Quantitative analysis of microbial contamination in private drinking water supply systems

Richard P. Allevi, Leigh-Anne H. Krometis, Charles Hagedorn, Brian Benham, Annie H. Lawrence, Erin J. Ling and Peter E. Ziegler

ABSTRACT

Over one million households rely on private water supplies (e.g. well, spring, cistern) in the Commonwealth of Virginia, USA. The present study tested 538 private wells and springs in 20 Virginia counties for total coliforms (TCs) and *Escherichia coli* along with a suite of chemical contaminants. A logistic regression analysis was used to investigate potential correlations between TC contamination and chemical parameters (e.g. NO_3^-, turbidity), as well as homeowner-provided survey data describing system characteristics and perceived water quality. Of the 538 samples collected, 41% ($n = 221$) were positive for TCs and 10% ($n = 53$) for *E. coli*. Chemical parameters were not statistically predictive of microbial contamination. Well depth, water treatment, and farm location proximate to the water supply were factors in a regression model that predicted presence/absence of TCs with 74% accuracy. Microbial and chemical source tracking techniques (*Bacteroides* gene Bac32F and HF183 detection via polymerase chain reaction and optical brightener detection via fluorometry) identified four samples as likely contaminated with human wastewater.

Key words | indicator organisms, microbial source tracking, optical brighteners, private drinking water, well

Richard P. Allevi
Leigh-Anne H. Krometis (corresponding author)
Brian Benham
Erin J. Ling
Department of Biological Systems Engineering,
200 Seitz Hall, Virginia Tech,
Blacksburg, VA 24061-0303,
USA
E-mail: *krometis@vt.edu*

Charles Hagedorn
Annie H. Lawrence
Department of Crop and Soil Environmental Science,
330 Smyth Hall, Virginia Tech,
Blacksburg, VA 24061-0404,
USA

Peter E. Ziegler
College of Agriculture and Life Sciences,
1060 Litton-Reaves Hall, Virginia Tech,
Blacksburg, VA 24061-0334,
USA

INTRODUCTION

According to the United States Census Bureau's most recent available housing survey, over 13 million occupied households in the United States rely on private household wells as a primary source of drinking water (USCB 2010). While the US Environmental Protection Agency (USEPA) regulates water quality in public water supplies (i.e. systems serving at least 25 people or having a minimum of 15 service connections) through the 1974 Safe Drinking Water Act, private water supply users are solely responsible for the care and maintenance of their water supply system (e.g. well, spring, cistern). Many private water supply owners in the United States live in rural regions where there is generally less access to the education and/or financial resources necessary to address those water quality issues unique to private water supplies (Gasteyer & Vaswani 2004; Wescoat *et al.* 2007). Perhaps not surprisingly, the US Centers for Disease Control recently reported that the proportion of annual waterborne disease outbreaks associated with non-community (i.e. individual) groundwater water supply systems increased between 1976 and 2006 relative to the total number of outbreaks reported in all system types (Craun *et al.* 2010). While this suggests that a public health issue of potentially increasing concern exists, relevant data on private system water quality and management remain scarce, rendering the issue difficult to address.

Direct monitoring for waterborne human pathogens is often impractical due to the associated low concentrations, wide variety of targets and high cost of laboratory analysis (Savichtcheva & Okabe 2006). Instead, monitoring strategies generally target fecal indicator bacteria (FIB), which are chosen based on their presence in the feces of

doi: 10.2166/wh.2013.152

warm-blooded animals, low rates of survival and/or natural presence in extra-intestinal habitats, and their association with human pathogens of concern. The presence of FIB in groundwater-fed private drinking supplies has been linked to human illness, most commonly acute gastrointestinal illness (AGI) (Raina *et al.* 1999; Macler & Merkle 2000). The USEPA currently recommends that municipal drinking water maintain a zero maximum contaminant level for *Escherichia coli* and contain no more than one sample positive for total coliforms (TCs) for systems that are tested less than 40 times per month (USEPA 2012). While not legally applicable, USEPA drinking water standards for water quality in public supplies are useful as guidelines when assessing water quality in private supplies.

Although focused primarily on surveying ambient groundwater quality as opposed to human exposure, a recent US Geological Survey study found that of approximately 400 wells sampled at the point of entry, 34% were contaminated with TCs and 8% were contaminated with *E. coli* (DeSimone 2009). In addition, a limited number of peer-reviewed studies are available that provide field information on the quality of private water supplies in the United States at the point of use, which is generally considered a more accurate measure of potential human exposure (Table 1). Some of these studies have directly examined the potential association of management or environmental factors with bacterial contamination (e.g. improper system placement, proximity of grazing animals, well depth, etc.). Shallow wells (Sworobuk *et al.* 1987) and improper well sealing (i.e. grout around the casing or well cap seal) (Lamka *et al.* 1980; Sworobuk *et al.* 1987) have been correlated with higher TC densities. Other factors that have been associated with bacterial contamination of private water supplies are improper system placement with respect to potential contamination sources, proximity of grazing animals and lack of knowledge as to the significance of contaminated water (Lamka *et al.* 1980). Collectively, these studies suggest that bacterial contamination is not uncommon in private water supply systems throughout the country, although efforts to correlate private water supply contamination with predictive factors (e.g. system type) have been limited. It is worth noting that all the studies presented in Table 1 collected presence/absence data for the targeted FIB rather than quantitative estimates of concentration.

Table 1 | Summary of previous private drinking water studies

Study	Location	Percent TCs +ve (%)	Total sources[a]
Sandhu *et al.* (1979)	South Carolina	85	460
Lamka *et al.* (1980)	Oregon	35[b]	78
Sworobuk *et al.* (1987)	West Virginia	68	155
Bauder *et al.* (1991)	Montana	40	1,300
Kross *et al.* (1993)	Iowa	45	686
Gosselin *et al.* (1997)	Nebraska	15[c]	1,808
Borchardt *et al.* (2003)	Wisconsin	28	50

[a]Repeat tests of an individual system are only counted once.
[b]Percent of samples positive for coliforms, fecal coliforms or *Staphylococcus aureus*; or with standard plate counts exceeding 500/mL.
[c]Percent of samples with 'bacterial contamination' (not necessarily total coliforms (TCs)).

As FIB presence suggests an immediate health risk, system owners are encouraged to address positive drinking water samples via system repair, enhancement or decontamination (Simpson 2004). Knowledge of primary contamination sources would therefore be helpful in the identification of efficient and long-term remediation method(s). Microbial source tracking (MST) is a collection of methods used to determine the likely source of contamination associated with FIB presence. Source tracking analyses may target microorganisms specific to a given host (e.g. polymerase chain reaction (PCR) detection of human-specific *Bacteroides*) or phenotypic differences in metabolism or enzymatic capabilities between different source strains (e.g. antibiotic resistance analysis, carbon utilization profiles) (Scott *et al.* 2002; Santo Domingo *et al.* 2007). In the majority of previously published studies, MST analyses have been used in the management of surface waters to prioritize watershed remediation strategies (Hagedorn *et al.* 1999; Simpson *et al.* 2002). The application of these techniques to private drinking water system management has not been previously explored.

Chemical source tracking (CST) is often utilized in conjunction with MST in order to provide additional information regarding potential contributors to fecal contamination of a given water sample. These methods target chemicals that are assumed to be solely associated with

human wastewater in order to distinguish between human and non-human contamination. Caffeine, pharmaceuticals and optical brighteners have all proven useful for this purpose. The presence of caffeine or human pharmaceuticals in well water along with elevated nitrate concentrations was identified as evidence of human wastewater contamination in domestic and public drinking water supplies near Reno, Nevada (Seiler *et al.* 1999). Fluorescent whitening agents (FWA) (optical brighteners) and sodium tripolyphosphate (STP) were associated with increased levels of coliform bacteria in private wells in New Zealand (Close *et al.* 1989), and were used to imply human wastewater infiltration via septic tank drainage.

The overall goal of this work was to better characterize the magnitude and incidence of microbial contamination in private wells sampled in conjunction with a state extension program, and to investigate the potential use of MST or CST techniques to allow residents to address fecal contamination more effectively. Three major objectives were to: (i) document the prevalence of FIB contamination in private water supply systems; (ii) identify statistical relationships between bacterial contamination and system or environmental characteristics; and (iii) demonstrate the application of MST and CST techniques in identifying likely sources of contamination.

METHODS

Sample collection

Water samples assessed for this research project were collected through the ongoing Virginia Household Water Quality Program (VAHWQP; www.wellwater.vt.edu), based at Virginia Tech. The program works with local Virginia Cooperative Extension educators to provide statewide water quality testing at a reduced cost to homeowners, who depend on private water supply systems, as well as appropriate education on the maintenance of systems. These services are provided to interested Virginia homeowners via periodic, county-based drinking water clinics. Participation in these clinics by homeowners is wholly voluntary. The present study considers data from VAHWQP drinking water clinics conducted during the 2011 calendar year (Figure 1). As clinics are scheduled based on local interest and educator availability, the 2011 clinics include counties with highly diverse geologic, land-use and socioeconomic characteristics.

At an initial meeting, clinic participants are provided with basic information on well, spring and cistern construction as well as the opportunity to purchase a water sampling kit (US$45). Each kit includes three sampling

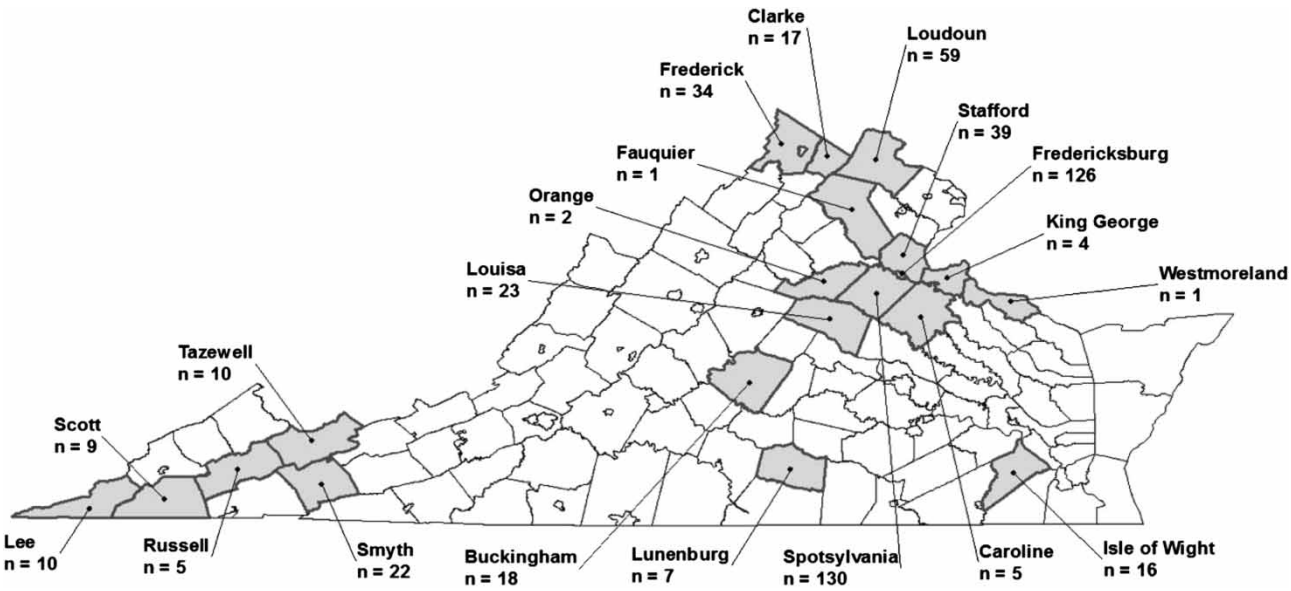

Figure 1 | Map of counties in which VAHWQP samples were taken for the present study.

bottles, sampling instructions and a survey with questions related to system construction, perceived local contamination sources, and perceived drinking water quality (Table 2). Two pre-sterilized sampling bottles are used for bacterial analysis – one bottle for *E. coli* and TC quantification and one bottle for filter capture and storage for later DNA analysis. The third bottle is used for chemical analysis, which includes testing for pH, conductivity [proxy for total dissolved solids], nitrate-N, chloride, fluoride, calcium, magnesium, sodium, manganese, copper, iron, sulfate and hardness. Participants are instructed to collect the water samples from the kitchen or bathroom faucet following an initial 5-minute flush on a pre-selected date. Each participant then brings his or her sample to a central location for transportation on ice to the water quality laboratory at the university. Samples are promptly refrigerated upon arrival in the laboratory. All bacterial analyses are performed within 8–12 hours of collection. Because the goal is to provide information on water quality representative of typical exposure at the point-of-use, homeowners are not required to bleach or otherwise disinfect their faucet prior to sample collection.

Sample processing

A total of 538 homeowner drinking water samples were processed during this study (Figure 2). Following the Virginia earthquake in the fall of 2011, an unexpectedly high volume of samples was received during the clinic servicing the counties of Caroline, Fauquier, Fredericksburg, King George, Louisa, Spotsylvania and Stafford; consequently, 181 of these samples were analyzed for presence/absence of TCs and *E. coli* only. The remaining 357 samples were quantified for TC and *E. coli* concentrations in order to document the typical magnitude of bacterial contamination. None of the samples received from the clinic following the earthquake was filter-captured and preserved for PCR analysis. Therefore, the total number of filtered samples available for later filter analysis was 207 (38% of total samples). Of these filtered samples, *E. coli* positive samples ($n = 26$) were analyzed via PCR for *Bacteroide*s because *E. coli* is considered to be a more specific indicator of contamination by mammalian species (Leclerc *et al.* 2001). Because the results of the initial clinics early in the year included very high concentrations of TCs and *E. coli*, fluorometry was added to the sample analysis procedure, beginning with the Isle of Wight clinic

Table 2 | Summary of VAHWQP participant survey

Private system characteristics	• What household water supply source was drawn for sample (well, spring, cistern)? • If 'well', is it a dug or bored well; a drilled well; or don't know? When was the well constructed? What is the approximate well depth? • Do other households share the same water supply? How many? • What water treatment devices are currently installed and affecting cold water only drawn at faucet for sample? • What pipe material is primarily used throughout your house for water distribution?
Presence of potential sources of contamination	• Describe the location of your home. Check one: on a farm; on a remote, rural lot; in a rural community; in a housing subdivision. • Do you have problems with corrosion or pitting of pipes or plumbing fixtures? • Is your water supply located within 100 ft of the following? Check all that apply: septic system drain field; home heating oil storage tank; pit privy or outhouse; pond or freshwater stream; cemetery; tidal shoreline or marsh. • Is your water supply located within a ½ mile of any of the following? Check all that apply: landfill; golf course; abandoned quarry, industry, etc.; illegal dump; crops/nursery; farm animal operation; active quarry; manufacturing/processing operation; commercial underground storage tank (UST) or supply lines.
Homeowner perceived water quality	• Does your water have an unpleasant taste? • Does your water have an objectionable odor? • Does your water have an unnatural color or appearance? • Does your water stain plumbing fixtures, cooking appliances/utensils, or laundry? • In a standing glass of water, do you notice floating, suspended, or settled particles?

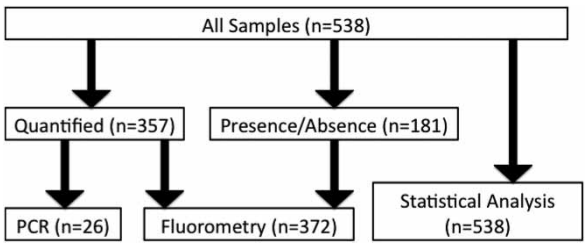

Figure 2 | Sample processing flowchart.

in June 2011, for a more in-depth investigation of possible contamination sources. Statistical efforts to identify correlations between microbial contamination and predictive factors used the entire available dataset ($n = 538$).

Microbial and source tracking analyses

Culture-based enumeration of TCs and *E. coli* was performed using the IDEXX Colilert 2000 method (www.idexx.com, Westbrook, MN, USA). The Colilert method has been proven to offer detection comparable to that of simple membrane filtration and growth on selective media, but with faster results (Cowburn *et al.* 1994). Most probable number (MPN) per 100 mL concentrations was determined using a statistical formula provided by Hurley & Roscoe (1983). Water samples that were positive for *E. coli* were confirmed by growth on eosin methylene blue (EMB) agar.

Between 24 and 36 hours of sample arrival, 250 mL of sample was captured on a sterilized 0.4 μm Isopore membrane filter (Millipore, Billerica, MA) for later molecular analysis (MST). Many MST methods involve the use of molecular tools (commonly PCR) to target and amplify genetic sequences from a source-specific host. DNA probes have been designed to target 16S rRNA gene sequences because these sequences are highly conserved and contain source-specific information (Kreader 1995; Kildare *et al.* 2007). *Bacteroides* spp. have been used as indicators of human fecal contamination as: (i) the most prevalent *Bacteroides* spp. found in the human gut are specific to humans (Allsop & Stickler 1985); and (ii) obligate anaerobes should theoretically indicate recent contamination (Fiksdal *et al.* 1985). Therefore, *Bacteroides* spp. were used as indicators of fecal contamination and were detected by amplification using end-point PCR followed by gel electrophoresis for the identification of PCR products. A DNA ladder and known positives were run in the agarose gel along with the PCR products. The Bac32F forward primer and Bac708 reverse primer were used to detect the general presence of *Bacteroides* (Bernhard & Field 2000a), while the HF183 forward primer and Bac708 reverse primer were used to detect human-specific *Bacteroides* (Bernhard & Field 2000b). DNA extraction was performed according to the QIAamp DNA Stool Handbook, under the 'Isolation of DNA from Stool for Pathogen Detection' protocol (QIAGEN Inc., Valencia, CA, USA).

Fluorometry was used to detect the presence of optical brighteners in water samples as an additional CST technique. Optical brighteners are useful in identifying sources of fecal contamination because they are found in laundry detergents, bleached toilet paper and other products that readily pass through household drains and into septic tanks. Several field studies have successfully used optical brighteners to help indicate sources of anthropogenic contamination in environmental samples (Close *et al.* 1989; McDonald *et al.* 2006; Dickerson Jr *et al.* 2007; Hartel *et al.* 2008). Following each drinking water clinic, 10 mL sample aliquots were refrigerated in the absence of UV light until analysis in order to prevent optical brighteners from degrading. Analysis was performed using a 10 AU Fluorometer (Turner Designs, Sunnyvale, FL, USA). Organic compounds in the environment that have been known to fluoresce naturally (Merker 1931; Hartel *et al.* 2007) can obscure the results of this test. As naturally fluorescing compounds are not degraded by sunlight, all positive samples were held under UV light for 4 hours to allow anthropogenic optical brighteners to degrade as confirmation. A positive result was confirmed if the reading decreased by approximately 30% of the original value, i.e. 30% of the fluorescence was attributable to degradable, likely anthropogenic, fluorescing compounds (Hartel *et al.* 2007).

Chemical analysis

As part of a VAHWQP drinking water clinic, all samples were tested for pH, conductivity [proxy for total dissolved solids], nitrate-N, chloride, fluoride, calcium, magnesium, sodium, manganese, copper, iron, sulfate and hardness in the Virginia Tech Biological Systems Engineering (BSE) Water Quality Lab and the Virginia Tech Crop and Soil Environmental Sciences (CSES) Soils Testing Laboratory using appropriate standard methods (WEF 2009). These

results were graciously provided to this study by the VAHWQP and were used for the statistical determination of relationships between chemical and bacterial analyses.

Statistical analysis

All statistical analyses were performed using JMP 9 Statistical Software (SAS Institute, Inc., Cary, NC). The goal of these analyses was to determine relationships between bacterial and chemical contamination of private drinking water supply and putative factors in order to identify drinking water system characteristics potentially associated with high contamination risk. Although *E. coli* could not be used as a dependent variable in this analysis because the available dataset for *E. coli* was ultimately too zero-heavy (i.e. 90% of samples negative for *E. coli*), TC presence was sufficiently high (41%). Using TC concentration as the dependent variable for further investigation, a cluster analysis was performed to determine if it was appropriate to break down the response into 'levels' of contamination, where each level represented a division of the total range of observed MPN/100 mL concentrations. The cluster analysis revealed that the only two appropriate levels were 'zero' and 'non-zero'. This was not surprising, as the concentration data for TCs were also somewhat zero-heavy, given that almost 60% of the 538 total samples were negative for the presence of TCs. Therefore, 'contamination' was considered as a binary response with 'present' representing any concentration of TCs in the water samples other than zero. This analysis also permitted the inclusion of all samples analyzed per Figure 2.

A logistic regression model was used to combine several predictive factors into one equation that predicted the presence or absence of TCs. This type of regression allowed for the mean response (π) to be constrained by $0 < \pi < 1$, in order to obtain a result in the form of a probability of contamination. The regression equation is given as follows (Montgomery *et al.* 2006):

$$\pi = \frac{e^{\beta_0+\beta_1 x_1 \ldots+\varepsilon}}{1+e^{\beta_0+\beta_1 x_1 \ldots+\varepsilon}} \quad (1)$$

where β_0 is the intercept, β_1 is the estimate for the coefficient of the predictor x_1, and ε is the error term. A response value (π) below 0.5 was considered to be a prediction of an uncontaminated well, while a response value above 0.5 was considered to be a prediction of a contaminated well.

Selection procedures were used to determine which variables were appropriate to include in the regression model based on a significance level of $\alpha = 0.05$. Forward, backward and stepwise selection procedures all resulted in the inclusion of the same variables in the regression model. Two separate logistic regression procedures were carried out – one for the prediction of the presence/absence of TCs based on chemical data, and one for the same prediction based on survey data.

RESULTS AND DISCUSSION

Bacterial enumeration

A total of 538 samples from 20 different VA counties (Figure 1) were processed and analyzed by the VAHWQP drinking water clinics during the 2011 calendar year. Overall, 41% ($n = 221$) of samples were positive for TCs and 10% ($n = 53$) of samples were positive for *E. coli*. This level of prevalence in Virginia is consistent with previous assessments of well water quality in peer-reviewed private water quality studies from around the United States (Table 1). No significant trends or correlations between FIB presence and location or time of year were identified. Participation rates varied widely by county/region (Figure 1) and seasonal effects may have been obscured as different regions were sampled during different months. While not generally pathogenic, the presence of TCs and/or *E. coli* in private drinking water samples does suggest possible sources of contamination such as breaches in well construction (e.g. broken cap, improper sealing) or poorly maintained water treatment systems, and does suggest a possible risk of exposure to actual human pathogens.

Distribution plots illustrating the total observed concentrations of TCs and *E. coli* are shown in Figures 3 and 4, respectively. Observed concentrations of FIB were remarkably high. The dotted lines represent the maximum detection limit of 2,081 MPN/100 mL. Fifty-three samples were above the USEPA's municipal drinking water standard for *E. coli* (zero). Approximately 50% of the quantified samples had TC concentrations greater than 40 MPN/100 mL and *E. coli* concentrations greater than 30 MPN/

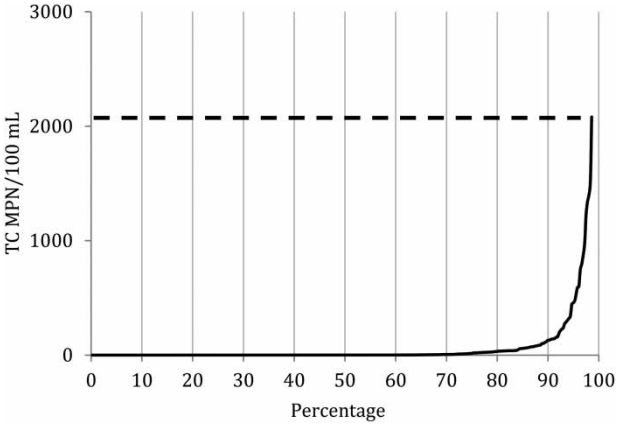

Figure 3 | Cumulative distribution plot for total coliforms ($n = 357$).

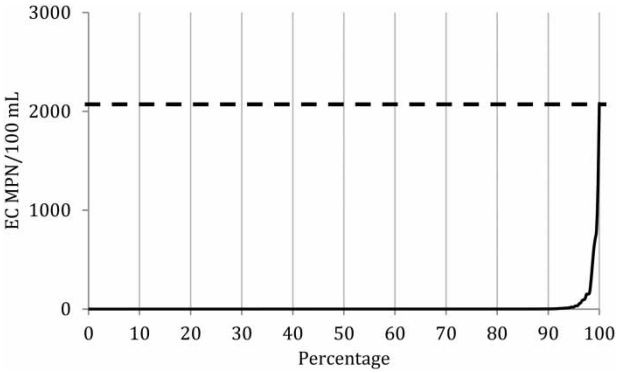

Figure 4 | Cumulative distribution plot for *E. coli* ($n = 357$).

100 mL. Six samples were above the maximum detection limit for TCs (2,081 MPN/100 mL) and one sample was above the maximum detection limit for *E. coli* (2,081 MPN/100 mL).

The average TC concentration in these *E. coli* positive samples was very high compared to the broader dataset average of 137 MPN/100 mL (Table 3). The majority of systems contaminated with *E. coli* were dug or bored wells (49%), and only 34% of homeowners indicated that their water supply system included any known water treatment device, none of which directly targeted bacteria by chlorination. Ninety-one percent of *E. coli* contaminated systems were located in a rural community, remote rural lot or on a farm and the average well depth for these systems was 34.1 m (112 ft), which is significantly shallower than the broader dataset average of 75.0 m (246 ft). There is no evidence that homeowners consistently noted any specific unpleasant physical water properties (taste, odor, color, suspended particles) associated with their drinking water.

Table 3 | Characteristics of the *E. coli* positive samples ($n = 53$ out of 538)

Variable	
Avg TC MPN/100 mL (of $n = 46$ quantified samples)	879[a]
Avg EC MPN/100 mL (of $n = 46$ quantified samples)	223[a]
System type	Springs: $n = 4$ (7.5%) Cisterns: $n = 1$ (2%) Drilled wells: $n = 18$ (34%) Dug/Bored wells: $n = 26$ (49%) No response: $n = 4$ (7.5%)
Treatment device	Yes: $n = 18$ (34%) Chlorinator: $n = 0$ (0%) Other: $n = 18$ (100%)
Average well depth (out of $n = 29$ responses)	112 ft
Average year built (out of $n = 27$ responses)	1980
Location	On a farm: $n = 14$ (26.5%) On a remote, rural lot: $n = 10$ (19%) In a rural community: $n = 24$ (45%) In a housing subdivision: $n = 4$ (7.5%) No response: $n = 1$ (2%)
Corrosion of piping	Yes: $n = 9$ (17%)
Unpleasant taste	Yes: $n = 7$ (13%)
Objectionable odor	Yes: $n = 4$ (7.5%)
Unnatural color	Yes: $n = 12$ (23%)
Water stains appliances	Yes: $n = 20$ (38%)
Visible particles in water	Yes: $n = 7$ (13%)
System is located within 100 ft of…	Septic system drain field: $n = 6$ (11%) Cemetery: $n = 2$ (4%) Home heating oil storage tank: $n = 2$ (4%) Pond or freshwater stream: $n = 2$ (4%)
System is located within 1/2 mile of…	Landfill: $n = 2$ (4%) Illegal dump: $n = 1$ (2%) Abandoned quarry: $n = 1$ (2%) Golf course: $n = 1$ (2%) Fruit orchard: $n = 12$ (23%) Farm animal operation: $n = 23$ (43%)

[a]Only 46 of the 53 *E. coli* positive samples were quantified.

Source tracking

Of the 372 samples tested for the presence of optical brighteners via fluorometry, three were positive [i.e. values were between 50 and 100 units]. After 4 hours of exposure to UV light to degrade anthropogenic optical brighteners, the samples showed declines in their readings of approximately 30% or greater, indicating that these readings were not completely the result of natural fluorescing compounds (i.e. anthropogenic sources were present). All three positive samples were collected from shallow dug or bored wells, were positive for both *E. coli* and TCs, and all three homeowners noted discoloration of water. Data collected for two of the three wells revealed that wells were constructed in 1945 and 1956, both considerably older than the broader dataset average of 1981. The age of the third well was not provided.

In order to further investigate possible sources of contamination, PCR targeting *Bacteroides* genes was performed on 26 *E. coli*-positive samples (Figure 2). One sample was positive for general *Bacteroides* (Bac32F) (Figure 5). No samples were positive for the human specific marker (HF183). The sample that was positive for general *Bacteroides* was not analyzed on the fluorometer because it was taken from the February 2011 clinic before fluorometry analysis was initiated. It was, however, sourced from a spring on a farm, and had a TC concentration of 128 MPN/100 mL and an *E. coli* concentration of 95 MPN/100 mL.

It is worth noting that no internal processing control was used during PCR, which means that there is no way to determine quantitatively whether any inhibition occurred, which might have resulted in false negatives. However, the DNA isolation protocol included treatment with InhibitEX (Inhibitex, Inc., Alpharetta, GA) tablets that bind typical inhibitors (e.g. humic acids) via chemical adsorption and then are removed via centrifugation and pelleting. It is also important to consider that negative source tracking results do not necessarily discount the possibility of fecal contamination from humans or animals, given that *Bacteroides* die off more rapidly than *E. coli* in the environment (Okabe & Shimazu 2007). Additional complications arise in characterizing human-specific *Bacteroides* because the human-specific marker (HF183) is not present in 100% of the human population. False negatives for the presence of human fecal contamination could result if the majority of the population living near a particular private drinking water supply system does not carry the HF183 marker. Although past studies have examined larger volumes of water for *Bacteroides* to increase specificity, the available water volume in this study was constrained by the practicalities of sample collection within the existing VAHWQP program.

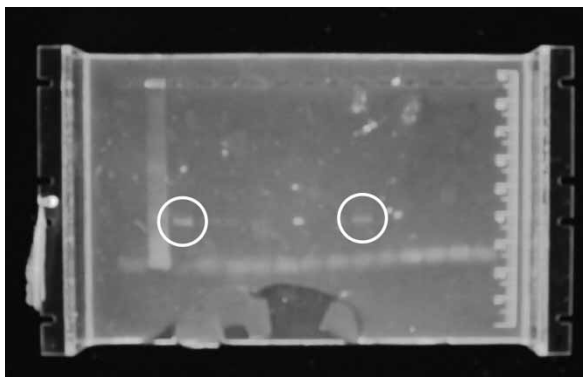

Figure 5 | Agarose gel showing control in well three and positive sample (Bac32F, general *Bacteroides* marker) in well ten.

Statistical analysis

Analysis of chemical data

A backwards selection procedure was performed to identify chemical water quality parameters appropriate for inclusion in a logistic regression model to predict the presence/absence of TCs in a water sample. All variables (e.g. pH, nitrate, etc.) were initially entered into the model, and variables were removed one at a time based on a significance threshold of $\alpha = 0.05$. At the conclusion of the selection procedure, no variables remained in the model (i.e. no parameters were significant predictors of microbial contamination). Construction of an initial scatterplot matrix (for basic observation of trends) in JMP 9 revealed no evidence of potential correlations. The scatterplot matrix did reveal expected correlations (e.g. between hardness and calcium and magnesium).

Although no previous studies have identified significant associations between bacterial contamination and concentrations of particular chemicals in private drinking water systems, these relationships may warrant further investigation,

as available datasets have been limited. Theoretically, correlations may exist between bacterial contamination and chemicals associated with anthropogenic surface water contamination, such as nitrates common in fertilizer.

Analysis of survey data

As VAHWQP participants sometimes chose not to answer all of the survey questions, the investigation of potential predictive relationships between survey responses and TC contamination began with an initial screening of data availability. Survey questions that had response rates of less than 5% (<27 respondents) were excluded from the analysis. The following questions were therefore not used in the statistical analysis of survey data:

1. Is your water source located within 100 ft of a (i) pit/privy or outhouse, (ii) cemetery?
2. Is your water source located within half a mile of a (i) landfill, (ii) illegal dump, (iii) active quarry, (iv) abandoned quarry, (v) commercial UST, (vi) golf course, (vii) manufacturing plant?
3. Is your household water distribution system primarily composed of (i) steel piping, (ii) lead piping?

A backwards selection procedure using the remaining variables was then performed to determine significant predictive factors ($\alpha = 0.05$). All the variables that remained in the model had *p*-values less than 0.02 (Table 4).

The selection procedure identified the following four variables as significantly associated with TC contamination: water supply system type ([dug/bored well] versus [drilled well *or* no response]); the presence of any type of water treatment device; well depth; and whether or not the water supply system was located within half a mile (approx. 1 km) of a farm animal operation. Using the nominal logistic regression procedure, an initial model was developed that could be used to predict whether a well would be contaminated with TCs based on these four associated factors. The parameter estimates from this model revealed that well type (i.e. drilled vs. dug/bored) was no longer considered to be significant based on $\alpha = 0.05$. Therefore, a second model was created using (i) treatment device [yes/no], (ii) well depth, and (iii) farm animal operation within half a mile of a well [yes/no] as the three predictive factors (Table 5). The resulting logistic regression model is given below:

Table 4 | Summary of the selection procedure used to determine significant predictors total coliform contamination from survey data

Parameter	Action	*p*-value
All	Entered	–
Heating oil storage tank within 100 ft [yes/no]	Removed	0.8346
Septic tank within 100 ft [yes/no]	Removed	0.7522
Objectionable odor [yes/no]	Removed	0.7396
Corrosion of pipes [yes/no]	Removed	0.5961
Water stains appliances [yes/no]	Removed	0.5749
Fruit orchard within ½ mile [yes/no]	Removed	0.548
Copper piping [yes/no]	Removed	0.3459
Community type (location)	Removed	0.2648
Particles in water [yes/no]	Removed	0.2457
Year system was built	Removed	0.1807
Stream/Pond/Lake within 100 ft [yes/no]	Removed	0.1622
Plastic piping [yes/no]	Removed	0.1577
Unusual water color [yes/no]	Removed	0.1486
Objectionable water taste [yes/no]	Removed	0.1316
System type ([Dug/Bored well] vs. [No response & Drilled well])	*Included*	*0.0035*
Treatment device [yes/no]	*Included*	*0.0193*
Well depth [ft]	*Included*	*<<0.0001*
Farm animal operation within ½ mile [yes/no]	*Included*	*0.0065*

Note: Eight percent ($n = 41$) of homeowners did not respond to the question regarding system type.

$$\pi = \frac{e^{0.0918+0.2807F-0.4739T-0.006D+\varepsilon}}{1+e^{0.0918+0.2807F-0.4739T-0.006D+\varepsilon}} \quad (2)$$

where ***F*** and ***T*** are binary representations of a farm animal operation being within a half mile and the presence of a water treatment system (1 = yes; 0 = no), respectively, and ***D*** represents the depth of a well in feet. The model was only able to make predictions for samples that had data for all the predictive variables that were included in the regression equation. Therefore, the model made 339 total predictions. Using this regression equation, the model's predictions of the presence/absence of TC contamination

Table 5 | Parameter estimates for the refined regression model predicting total coliform contamination

Term	Estimate	Prob > ChiSq
Intercept	0.9108	< 0.0001
Farm animal operation [yes]	0.2807	0.0343
Treatment [yes]	− 0.4739	0.0002
Depth [ft]	− 0.0060	< 0.0001

were compared to observed presence/absence of TC contamination to determine the accuracy of the model. Results show that the model predicted the presence/absence of TC contamination in the 339 total samples with 74% accuracy. Of the total predictions produced by the model, 9% were false positives and 17% were false negatives. Because of the possible health-related implications of predicting false negatives (i.e. predicting that a contaminated well is not contaminated), the goal should be to minimize this rate in the future.

Although this statistical analysis considered the presence or absence of a water treatment device, there was not a large enough sample size to consider 'type of device'. Some of the treatment devices provided as options in the VAHWQP survey not designed to target bacteria (e.g. water softeners). However, the presence of a water treatment device of any type may reveal that the private water supply system owner is more conscious of their water quality. This may indirectly correlate with lower levels of bacterial contamination due to additional proactive measures potentially taken by the private water supply system owner; however, targeted research would be required to confirm this hypothesis. In communicating to homeowners the importance of water treatment, it must be made clear that while the presence of an appropriate water treatment device can improve water quality, neglected or malfunctioning water treatment devices can become a *source* of bacterial contamination themselves (e.g. clogged filters).

CONCLUSIONS

One of the major goals of this study was to determine if the prevalence of coliform contamination in private water supplies that were tested as part of the VAHWQP Cooperative Extension effort was similar to rates reported in recent peer reviewed literature. The results of this study were consistent with those of previous peer reviewed literature (Table 1) given that 41% of samples were positive for TCs and 10% of the samples were positive for *E. coli*. As current USEPA drinking water standards for municipal waters simply require confirmation of coliform absence, no previous study on private drinking water reported in the literature has attempted to quantify bacterial contamination. Although quantitative measures of FIB (rather than presence/absence reporting) did not prove useful in predictive model development in this case, collection of concentration levels did provide insights into the relative magnitude of contamination, and may prove useful during future analyses of larger datasets or in studies considering dose–response predictions. Concentrations of TCs and *E. coli* observed for this study period were high, with 53 samples above the USEPA's municipal drinking water standards for *E. coli*, which require a zero maximum contaminant level. Six samples were above the maximum detection limit for TCs, and one sample was above the maximum detection limit for *E. coli*.

Three out of 372 samples tested positive for the presence of optical brighteners, and one out of 26 *E. coli* positive samples was positive for general *Bacteroides*. While these results were promising for future source tracking efforts, research should be focused on finding methods that increase the specificity of results when analyzing drinking water samples that often contain low concentrations of target organisms. Analysis should also be performed using quantitative polymerase chain reaction (qPCR) as opposed to endpoint PCR in order to eliminate the need for running the products through a gel, and so a copy number can be estimated.

Results from logistic regression analysis revealed that the chemical water analyses were not useful in predicting TC contamination in the water samples. The survey data showed more promise in its ability to predict TC contamination. A logistic regression was performed on the survey data, and the final regression model included three significant ($\alpha = 0.05$) predictors of TC contamination – well depth, whether the owner had any type of water treatment device and whether the well was located within 0.5 mile of a farm animal operation. The final regression model was able to predict the presence/absence of TC contamination in wells with 74% accuracy. Although model accuracy must be improved, these relationships do suggest private system characteristics that warrant further

investigation. Homeowners who are looking to construct new drinking water systems can consider these factors in order to prevent future contamination issues. It is important to note that the data collected through this survey focused solely on water quality perception, known environmental risks and construction factors. The inclusion of questions related to household demographics (e.g. income, education) in future surveys may provide useful insight to those designing public health interventions to improve point of use water quality in homes reliant on private water supplies, given reported links between environmental health issues and socio-economic factors.

In order to better understand public health issues associated with private drinking water supply systems and to design appropriate interventions, both outreach-based (e.g. extension) and epidemiological studies will be useful. Though extension and outreach programs do not usually target specific populations, rendering their findings less generalizable, over time they can gather an immense amount of participant-provided data on private system water quality and associated system characteristics. In addition, these programs empower private water supply system owners to become active stewards of their health and the health of the local environment. Through education, this approach bridges the gap between scientific investigation and implementation of preventative and remediation strategies to produce tangible improvements in public health.

ACKNOWLEDGEMENTS

This project was supported by the Virginia Tech College of Agriculture and Life Sciences Integrated Research Program and the Rural Health and Safety Education Competitive Program of the USDA National Institute of Food and Agriculture (NIFA), grant number 2011-05026.

REFERENCES

Allsop, K. & Stickler, D. J. 1985 An assessment of *Bacteroides fragilis* group organisms as indicators of human faecal pollution. *Journal of Applied Microbiology* **58**, 95–99.

Bauder, J. W., White, B. A. & lnskeep, W. P. 1991 Montana extension initiative focuses on private well quality. *Soil and Water Conservation Society* **46**, 69–74.

Bernhard, A. E. & Field, K. G. 2000a Identification of nonpoint sources of fecal pollution in coastal waters by using host-specific 16s ribosomal DNA genetic markers from fecal anaerobes. *Applied and Environmental Microbiology* **66**, 1587–1594.

Bernhard, A. E. & Field, K. G. 2000b A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides–Prevotella* genes encoding 16S rRNA. *Applied and Environmental Microbiology* **66**, 4571–4574.

Borchardt, M. A., Bertz, P. D., Spencer, S. K. & Battigelli, D. A. 2003 Incidence of enteric viruses in groundwater from household wells in Wisconsin. *Applied and Environmental Microbiology* **69**, 1172–1180.

Close, M. E., Hodgson, L. R. & Tod, G. 1989 Field evaluation of fluorescent whitening agents and sodium tripolyphosphate as indicators of septic tank contamination in domestic wells. *New Zealand Journal of Marine and Freshwater Research* **23**, 563–568.

Cowburn, J. K., Goodall, T., Fricker, E. J., Walter, K. S. & Fricker, C. R. 1994 A preliminary study of the use of Colilert for water quality monitoring. *Letters in Applied Microbiology* **19**, 50–52.

Craun, G. F., Brunkard, J. M., Yoder, J. S., Roberts, V. A., Carpenter, J., Wade, T., Calderon, R. L., Roberts, J. M., Beach, M. J. & Roy, S. L. 2010 Causes of outbreaks associated with drinking water in the United States from 1971 to 2006. *Clinical Microbiolical Review* **23**, 507–528.

DeSimone, L. A. 2009 Quality of water from domestic wells in principal aquifers of the United States, 1991–2004: overview of major findings. U.S. Geological Survey Circular 1332. Available at: http://pubs.usgs.gov/sir/2008/5227 (accessed March 2013).

Dickerson Jr, J. W. , Hagedorn, C. & Hassall, A. 2007 Detection and remediation of human-origin pollution at two public beaches in Virginia using multiple source tracking methods. *Water Research* **41**, 3758–3770.

Fiksdal, L., Maki, J. S., Lacroix, S. J. & Staley, J. T. 1985 Survival and detection of *Bacteroides* spp., prospective indicator bacteria. *Applied and Environmental Microbiology* **49**, 148–150.

Gasteyer, S. & Vaswani, R. 2004 Still Living Without the Basics in the 21st Century: Analyzing the Availability of Water and Sanitation Services in the United States. Rural Community Assistance Partnership: Washington, DC, Available at: http://www.rcap.org/sites/default/files/rcap-files/StillLiving/Still_Living_full.pdf (accessed March 2013).

Gosselin, D. C., Headrick, J., Tremblay, R., Chen, X. & Summerside, S. 1997 Domestic well water quality in rural Nebraska: Focus on nitrate-nitrogen, pesticides, and coliform bacteria. *Ground Water Monitoring & Remediation* **17**, 77–87.

Hagedorn, C., Robinson, S. L., Filtz, J. R., Grubbs, S. M., Angier, T. A. & Reneau, R. B. 1999 Determining sources of fecal pollution in a rural Virginia watershed with antibiotic resistance patterns in fecal *Streptococci*. *Applied and Environmental Microbiology* **65**, 5522–5531.

Hartel, P. G., Hagedorn, C., McDonald, J. L., Fisher, J. A., Saluta, M. A., Dickerson Jr, J. W., Gentit, L. C., Smith, S. L., Mantripragada, N. S., Ritter, K. J. & Belcher, C. N. 2007 Exposing water samples to ultraviolet light improves fluorometry for detecting human fecal contamination. *Water Research* **41**, 3629–3642.

Hartel, P. G., Rodgers, K., Moody, G. L., Hemmings, S. N. J., Fisher, J. A. & McDonald, J. L. 2008 Combining targeted sampling and fluorometry to identify human fecal contamination in a freshwater creek. *Journal of Water & Health* **6**, 105–116.

Hurley, M. A. & Roscoe, M. E. 1983 Automated statistical analysis of microbial enumeration by dilution series. *Journal of Applied Bacteriology* **55**, 159–164.

Kildare, B. J., Leutenegger, C. M., McSwain, B. S., Bambic, D. G., Rajal, V. B. & Wuertz, S. 2007 16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal Bacteroidales: a Bayesian approach. *Water Research* **41**, 3701–3715.

Kreader, C. 1995 Design and evaluation of *Bacteroides* DNA probes for the specific detection of human fecal pollution. *Applied and Environmental Microbiology* **61**, 1171–1179.

Kross, B. C., Hallberg, G. R., Bruner, D. R., Cherryholmes, K. & Johnson, J. K. 1993 The nitrate contamination of private well water in Iowa. *American Journal of Public Health* **83**, 270–272.

Lamka, K. G., LeChevallier, M. W. & Seidler, R. J. 1980 Bacterial contamination of drinking water supplies in a modern rural neighborhood. *Applied and Environmental Microbiology* **39**, 734–738.

Leclerc, H., Mossel, D. A. A., Edberg, S. C. & Struijk, C. B. 2001 Advances in the bacteriology of the coliform group: their suitability as markers of microbial water safety. *Annual Review of Microbiology* **55**, 201–234.

Macler, B. A. & Merkle, J. C. 2000 Current knowledge on groundwater microbial pathogens and their control. *Hydrogeology Journal* **8**, 29–40.

McDonald, J. L., Hartel, P. G., Gentit, L. C., Belcher, C. N., Gates, K. W., Rodgers, K., Fisher, J. A., Smith, K. A. & Payne, K. A. 2006 Identifying sources of fecal contamination inexpensively with targeted sampling and bacterial source tracking. *Journal of Environmental Quality* **35**, 889–897.

Merker, E. 1931 Die Fluoreszenz und die Lichtdurchiassigkeit der bewohnten Gewasser. *Zoologische Jahrbücher. Abteilung für allgemeine Zoologie und Physiologie der Tiere* **49**, 69.

Montgomery, D. C., Peck, E. A. & Vining, G. G. 2006 *Introduction to Linear Regression Analysis*. John Wiley and Sons, Inc., Hoboken, NJ.

Okabe, S. & Shimazu, Y. 2007 Persistence of host-specific *Bacteroides–Prevotella* 16S rRNA genetic markers in environmental waters: effects of temperature and salinity. *Applied Microbiology and Biotechnology* **76**, 935–944.

Raina, P. S., Pollari, F. L., Teare, G. F., Goss, M. J., Barry, D. A. & Wilson, J. B. 1999 The relationship between *E. coli* indicator bacteria in well-water and gastrointestinal illnesses in rural families. *Canadian Journal of Public Health* **90**, 172–175.

Sandhu, S. S., Warren, W. J. & Nelson, P. 1979 Magnitude of pollution indicator organisms in rural potable water. *Applied and Environmental Microbiology* **37**, 744–749.

Santo Domingo, J. W., Bambic, D. G., Edge, T. A. & Wuertz, S. 2007 Quo vadis source tracking? Towards a strategic framework for environmental monitoring of fecal pollution. *Water Research* **41**, 3539–3552.

Savichtcheva, O. & Okabe, S. 2006 Alternative indicators of fecal pollution: relations with pathogens and conventional indicators, current methodologies for direct pathogen monitoring and future application prespectives. *Water Research* **40**, 2463–2476.

Scott, T. M., Rose, J. B., Jenkins, T. M., Farrah, S. R. & Lukasik, J. 2002 Microbial source tracking: current methodology and future directions. *Applied and Environmental Microbiology* **68**, 5796–5803.

Seiler, R. L., Zaugg, S. D., Thomas, J. M. & Howcroft, D. L. 1999 Caffeine and pharmaceuticals as indicators of waste water contamination in wells. *Ground Water* **37**, 405–410.

Simpson, H. 2004 Promoting the management and protection of private water wells. *Journal of Toxicology and Environmental Health: Part A* **67**, 1679–1704.

Simpson, J. M., Santo Domingo, J. W. & Reasoner, D. J. 2002 Microbial source tracking: state of the science. *Environmental Science & Technology* **36**, 5279–5288.

Sworobuk, J. E., Law, C. B. & Bissonnette, G. K. 1987 Assessment of the bacteriological quality of rural groundwater supplies in Northern West Virginia. *Water, Air, & Soil Pollution* **36**, 163–170.

USCB 2010 *American housing survey for the United States: 2009-current housing report*. Available at: http://www.census.gov/prod/2011pubs/h150-09.pdf (accessed March 2013).

USEPA 2012 National Primary Drinking Water Regulations. Drinking Water Contaminants. Available at: http://water.epa.gov/drink/contaminants/index.cfm (accessed March 2013).

WEF 2009 *Standard Methods for the Examination of Water and Wastewater*. WEF, Alexandria, VA,. Available at: www.standardmethods.org (accessed March 2013).

Wescoat, J. L., Headington, L. & Theobald, R. 2007 Water and poverty in the United States. *Geoforum* **38**, 801–814.

First received 31 July 2012; accepted in revised form 5 March 2013. Available online 9 April 2013