Methane in groundwater at concentrations high enough to be flammable can occur naturally or be related to natural gas wells. Flaming tap water was found in Granville Summit, Pennsylvania, March 2012.

PHOTO: MELANIE STETSON FREEMAN/THE CHRISTIAN SCIENCE MONITOR VIA GETTY IMAGES

ENERGY AND THE ENVIRONMENT

Engaging over data on fracking and water quality

Data alone aren't the solution, but they bring people together

By **S. L. Brantley,[1,2] R. D. Vidic,[3] K. Brasier,[4] D. Yoxtheimer,[5] J. Pollak,[6] C. Wilderman,[7] T. Wen[1]**

Extraction of unconventional oil and gas using high-volume hydraulic fracturing (HVHF)—"fracking"—is a "wicked" problem: Science and policy-making are complex and opaque; problems are unstructured, cross areas of policy jurisdiction, require coordinated action among various stakeholders who disagree about values, and tend to result in limited solutions with complex consequences (*1*). Public participation in decision-making about hydrocarbon extraction is limited by the largely private nature of transactions among mineral rights owners and industry and the narrow opportunity for public input into procedures. Likewise, obstacles to accessing water-quality data and the dearth and diversity of such data limit shared understanding. We found, however, that, although data alone do not resolve wicked problems, shared interest in gathering, discussing, and improving water-quality data can lead to productive discussions among scientists, citizens with local knowledge, regulators, and industry practitioners. Such opportunities to "pull back the curtain" on science, funded and facilitated by honest brokers, could build trust and develop procedural fairness as foundations for social license.

The rate of HVHF in Pennsylvania (PA) increased exponentially from 2004 to 2011 because of rapid technological advances that accelerated development of the Marcellus formation, the largest such shale-gas play in the United States. Almost 11,000 shale-gas wells now dot the valleys and ridges across half of PA. Spills and leaks of fracking fluids and wastes occasionally occur, and although most are small, they add to the risk of cumulative impact. Events such as well blowouts and burning tap water amplified the public's perception of risks to water quality. Although many community-based watershed organizations monitor streams and are concerned about fracking, the state has struggled to build capacity to document pre-drilling water quality and postdrilling impacts.

Although some communication among scientists and nonscientists has focused on the impact of shale-gas development, this has generally targeted seismic risk rather than water quality. In 2011, we had the idea that assessing water-quality data in PA might help address public concerns if data from academics, consultants, industry practitioners, government, and nonscientists from watershed groups were compiled in one public, online database. Although the resulting Shale Network database itself has

[1]Earth and Environmental Systems Institute, Pennsylvania State University, University Park, PA, USA. [2]Department of Geosciences, Pennsylvania State University, University Park, PA, USA. [3]Department of Civil and Environmental Engineering, University of Pittsburgh, Pittsburgh, PA, USA. [4]Department of Agricultural Economics, Sociology, and Education, Pennsylvania State University, University Park, PA, USA. [5]Marcellus Center for Outreach and Research, Pennsylvania State University, University Park, PA 16802, USA. [6]Consortium of Universities for the Advancement of Hydrologic Sciences, Inc., Cambridge, MA 02140, USA. [7]Department of Environmental Studies, Dickinson College, Carlisle, PA, USA. Email: brantley@eesi.psu.edu

proven useful, even more important has been the process of building it. We design workshops to engage data providers, recognizing that no one person understands all aspects of water data, from landscape to laboratory to computer to sharing to interpretation. These workshops forge a social network for volunteers, industry practitioners, consultants, regulators, and academic scientists voicing diverse perspectives and concerns about water quality, and the focus on data and observation keeps conversations productive. To our knowledge, no such network exists in other U.S. shale plays. The lessons we describe below from the effort cross the biophysical and social sciences, creating community among stakeholders.

Today, our database (2) spans ~28,000 sites mostly within PA and contains more than a million data values derived from multiple universities, government entities, volunteer groups, energy companies, and consultants (3, 4). All location-specific water-quality data with sufficient quality control can be published in the database, which is run collaboratively with CUAHSI (Consortium of Universities for the Advancement of Hydrologic Science, Inc.) as part of its National Science Foundation–funded work to provide a search engine that finds water data in online databases such as ours, as well as data from the U.S. Geological Survey (USGS), the U.S. Environmental Protection Agency (EPA), and universities and other entities.

Much of the surface-water data in the Shale Network database derive from sensors run by watershed groups, as well as from state agencies that had not previously published data online. Some industry data for PA—so-called "pre-drill" data that are shared with state regulators but are collected by industry before drilling to protect against potential liability—are now being uploaded, from tens of thousands of groundwater samples with up to 50 analytes each.

LESSONS ON WATER AND DATA

Given that water quality shows extreme spatial and temporal variability caused by both natural and anthropogenic factors, contamination from shale-gas activities can be difficult or impossible to document. Hundreds of HVHF chemicals are used that require diverse sampling and analytical strategies. In addition to HVHF chemicals, the wastewaters can carry naturally occurring radioactive materials, organics, metals, salts, and sediments (5). Toxic compounds that can transform in the environment are present at low, difficult-to-detect concentrations [e.g., (6)].

The constituents that are most likely to pollute waterways—sediments, salts, and methane—are already ubiquitous in the environment, derived from soils (sediments), oceanic aerosols, natural brines, road deicing (salts), and bacteria or rock formations (methane). In PA, methane also emits from many of the hundreds of thousands of mapped and unmapped legacy coal mines and gas or oil wells. Documenting new contamination in precontaminated waters is difficult (see the figure).

For many of the tens of thousands of stream kilometers and aquifer hectares near well pads in PA, there are insufficient water-quality data during relevant time periods and in the public domain to assess impacts. This is partly because the sources of potential contamination are widely distributed amid a complex stream network, aquifers are extremely heterogeneous, and contamination can be transient (7). Many of the measurements came from sensors located in limited locations, are focused on irrelevant analytes, or were only completed once, before or after drilling. Mostly, water quality is monitored in the United States in time-limited projects driven by concerns about specific phenomena (e.g., coal mining, acid rain).

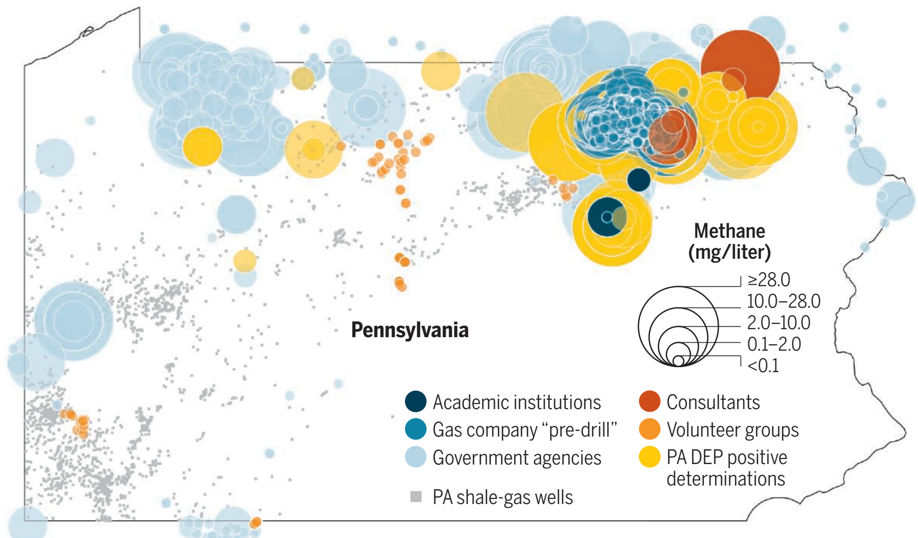
Overall, we observed that stakeholders generally hesitate to share data (3). Volunteer groups lack expertise or time. Homeowners fear lower property resale values. Academics fear competition around publications and funding. Government and industry practitioners lack resources or express concerns about liability, controversy, and nondisclosure agreements. The result is that when papers are published without explicit data values and locations, findings are highly controversial (8).

Some Shale Network participants began to teach better data management for watershed groups (4), whereas others worked to broker the agreements that led to publication of "pre-drill" industry data. But data providers generally found it arduous to compile metadata to describe their measurements. Most preferred using their own databases. Water metadata are particularly complex in comparison to other monitored features of fracking such as seismicity because of the multitude of contaminants, names, and reporting conventions for contaminants, sampling strategies, and analytical methods. The common constituent nitrate is published in 13 different reporting conventions in PA data alone. Such issues also hamper new, innovative industry-government partnerships to publicize fracking chemicals (9).

Given these factors, it is not surprising that we know of <10 incidents where data in the database marginally document contamination effects through 2017. Neither has the database ever documented a previously unreported event. Instead, incidents are reported by the public, media, regulators, or industry. After discussing such observations at workshops, PA researchers began to focus on improving baseline stream chemistry estimates (10) or documenting cumulative impacts using sediments, isotopes, or

Shale-gas wells and water methane concentrations in Pennsylvania

Data showing locations of wells drilled since 2004 are from the Pennsylvania Department of Environmental Protection (PA DEP). Methane data are from the Shale Network database (2), contributed by sources as indicated in the legend. All methane data describe groundwater, except data from volunteer groups that measured surface water. PA DEP positive determinations indicate that the PA DEP investigated and concluded that a shale-gas company was presumed responsible for contamination. Concentrations above 10 mg/liter in drinking water wells are considered by the USGS as requiring action.



GRAPHIC: J. YOU/SCIENCE

ecological indicators (*11*). Workshop discussions also introduced new techniques to use gaining streams to monitor for methane in groundwater over wider areas (*12*).

These issues, as well as the difficulties and expense of monitoring, began to drive some conversations about coordinating plans for where and what to measure. But monitors tended to maintain local, expedient sampling designs. The Shale Network effort and the new search engine for water data nonetheless provide the first step in coordination by making data easier to find. Other monitoring groups are now looking at using the CUAHSI online data system. And the ongoing publication of large volumes of pre-drill data is driving conversations about better methods of data storage while highlighting anomalies in groundwater chemistry that might bear further investigation (*13*).

FACILITATING DATA-DRIVEN DIALOGUE

Several attributes of the workshops—some noted previously by social scientists studying other initiatives and issues—led to this success. First, workshops were facilitated by funding from an agency and universities identified as honest brokers (*14*), allowing the framework to be open and not driven by an agenda. No one pushed for decisions or consensus. Scientific analyses were presented that spoke to many policy options.

Second, all data with quality control were welcomed. We stressed participation by academics, government scientists, industry practitioners, consultants, and watershed groups. Everyone was encouraged to explore data in hands-on modules and to make oral or poster presentations. The focus was on observations emphasizing distinctive expert and local knowledge (*15*).

Third, we encouraged participants to articulate new foci for future workshops. Although motivations differed among stakeholders—scientists attended workshops to present or find data whereas nonscientists attended to understand issues in their backyards—conversations facilitated cooperative agenda setting (*15*) that built trust. In 2017, for example, a nonscientist expressed concerns about disposal practices for the ~1200 tons of rock fragments accumulated per gas well, and researchers now investigate where this rock is buried.

As understanding grew that the database did not reveal much evidence of contamination, we considered if the reason was lack of data or lack of incidents (*7*). Media and government reports for PA were scrutinized for incidents where wastes contaminated water resources through spills or leaks at the surface: In the 10-year period when ~10,000 wells were drilled after 2004, <100 spills or leaks greater than 400 gallons (*7*) transiently contaminated <200 km of the ~70,000 km of PA streams overlying the Marcellus play. Using the state online reports, the average distance between spills and the nearest stream was <300 m, and many of the streams had high ecological or drinking water value (*16*). Incidents were widely but unevenly distributed throughout the gas play and appeared more frequent early in the boom (*7, 16*). Fracking chemicals were never reported to migrate from deep shale into drinking water aquifers.

Given that these incidents largely could not be corroborated by data in our database, we began to work more closely with the state to learn about regulatory actions

> ***"...recognition that groundwater is a public resource that requires public data transcends issues around shale gas..."***

and difficulties. As reported by the state, methane migration was the most commonly reported problem: ~39 of the ~9600 shale wells drilled into the Marcellus in PA between 2004 and 2015 allowed methane migration, affecting 108 drinking water supplies. But some suspect that state reporting is limited by financial and human resource constraints. So, we began developing data mining tools to investigate both natural and non-natural sources of methane (*13*).

When water-quality information for incidents was released publicly, we discussed the data (*6*). Discussions taught everyone about difficulties in determining causation, highlighting the need for multiple lines of evidence and state-of-the-art analyses. The public, regulators, academic scientists, consultants, and industry workers were together exposed to the knowns, unknowns, and gray areas.

Workshop participants remained concerned about the lack of disclosure when incidents were litigated and nondisclosure agreements signed. No regulations require data disclosure that could allow scientists in industry and academia to learn to improve best practices. In addition, no database of spill timing, volume, or cause is maintained in PA (*16*), and media and government reports of these incidents can be discrepant. These factors exacerbate our conundrum: How can we maintain public trust in water quality knowing that all data are not released and we cannot monitor everywhere a spill or leak might occur?

BARRIERS TO ENGAGEMENT

Data sharing can promote understanding and trust among stakeholders—but technical and nontechnical barriers around data must be surmounted. Water-quality investigators must follow the path that academic and USGS seismologists have followed by agreeing on standards for measuring and reporting.

Changing the norms for reporting is especially important for groundwater because, in many areas worldwide, including PA, groundwater is treated as private property yet moves across ownership boundaries. The recognition that groundwater is a public resource that requires public data transcends issues around shale gas and has already galvanized legislation in states experiencing drought, such as California.

Why doesn't data-driven engagement among scientists and nonscientists happen more often? After all, the American Petroleum Institute encourages community engagement early in exploration, and some gas companies participate in community programs. Rapid rates of technological change may outpace the rate of public engagement and policy implementation. Scientists often train to understand only select types of data and thus develop only fragmentary understanding, hampering communication with nonscientists. Limited funding from honest brokers hampers creation of unbiased forums for stakeholders to assess risks by publicly engaging around shared data. Forging opportunities to discuss science by focusing on data can nonetheless build trust and be part of the solution. ■

REFERENCES AND NOTES

1. E. P. Weber, A. M. Khademian, *Public Adm. Rev.* **68**, 334 (2008).
2. ShaleNetwork Database, DOI: 10.4211/his-data-shalenetwork.
3. K. J. Brasier *et al*., *J. Environ. Plann. Manage.* **60**, 2103 (2017).
4. C. C. Wilderman, J. Monismith, *Citizen Science: Theory and Practice* **1**, 7 (2016).
5. R. D. Vidic, S. L. Brantley, J. M. Vandenbossche, D. Yoxtheimer, J. D. Abad, *Science* **340**, 826 (2013).
6. G. T. Llewellyn *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6325 (2015).
7. S. L. Brantley *et al*., *Int. J. Coal Geol.* **126**, 140 (2014).
8. S. G. Osborn, A. Vengosh, N. R. Warner, R. B. Jackson, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 8172 (2011).
9. K. Konschnik, *Nat. Resour. J.* **54**, 319 (2014).
10. X. Niu *et al*., *Environ. Geochem. Health* 10.1007/s10653-017-0031-6 (2017).
11. W. D. Burgos *et al*., *Environ. Sci. Technol.* **51**, 8851 (2017).
12. V. M. Heilweil *et al*., *Environ. Sci. Technol.* **49**, 4057 (2015).
13. Z. Li *et al*., Corrigendum to "Searching for anomalous methane in shallow groundwater near shale gas wells" [*J. Contam. Hydrol.* **195**, 23 (2016)]. *J. Contam. Hydrol.* **207**, 50 (2017).
14. R. A. Pielke Jr., *The Honest Broker: Making Sense of Science in Policy and Politics* (Cambridge Univ. Press, New York, 2007).
15. F. Fischer, *Citizens, Experts, and the Environment* (Duke Univ. Press, 2000).
16. K. O. Maloney *et al*., *Sci. Total Environ.* **581-582**, 369 (2017).

ACKNOWLEDGMENTS

Funding derived from National Science Foundation grants OCE-11-40159, EAR-13-31726, and IIS-16-39150 to S.L.B.

10.1126/science. aan6520

Engaging over data on fracking and water quality

S. L. Brantley, R. D. Vidic, K. Brasier, D. Yoxtheimer, J. Pollak, C. Wilderman and T. Wen

Science **359** (6374), 395-397.
DOI: 10.1126/science.aan6520