PubMed: 23479604]
- Poth, CW. Summary, ground-water resources of Butler County, Pennsylvania. Pennsylvania Geological Survey; Harrisburg: 1973.
- Rasmussen SG, Ogburn EL, McCormack M, Casey JA, Bandeen-Roche K, Mercer DG, et al. Association Between Unconventional Natural Gas Development in the Marcellus Shale and Asthma Exacerbations. *JAMA Internal Medicine*. 2016
- Reese SO, Neboga VV, Pelepko S, Kosmer WJ, Beattie S. Groundwater and petroleum resources of Sullivan County, Pennsylvania: Pennsylvania Geological Survey, Water Resource Report 71. Pennsylvania Geological Survey. 2014
- Rojstaczer S, Wolf S. Permeability changes associated with large earthquakes: An example from Loma Prieta, California. *Geology*. 1992; 20:211–214.

- Siegel D, Smith B, Perry E, Bothun R, Hollingsworth M. Pre-drilling water-quality data of groundwater prior to shale gas drilling in the Appalachian Basin: Analysis of the Chesapeake Energy Corporation dataset. *Applied Geochemistry*. 2015a; 63:37–57.
- Siegel DI, Azzolina NA, Smith BJ, Perry AE, Bothun RL. Methane Concentrations in Water Wells Unrelated to Proximity to Existing Oil and Gas Wells in Northeastern Pennsylvania. *Environmental science & technology*. 2015b; 49:4106–4112. [PubMed: 25764141]
- Stacy SL, Brink LL, Larkin JC, Sadovsky Y, Goldstein BD, Pitt BR, et al. Perinatal outcomes and unconventional natural gas operations in Southwest Pennsylvania. *PloS one*. 2015; 10:e0126425. [PubMed: 26039051]
- U.S.EPA. Assessment of the Potential Impacts of Hydraulic Fracturing for Oil and Gas on Drinking Water Resources (External Review Draft). U.S. Environmental Protection Agency; Washington, DC: 2015. EPA/600/R-15/047
- Vidic RD, Brantley SL, Vandenbossche JM, Yoxtheimer D, Abad JD. Impact of Shale Gas Development on Regional Water Quality. *Science*. 2013:340. [PubMed: 23888018]
- Warner NR, Christie CA, Jackson RB, Vengosh A. Impacts of Shale Gas Wastewater Disposal on Water Quality in Western Pennsylvania. *Environmental Science & Technology*. 2013; 47:11849–11857. [PubMed: 24087919]
- Warner NR, Jackson RB, Darrah TH, Osborn SG, Down A, Zhao KG, et al. Geochemical evidence for possible natural migration of Marcellus Formation brine to shallow aquifers in Pennsylvania. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109:11961–11966. [PubMed: 22778445]
- Yerkes R, Castle R. Surface deformation associated with oil and gas field operations in the United States. *Land subsidence*. 1969; 1:55–64.

Highlights

- 1) Groundwater (GW) quality was associated with distance to UGD gas wells and topography
- 2) Ca, Cl, & SO₄ levels are higher in GW near gas wells, especially in valley settings
- 3) Enhanced mixing of fresh and saline waters in valley may cause the change

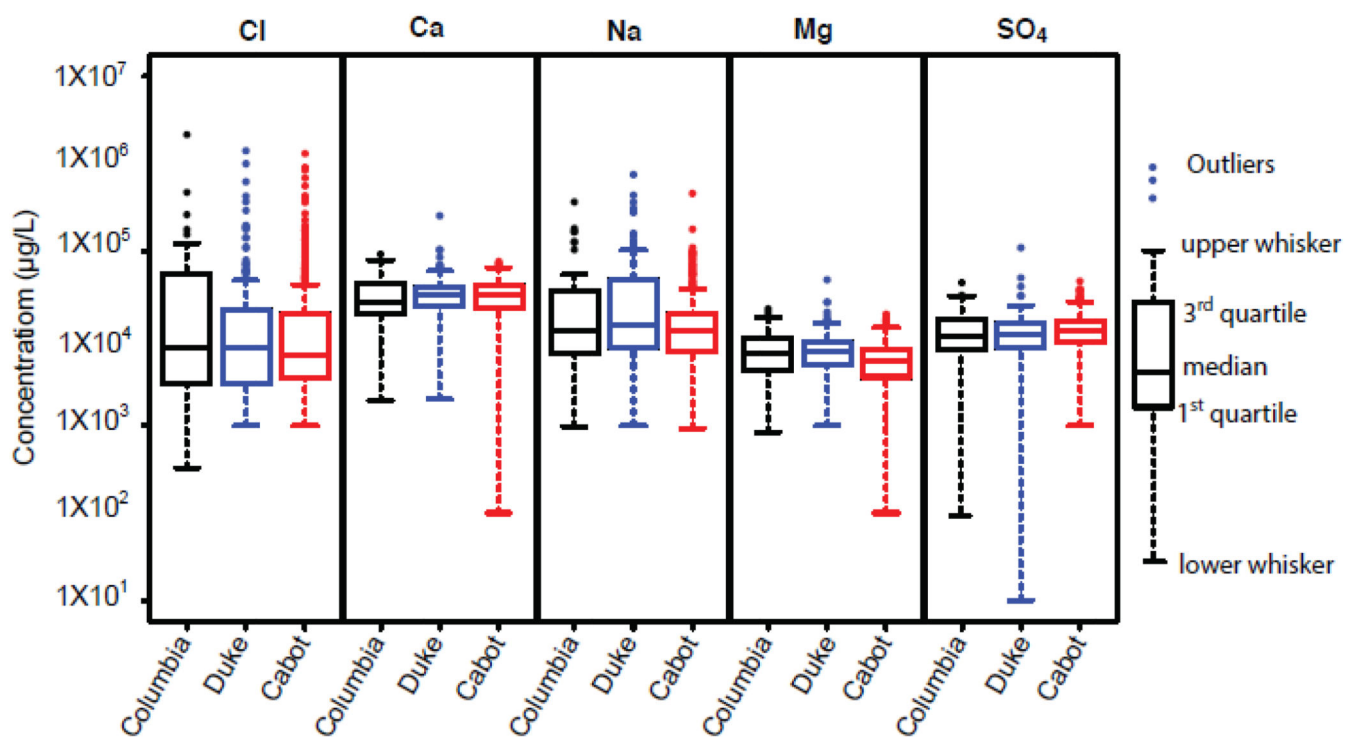


Figure 1.

Boxplots of concentration of five major constituents (Cl, SO_4 , Ca, Na, and Mg) in three studies. The upper whisker is the maximum observation below upper fence (1.5 IQR above 3rd quartile) and the lower whisker is the minimum observation above lower fence (1.5 IQR below 1st quartile). IQR is the interquartile range, a distance between 3rd and 1st quartiles.

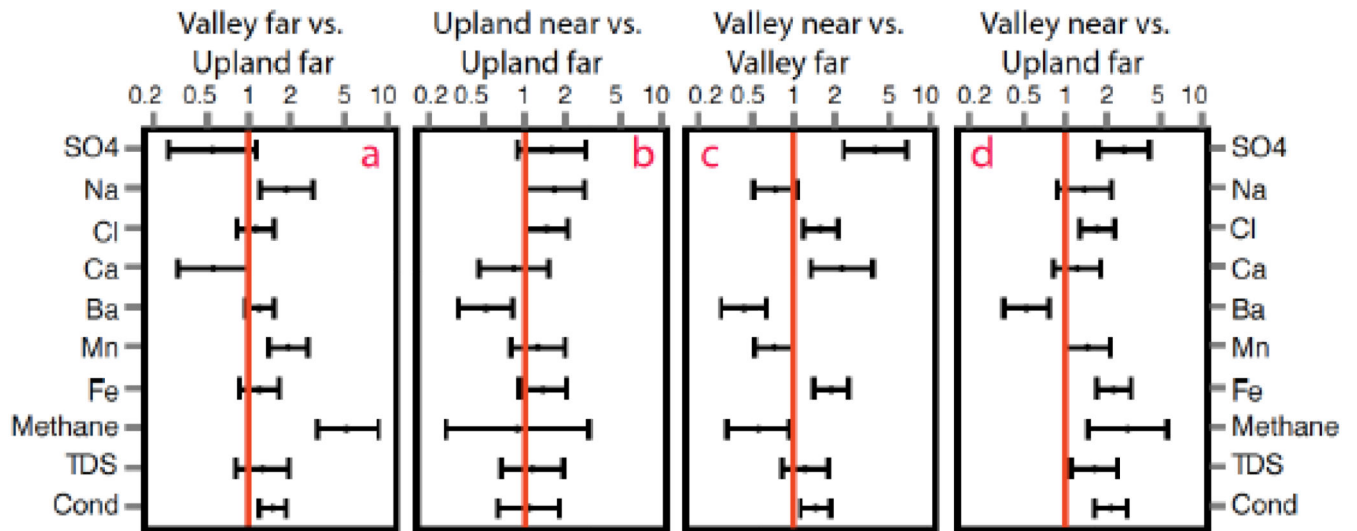


Figure 2. Ratio of proportions (RP) of 10 constituents between groups in the Cabot study. Error bar of each constituent is the 95% confidence interval of RP.

Table 1

The EPA guideline concentrations and threshold values of components used to compute the relative proportions

Symbols	Chemical names	Guidelines (mg/L)	Type of guidelines	Percentage exceeding guidelines	Threshold values (at 75 percentile-mg/L)
Na	sodium	20	HA	13.70	18
Mg	magnesium	n.a		NA	7.6
Ca	calcium	n.a		NA	40
Cl	chloride	250	SMCL	0.89	18.8
SO ₄	sulfate	250	SMCL	0	16
Ba	barium	2	MCL	1.1	0.17
Mn	manganese	0.05	SMCL	15.61	0.026
Fe	iron	0.3	SMCL	10.95	0.09
Pb	lead	0.015	TT	2.87	0.0012
Turb	turbidity	5	TT	8.30	1.1
TDS	total dissolved solids	500	SMCL	1.2	190
Cond	Specific conductivity	n.a		NA	273
Alk	alkalinity	n.a		NA	140
CH ₄	methane	28	AL	0.18	1.0 *

Note:

Health Advisory (HA) is an estimated level of a contaminant acceptable in drinking water based on health effect information. These are not legally enforceable. (EPA 2009)

Secondary Maximum Contaminant Level (SMCL) is set by the EPA as guideline for management of aesthetic considerations of water quality.

Maximum Contaminant Level (MCL) is the highest level of a contaminant that is allowed in drinking water. These are set and enforced by the Environmental Protection Agency (EPA).

Treatment Technique (TT) regulates this contaminant, by requiring that water systems take steps to control the corrosiveness of the water. The process of TT should reduce the contaminant level below the set action level.

Action Level (AL) has been suggested by the U.S. Department of Interior, Office of Surface Mining, for methane levels in water. When methane level is greater than 28 mg/L, immediate action to reduce the methane level in the water is suggested. When methane level is within 10- 28 mg/L, investigative monitoring is suggested. When methane level is below 10 mg/L, well usage is considered to be generally safe

* In this study we select the threshold value as 1 mg/L, the level high enough to overcome sampling and analytical artifacts.

Table 2

Proportions of high level of constituents (percentage above thresholds) among four groups in Cabot data

Symbols	Upland far % (n/N)	Upland near % (n/N)	Valley far % (n/N)	Valley near % (n/N)	p-value*	FDR-adjusted p-value [#]
SO₄	19.69 (25/127)	34.48 (20/58)	13 (13/100)	54.29 (76/140)	<0.0001	0.0003
Ba	28.00(91/329)	15.25 (18/118)	33.21 (91/274)	15.22 (28/184)	<0.0001	0.0003
CH₄	2.54 (18/710)	2.38 (3/126)	13.17 (88/668)	7.69 (15/195)	<0.0001	0.0003
Cond	18.67 (98/525)	21.31 (13/61)	27.95 (135/483)	42.74 (50/117)	<0.0001	0.0003
Fe	17.63 (58/329)	25 (29/116)	20.82 (56/269)	41.44 (75/181)	<0.0001	0.0003
Cl	20.12 (65/323)	29.82 (34/114)	21.98 (60/273)	36.36 (68/187)	0.0003	0.0006
TDS	19.57 (63/322)	27.35 (32/117)	24.54 (67/273)	35.11 (66/188)	0.0003	0.0006
Mn	14.24 (47/330)	18.8 (22/117)	28.83 (79/274)	21.98 (40/182)	0.0006	0.0011
Ca	25.4 (32/126)	24.56 (14/57)	15.15 (15/99)	36.03 (49/136)	0.0161	0.0250
Na	22.22 (28/126)	33.33 (19/57)	37 (37/100)	28.68 (39/136)	0.0293	0.0410
Turb	16.82 (36/214)	20.93 (18/86)	22.14 (31/140)	28.4 (46/162)	0.0425	0.0541
Mg	25.16 (81/322)	32.23 (39/121)	24.39 (70/287)	19.08 (33/173)	0.1226	0.1430
Alk	22.05 (28/127)	31.03 (18/58)	29 (29/100)	28.78 (40/139)	0.3341	0.3598
Pb	26.19 (33/126)	28.07 (16/57)	28.28 (28/99)	25 (34/136)	0.9019	0.9019

* p-value was from Chi-square test for group differences in proportion of high levels.

[#] False discovery rate, defined as the multiple tests adjusted p-values using the approach (Benjamini and Hochberg, 1995) to control for false positive error rate.

Table 3

Proportions of high level of constituents (percentage above thresholds) among four groups in Columbia and Duke data.

Symbols	Upland far	Upland near	Valley far	Valley near	p-value *	FDR-adjusted P-value #
Columbia	% (n/N)	% (n/N)	% (n/N)	% (n/N)		
Mg	38.89 (7/18)	60.0 (6/10)	32.0 (8/25)	80.0 (4/5)	0.1602	0.4005
Na	31.03 (4/18)	50.0 (5/10)	44.0 (11/25)	60.0 (3/5)	0.2711	0.4518
Cl	29.82 (7/17)	50.0 (5/10)	40.0 (10/25)	20.0 (1/5)	0.7932	0.7932
Ca	27.78 (5/18)	50.0 (5/10)	28.0 (7/25)	40.0 (2/5)	0.5711	0.7139
SO4	23.53 (4/17)	50.0 (5/10)	20.0 (5/25)	60.0 (3/5)	0.1306	0.4005
Duke						
Mg	37.5 (30/80)	32.26 (10/31)	45.16 (14/31)	71.43 (5/7)	0.2468	0.3085
Na	35.0 (28/80)	41.94 (13/31)	48.39 (15/31)	71.43 (5/7)	0.2135	0.3085
Cl	20.0 (16/80)	18.75 (6/32)	41.94 (13/31)	85.71 (6/7)	0.0006	0.0030
Ca	17.5 (14/80)	15.63 (5/32)	19.35 (6/31)	71.43 (5/7)	0.0197	0.0493
SO4	27.5 (22/80)	25.0 (8/32)	13.33 (4/30)	14.29 (1/7)	0.4590	0.4590

* p-value was from Chi-square test for group differences in proportion of high levels..

False discovery rate, defined as the multiple tests adjusted p-values using the approach (Benjamini and Hochberg, 1995) to control for false positive error rate.

Table 4

Ratio of proportions and 95% CI for Cl and Ca in Duke data

Duke	Valley far vs. Upland far RP (95% CI)	Upland near vs. Upland far RP (95% CI)	Valley near vs. Valley far RP (95% CI)	Valley near vs. Upland far RP (95% CI)
Cl	2.10 (1.15, 3.83)	0.94 (0.40, 2.18)	2.04 (1.22, 3.41)	4.29 (2.52, 7.30)
Ca	1.11 (0.47, 2.62)	0.89 (0.35, 2.28)	3.69 (1.57, 8.70)	4.08 (2.09, 7.96)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript