The risk of pulmonary NTM infections and water-quality constituents among persons with cystic fibrosis in the United States, 2010-2019

Ettie M. Lipner[a], Joshua P. French[b], Rachel A. Mercaldo[a], Stephen Nelson[c], Adrian M. Zelazny[a], Julia E. Marshall[a], Michael Strong[d], Joseph O. Falkinham III[e], D. Rebecca Prevots[a]

[a]National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA; [b]Department of Mathematical and Statistical Sciences, University of Colorado Denver, Denver, CO, USA; [c]Department of Geological Sciences, Brigham Young University, Provo, UT, USA; [d]Center for Genes, Environment and Health, National Jewish Health, Denver, CO, USA; [e]Virginia Tech, Blacksburg, VA, USA.

Address for correspondence:

Ettie M. Lipner, Ph.D., M.P.H.; Epidemiology and Population Studies Unit, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 5601 Fishers Lane, Room 7D10, Rockville, MD 20892. Tel: 202-746-5869, Email: ettie.lipner@nih.gov

Running title: NTM infection risk and water-quality constituents

Conflicts of interest: None declared.

Source of Funding: EML, RAM, JM, AMZ, and DRP were supported by the Division of Intramural Research, NIAID. JF was supported by NSF award [1915277]. MS was supported by NSF award [1743597]. Authors receive funding to support open access publishing.

This study was classified as nonhuman subjects research by the National Institutes of Health, Office of Human Subjects Research Protection because data were de-identified and investigators could not link back to identifiable data.

The water-quality dataset and the R code for the analysis are available from the first author. The patient dataset cannot be made publicly available due to unique identifiers in the data.

ABSTRACT

Rationale: The prevalence of nontuberculous mycobacterial (NTM) pulmonary disease varies geographically in the United States (U.S.). Previous studies indicate that the presence of certain water-quality constituents in source water increase NTM infection risk.

Objective: To identify water-quality constituents that influence the risk of NTM pulmonary infection in persons with cystic fibrosis (pwCF) in the U.S.

Methods: We conducted a population-based case-control study using NTM incidence data collected from the Cystic Fibrosis Foundation Patient Registry (CFFPR) during 2010-2019. We linked patient zip code to county and associated patient county of residence with surface water data extracted from the Water Quality Portal. We used logistic regression models to estimate odds of NTM infection as a function of water-quality constituents. We modeled two outcomes: pulmonary infection due to *Mycobacterium avium* complex (MAC) and *Mycobacterium abscessus* species.

Results: We identified 484 MAC cases, 222 *M. abscessus* cases and 2816 NTM-negative CF controls resident in 11 states. In multivariable models, we found that for every 1-standardized unit increase in the log concentration of sulfate and vanadium in surface water at the county level, the odds of infection increased by 39% and 21%, respectively, among pwCF with MAC compared with CF-NTM-negative controls. When modeling *M. abscessus* as the dependent variable, every 1-standardized unit increase in the log concentration of molybdenum increased the odds of infection by 36%.

Conclusions: These findings suggest that naturally-occurring and anthropogenic water-quality constituents may influence the NTM abundance in water sources that supply municipal water systems, thereby increasing MAC and *M. abscessus* infection risk.

Abstract word count: 250

Manuscript word count: 3719

Key words: NTM; MAC; *M. abscessus*; environmental epidemiology; trace metals

What this study adds

This study adds to our understanding of water-quality constituents which increase the risk of NTM pulmonary infections. We found that across states, at a subnational level, molybdenum, sulfate, and vanadium in surface water are associated with an increased risk of NTM pulmonary infections. These water-quality constituents in surface water may act by influencing mycobacterial growth, leading to increased risk of exposure and infection, and/or though host susceptibility.

1. INTRODUCTION

The incidence and prevalence of nontuberculous mycobacteria (NTM) pulmonary infection (NTM PI) has increased globally over the past 20 years[1] . Persons with cystic fibrosis (pwCF) are a susceptible population at high risk of NTM PI, which remains one of the most challenging to treat infections, requiring prolonged treatment courses[2]. NTM are environmental opportunistic bacteria, widespread in both natural and engineered environments including in soil, natural water, water distribution systems, and biofilms in municipal water supplies[3-5]. NTM infections have been found to be the leading cause of drinking water-associated illnesses, ED visits, hospitalizations, and the most common cause of death for all types of water exposure routes[5]. Although NTM are ubiquitous, the distribution of NTM in the environment and the prevalence of NTM-associated infection varies geographically in the United States (U.S.)[6-12]. As NTM are natural inhabitants of soil and water[7,13-15], certain environmental conditions likely contribute to higher abundance of NTM, leading to increased exposure with an increased risk of NTM infection. Some environmental factors associated with NTM pulmonary infection (NTM PI) have been identified in previous studies, including evapotranspiration and geochemical soil properties[12,16]. Recent studies have demonstrated consistent and robust associations between trace metals in environmental water sources and increased risk of NTM PI. These studies, conducted in three U.S. states (Colorado, Oregon and Hawaii)[17-20], found that molybdenum in surface water was associated with increased risk of *Mycobacteroides abscessus* (formerly *Mycobacterium abscessus)* infection and vanadium in surface and groundwater was associated with an increased risk of *Mycobacterium avium* complex (MAC) infection. Molybdenum is essential for the metabolism, persistence, and pathogenesis of *Mycobacterium tuberculosis*, an organism phylogenetically related to NTM. Given the genetic relatedness of *M. tuberculosis* and NTM[21-24], and our consistently observed associations across multiple geographic regions, we hypothesize that infection rates would be higher in regions with high concentrations of these trace metals in the source water supply. Therefore, we tested whether water-quality constituents in environmental surface water are associated with NTM PI in persons with cystic fibrosis (pwCF) at a national level. Here, we describe a population-based, case-control study using water-quality data from the Water Quality Portal, a water-quality database sponsored by the U.S. Geological Survey, U.S. Environmental Protection Agency, and the National Water Quality Monitoring Council, together with national cystic fibrosis (CF) patient data extracted from the Cystic Fibrosis Foundation Patient Registry (CFFPR). We then discuss geological factors, NTM high-risk areas, and plausible biological explanations of the effect of specific water-quality constituents on the risk for NTM infection.

2. MATERIALS AND METHODS

2.1 Data Collection

Study design and patient data.

This study was a case-control study using CFFPR data[25], which includes data on >90% of pwCF resident in the United States (U.S.). Because our study population is representative of all pwCF in the U.S., we therefore describe our design as population-based. This study was classified as nonhuman subjects research by the National Institutes of Health, Office of Human Subjects Research Protection because data were de-identified and investigators could not link back to identifiable data.

Our study population included all pwCF >= 12 years with a U.S. zip code listed in the CFFPR from January 2010 through December 2019. We excluded 795 patients with invalid zip codes. We abstracted data related to patient zip code, NTM cultures and patient demographics.

We defined incident cases in two ways: 1) pwCF who had at least one positive NTM culture preceded by two negative cultures during the study period, 2) pwCF who had at least one positive NTM culture preceded by two negative cultures during the study period and lived in the same county for at least one year prior to their first positive culture. Controls were defined as pwCF who had at least two negative cultures within a single county over a period of at least two years without a history of any positive cultures. Our study population comprised 4,829 CF patients across the conterminous U.S., defined as the 48 U.S. states excluding Alaska and Hawai'i, referred to hereafter as the U.S.

2.1.1 Water-quality data compilation

We obtained water quality data from the Water Quality Portal (WQP)[26], a water-quality database collected and hosted by the U.S. Geological Survey, the U.S. Environmental Protection Agency, and the National Water Quality Monitoring Council. We accessed the WQP on 28 December 2021 and extracted water-quality data for the U.S. The majority of the data were surface water measurements; therefore, we chose to focus the water exposure dataset on only surface water samples. The cleaned dataset used in this analysis required correcting identified unit errors, as well as identifying and excluding particular sampling locations and sampling types that were not representative of recreational or municipal drinking water. All cleaning steps were done in conjunction with a USGS scientist (WS McBride) who had unique knowledge on water sampling. Appendix 1 details the cleaning steps taken to transform raw data to the cleaned dataset.

Since our previous findings focused on trace metals[17-20], we limited our water exposure dataset to states that had at least 50% of counties with trace metal water data available for analysis. This narrowed our analysis to 11 states: Arizona, California, Colorado, Connecticut, Maine, Massachusetts, Nevada, New Mexico, Rhode Island, Utah, Wyoming.

This dataset included water-quality constituents collected from sampling sites across the conterminous U.S. from January 1, 2000 through December 31, 2019. All water-sample sites were aggregated by county. Water-quality constituents were eliminated if data were not available for more than 50 percent of counties. Following these curation steps, 29 water-quality constituents remained for analysis (Supplementary Table 1). We used a natural log transformation on all individual measurements. Subsequently, we calculated the median value of each water-quality constituent for each county. We standardized all the water-quality constituent log concentrations to have a mean of 0 and standard deviation of 1. For counties with missing data, we imputed the median value for all water-quality constituents. Each patient was assigned the water quality value of each constituent for his or her respective county of residence. The states and counties with available data are shown in Figures 1 and 2.

2.2 Statistical analysis

2.2.1 Logistic Regression Models with Backward Elimination

We used logistic generalized linear models (GLMs) with backward elimination to model two separate outcomes (infection due to MAC and *M. abscessus* species) as a function of water-quality constituents in surface water. For each subject, county-level median values of each water-quality constituent (standardized, imputed) were included. Because demographic variables, including age, age of CF diagnosis, sex, chronic macrolide use, median income, are associated with NTM infection risk and could also influence county of residence, we included these as confounders in our models. For each outcome, we constructed a model (Table 2; Model 1) that included only those water-quality constituents whose variance inflation factor was less than 10;

values above 10 indicate high collinearity. For these models, we sequentially removed a single constituent with the highest variance inflation factor over 10. By removing these variables with high variance inflation factors, we mitigated the potential impact of collinear covariates, making it easier to assess the role of the remaining variables in the model. This procedure eliminated magnesium and sodium, leaving the following 27 water-quality constituents as explanatory variables: aluminum, arsenic, boron, beryllium, boron, cadmium, calcium, chloride, chromium, cobalt, copper, iron, lead, manganese, mercury, molybdenum, nickel, nitrate, phosphorus, potassium, selenium, silver, strontium, sulfate, vanadium, zinc, pH, as well as several demographic variables. Next, the backwards elimination variable selection procedure was used to determine a final model for both groups of NTM species. The backwards elimination algorithm removes explanatory variables one-at-a-time based on which variable removal minimizes the AIC selection criterion. This process continues until removing an explanatory variable that results in a worse fitted model. The final models for both groups of NTM species are referred to as Model 1 and shown in Table 2. The correlation matrix for all water-quality constituents is shown in Supplementary Table 1. Next, we constructed separate single-constituent logistic GLMs for the water-quality constituents that were significant in Model 1 (Table 2) to better assess their effect in the absence of other potentially correlated constituents. We estimated the odds of NTM infection among pwCF given exposure to water-quality constituents in surface water sources in 11 U.S. states. We present an odds ratio, 95% confidence interval (CI), and p-value for each model variable. Finally, we predicted the probability of NTM infection by county among pwCF based on the elemental composition of water. The results are presented in probability maps across counties in the 11 U.S. states analyzed (Figures 1 & 2).

3. RESULTS

3.1 Study Population Characteristics

Our study population comprised 3,897 pwCF in 11 states including 2,818 NTM culture-negative controls, 520 incident MAC cases, and 239 incident *M. abscessus* cases. We restricted our analysis to 484 incident MAC cases and 222 incident *M. abscessus* cases who lived in the same county for at least one year prior to their first positive culture. Our study participants were resident in 11 states (Arizona, California, Colorado, Connecticut, Maine, Massachusetts, Nevada, New Mexico, Rhode Island, Utah, Wyoming) at any point during 2010 through 2019. Regression model results using all 520 MAC and 239 *M. abscessus* cases are presented in the Supplement (Supplementary Table 2; Model 1 & Supplementary Table 3; Model 2).

Demographic characteristics of cases and controls are shown in Table 1. We observed a younger mean age, younger mean age of CF diagnosis, higher proportion of chronic macrolide use, and a higher proportion of males among pwCF with *M. abscessus* infection compared to those with MAC infection.

3.2 Logistic Regression Models

For MAC, we found that sulfate, vanadium, and zinc were significantly positively associated with increased odds of infection, while chromium, copper, and manganese demonstrated significantly reduced odds of infection. For every 1-standardized log unit increase in concentration, the odds of MAC infection increased by 39% for sulfate, 21% for vanadium, and 41% for zinc. For every 1-standardized log unit increase in chromium, copper, and manganese, the odds of MAC infection decreased by 26% for chromium, 23% for copper, and 20% for manganese. (Table 2; Model 1).

For *M. abscessus*, mercury, molybdenum, and phosphorus were significantly positively associated with odds of infection, while arsenic, chromium, and manganese were significant protective factors. For every 1- standardized log unit increase, the odds of *M. abscessus* infection increased by 45% for mercury, 36% for molybdenum and 25% for phosphorus. In contrast, for every 1- standardized log unit increase in the odds of *M. abscessus* infection decreased by 40% for arsenic, 24% for chromium and 26% for manganese (Table 2; Model 1).

In separate single-constituent models, controlling for the significant demographic and clinical variables (Table 3; Model 2) we found that for MAC, sulfate, vanadium, chromium, and manganese remained statistically significant. For every 1-standardized log unit increase, the odds of MAC infection increased by 36% for sulfate and 13% for vanadium. In contrast, for every 1-standardized log unit increase in chromium and manganese concentrations in surface water, the odds of MAC infection decreased by 22% for chromium and 21% for manganese. For *M. abscessus*, only molybdenum and chromium remained statistically significant in single-constituent models. For every 1-standardized log unit increase, the odds of *M. abscessus* infection increased by 28% for molybdenum. In contrast, the odds of *M. abscessus* infection decreased by 19% for every 1-standardized log increase in chromium (Table 3; Model 2). Several explanatory variables significant at a level of 0.05 were also significant in our analysis after adjusting for the multiple comparisons problem. Specifically, we used the Bonferroni correction to adjust the significance level of 0.05 by the number of models considered (14 models) so that a variable was only considered significant when the p-value is less than $\alpha = 0.05/14 = 0.004$. Chromium remained significant in Models 1 & 2 for MAC, but not for *M. abscessus*. In Model 1, manganese remained significant for MAC and *M. abscessus*. Although, in Model 2, manganese remained significant for MAC, but not for *M. abscessus*. In Models 1 & 2, vanadium was not statistically significant for MAC. In Model 1, sulfate was not statistically significant for MAC, but was significant in Model 2. Molybdenum remained statistically significant for *M. abscessus* in both Models 1 & 2.

We predicted county-level risk of MAC (Figure 1) and *M. abscessus* (Figure 2) infection among pwCF based on Model 1 (Table 2). We used Model 1 here since it accounts for the presence of multiple water-quality constituents that more accurately represents the natural environment. Counties with the highest predicted probability of MAC infection were clustered in central and southern California, central and southern Nevada, southern Arizona, central and northern New Mexico as well as numerous counties throughout Colorado and Wyoming. Most counties in Utah demonstrated low predicted probability of MAC infection. Two counties in Massachusetts and Maine showed higher predicted probability of MAC infection compared with all other northeastern counties. Counties with the highest predicted probability of *M. abscessus* infection appear throughout much of Colorado and New Mexico. High predicted probability for *M. abscessus* infection is also seen in central and southern counties in California. Numerous counties in Nevada, Utah, and Wyoming show elevated predicted probability of *M. abscessus* infection. Counties in Arizona and the 4 northeastern states show low predicted probability for *M. abscessus* infection.

4. DISCUSSION

We found that increasing concentrations of sulfate and vanadium in surface water sources were associated with an increased risk of MAC infections and that increasing concentrations of molybdenum were associated with increased risk of *M. abscessus* infection among pwCF across 11 U.S. states. This study confirms our previous findings, and further strengthens the evidence for the increased risk of NTM infection with vanadium and/or molybdenum that we have observed in both CF and non-CF populations[17-20]. Increasing concentrations of vanadium in

Oregon surface water[19] and Hawai'i groundwater[20] were significantly associated with increased risk of MAC infection, while increasing concentrations of molybdenum in surface water in both Colorado[17,18] and Oregon[19] were significantly associated with increased risk of *M. abscessus* infection. Interestingly, in our current study, sulfate demonstrated the strongest risk association for MAC while vanadium was only borderline significant. In Hawai'i[20], however, we observed the opposite effect, where vanadium was by far the strongest risk factor and sulfate demonstrated only borderline significance. This is expected as basalt rocks, the predominant rock type on the island, are not major sources of sulfate, but rather sulfate comes from sea spray aerosols deposited inland on soils which then makes its way to streams and groundwater[27]. In Oregon[19], we observed a strong association between vanadium and MAC infection, but we did not observe a signal for sulfate. Our current findings also confirmed our previous results, conducted in Colorado[17,18] and Oregon[19], demonstrating that molybdenum is a significant risk factor for *M. abscessus* pulmonary infections. The protective effect observed with manganese may be related to its ability to oxidize and precipitate, which yields black particulates, trapping NTM and thereby removing it from the water[28]. While chromium demonstrated a significant protective association for both MAC and *M. abscessus*, we are much less certain of its effects in NTM and our results suggest that further investigation of this metal is warranted.

This study focused on water-quality constituent concentrations in surface water. While the database providing the patient data did not identify those individuals whose water source is groundwater rather than surface water, we observed that vanadium was highly associated with MAC infection in both Oregon surface water[19] and Hawaiian groundwater[20]. In both states, basalt is a primary rock formation, which contains high concentrations of vanadium. Ultimately as rocks weather and soil erodes, minerals, like molybdenum and vanadium, are naturally released into our waterways.

4.1 Geochemical Controls on Solutes

In this section we review determinants of sulfate, vanadium, and molybdenum concentrations in surface waters in general, followed by a discussion of selected counties where the local geology appears to influence MAC infection rates. The aim of this discussion is to help identify geologic determinants of high-risk areas in the U.S. and elsewhere globally.

Although this study concentrates on surface waters, perennial streams (those that continually flow) have a baseflow (i.e., groundwater) component where waters have interacted with rocks and sediments in the subsurface. Indeed, some cities rely almost entirely on groundwater. For example, Tucson, Arizona relies almost entirely on groundwater derived from local aquifers[29] and lies in a high-risk county in southern Arizona. In areas like these, we infer that surface waters interact with similar materials as groundwater, although the longer water-rock contact times in aquifers may result in higher solute leaching[20,30].

Vanadium Geochemistry. A source for elevated vanadium (V) in streams includes rocks and sediments that are vanadium rich and have contact with stream waters. Such rocks include basalts and shales, which can be the most enriched of any common rock type (Supplementary Figure 1). Equally important may be the oxidation state or pH of the water. Vanadium solubility is enhanced in basic and oxidizing waters[31,32].

Molybdenum. Similarly, molybdenum (Mo) depends on Mo-rich source rocks, oxidizing conditions and a basic pH. Molybdenum concentrations in basalts and granite/rhyolite rocks are similar and elevated in shales (Supplementary Figure 2). Highly Mo-enriched waters may occur in oxidized, alkaline volcanogenic aquifers (e.g., Hawai'i)[33,34], as a result of reactive dissolution of Mn- and Fe-oxides or the degradation of organic matter. Long aquifer residence times (>300 year[30]) also produce enrichment. Hodge et al.[35] suggested that carbonate rocks sequester molybdenum during precipitation and later release them during dissolution.

Sulfate. Sulfate, along with chloride, bicarbonate and carbonate, is one of the major anions in natural waters. Sulfate is stable under a wide range of conditions, although S^{2-} (H_2S, HS^-) is common in very reducing environments.[36]. Sulfate in continental surface and groundwaters can result from the dissolution of sulfate minerals such as gypsum ($CaSO_4{\bullet}2H_2O$) and anhydrite ($CaSO_4$) in sedimentary rocks or by oxidation of sulfide minerals. Gypsum and anhydrite are naturally occurring, soluble sulfate salts that are widespread (Figure 3).

Figure 3 illustrates potential sources of vanadium, molybdenum and sulfate. However, some caution should be exercised in interpreting this map in fine detail: some mapping is discontinuous across state boundaries, because these maps are compiled from individual state survey maps that emphasize the local geology differently. For example, one state may emphasize mapping unconsolidated surficial deposits over bedrock.

4.2 Assessment of Selected High-Risk Counties

Utah. Although Utah has relatively low NTM infection risks (Figures 1, 3), in depth knowledge of the geology of this state allows us to make interpretations of several counties and county-groups that are examined as examples of geochemical controls on MAC infection. The cluster of counties in eastern and south-central Utah with MAC risks among pwCF of 0.0901 - 0.14 could be affected by widespread gypsum-bearing rocks in the water sources of these counties. Two counties near the southwestern Wyoming border with higher MAC risks among pwCF of 0.1401- 0.20 may be influenced by both gypsum-bearing rocks and vanadium-rich shales (Figure 3).

California. A small cluster of counties in central California exhibits an MAC risks among pwCF of 0.201 - 0.30 (Figures 1, 3). These areas may be influenced by basalt and shale bedrock, as well as unconsolidated deposits in the Central Valley, where shale may be a source of both sulfate and vanadium, with basalt as an additional source. Erosion of sediment [37]and unconsolidated deposition may also be sources of sulfate and vanadium. Thus, it is not just bedrock that may be important, but the material reworked and redeposited from bedrock.

4.3 Water Imports and NTM Risk

Nevada acts as a warning that not all risks can be attributed to local rocks. This may occur when a large portion of water consumed has been imported from other drainage basins rather than being locally derived. This may occur frequently in arid and semi-arid urban centers where local water resources are insufficient to sustain the population and industry. Nevada's most populous county, Clark County, which includes Las Vegas, has a predicted MAC risk among pwCF of 0.201 - 0.30, the highest in the state (Figures 1, 3). With a county population of >2.3 million[38], average rainfall in Las Vegas City is only 106 mm/yr[39]. In such an arid setting, most water must be imported. In fact, 90% of the water used in the city is from the Colorado River [40]. Thus, water-borne NTM risks in this urban area are imparted by the entire Colorado River drainage in Wyoming, western Colorado and eastern Utah. Local geology probably plays little role in V, Mo and sulfate concentrations. The Colorado River drainage upstream from Clark County is underlain by gypsum-bearing rocks, as well as shales that impart these aqueous species to the Colorado River. A key takeaway from this is that NTM risk must be considered in terms of water sources, whether they be local or imported.

4.3 Role of selected trace metals on mycobacterial metabolism, pathogenicity and virulence

While the distribution of minerals is consistent with the distribution of CF-NTM patients, the relevance of these constituents in relation to NTM is also borne out by its relevance to mycobacterial physiology and pathogenesis. ABC transporters are a major class of a bacterial translocation machinery that transport a variety of substrates such as metals, nutrients, and vitamins inside the cell for survival and growth functions[41]. In most bacteria, specific ABC transporter genes, ModA, B and C, mediate high affinity uptake of molybdate (the bioavailable form of molybdenum) from the environment. Sulfur uptake, in the form of sulfate is mediated by an ABC transporter composed of genes CysW, T, and A. ModABC and CysWTA transport system genes are present in multiple reference genomes of *M. avium* and *M. abscessus* (i.e. *M. avium* 104 and *M. abscessus* ATCC 19977)[37].

Metals have an important effect on the intracellular growth of bacteria[42]; specifically, for *Mycobacterium tuberculosis,* metals play an important role in intracellular survival, pathogenesis and virulence and are required for growth within human macrophages[43]. Similarly to *M. avium* and *M. abscessus*, *M. tuberculosis* has specific molybdate and sulfate ABC transporters, ModABC and CysWTA[42]. Molybdenum, a trace metal and key component of several enzymes involved in nitrogen, sulfur, and carbon metabolism, is essential for *M. tuberculosis* pathogenicity and survival in the host[41,44]. Moreover, mutations of the *modA* gene, which is responsible for fixation of molybdate, resulted in reduced survival of *M. tuberculosis* in the lungs of mice[45]. The high conservation of the ModABC system within mycobacteria suggests a critical role of molybdenum among NTM species as well.

The uptake and assimilation pathway of sulfate (i.e., sulfate reduction), which provides the bacteria with the essential nutrient sulfur, is critical for cysteine and methionine biosynthesis. In addition, sulfur-containing metabolites have been shown to be important for pathogenesis and survival of *M. tuberculosis*. For example, the thiol-containing small molecule, mycothiol protects *M. tuberculosis* against antimicrobial agents and oxidative stress[42,43,46]. Sulfate binding periplasmic protein SubI with the ABC transporter CysTWA system mediate sulfate uptake in *M. tuberculosis*. Mutations in genes of this transporter affect the bacterial growth in vitro and its survival in macrophages during infection[41]. The presence of this sulfate transporter in *M. avium* and *M. abscessus* is consistent with a critical role of sulfate in survival and virulence of certain NTM species.

An alternative mechanism for the effect of molybdenum and vanadium may be on the growth of NTM in the natural and human-engineered environment. Surface waters are nutrient poor and may be limited for nitrogen, an essential mineral. In oxygenated surface waters, nitrate is the predominant nitrogen species. NTM are capable of utilizing nitrate (and nitrite) as the sole nitrogen source through the action of nitrate reductase, a molybdenum-dependent enzyme[47,48]. The enzyme may also use vanadium in place of molybdenum. Thus, the presence of molybdenum and vanadium in surface waters may provide the metal-cofactor for the use of nitrate as a sole nitrogen source.

5. Conclusion

We hypothesize that increasing concentrations of specific water-quality constituents, namely molybdenum, sulfate, and vanadium, influence the metabolism and growth of MAC and *M. abscessus* species in environmental water sources. Our hypothesis assumes that increased metabolism and growth leads to higher NTM abundance at specific regions, thereby increasing the risk of NTM exposure and subsequent infection in humans in those same regions. In the absence of environmentally measured NTM data, we used incidence of NTM infection as a proxy for NTM abundance in our studies. Two studies support this approach: 1) high NTM pulmonary disease prevalence was significantly correlated with regions of high NTM abundance across the U.S. and Europe[10], 2) isolation of NTM from shower aerosols was significantly associated with NTM pulmonary disease in a case control study (matched for age, sex, and geographic location of residence)[49]. While more research is needed to confirm our results in other geographic regions of the U. S., these findings may be useful to inform susceptible patients of their risks in residing in specific regions.

While the overlap of county-specific risk estimates (Figure 1,2) and the geology of our study region (Figure 3) supported our reported associations (Table 2 & Table 3), the presence of dedicated sulfate and molybdate transporter genes in the NTM genome further support our hypothesis of biological relevance and findings[37]. The ABC transporters provide an active mechanism for bacteria to acquire essential elements in the environment or within the host to improve its survival and pathogenicity. Although little is known about the role of sulfate and molybdenum in NTM metabolism and pathogenesis, the high conservation of ABC transport genes within mycobacteria as well as the well documented role that sulfate and molybdenum play in *M. tuberculosis* metabolism and pathogenesis, offers biologic plausibility, and lends support for causality.

FIGURE CAPTIONS

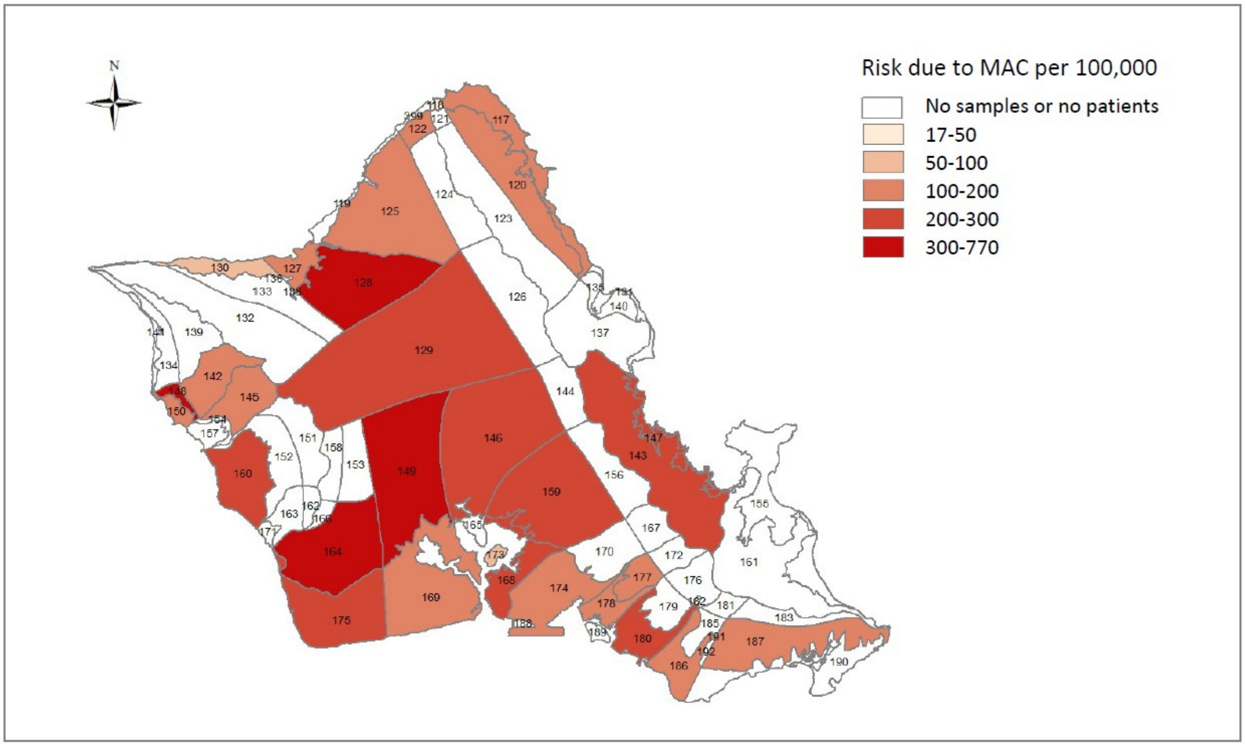


Figure 1. Predicted probability of MAC infection for counties where pwCF resided. Orange lines indicated state line boundaries. Gray lines represent county line boundaries in the U.S. State names are printed in *black*.

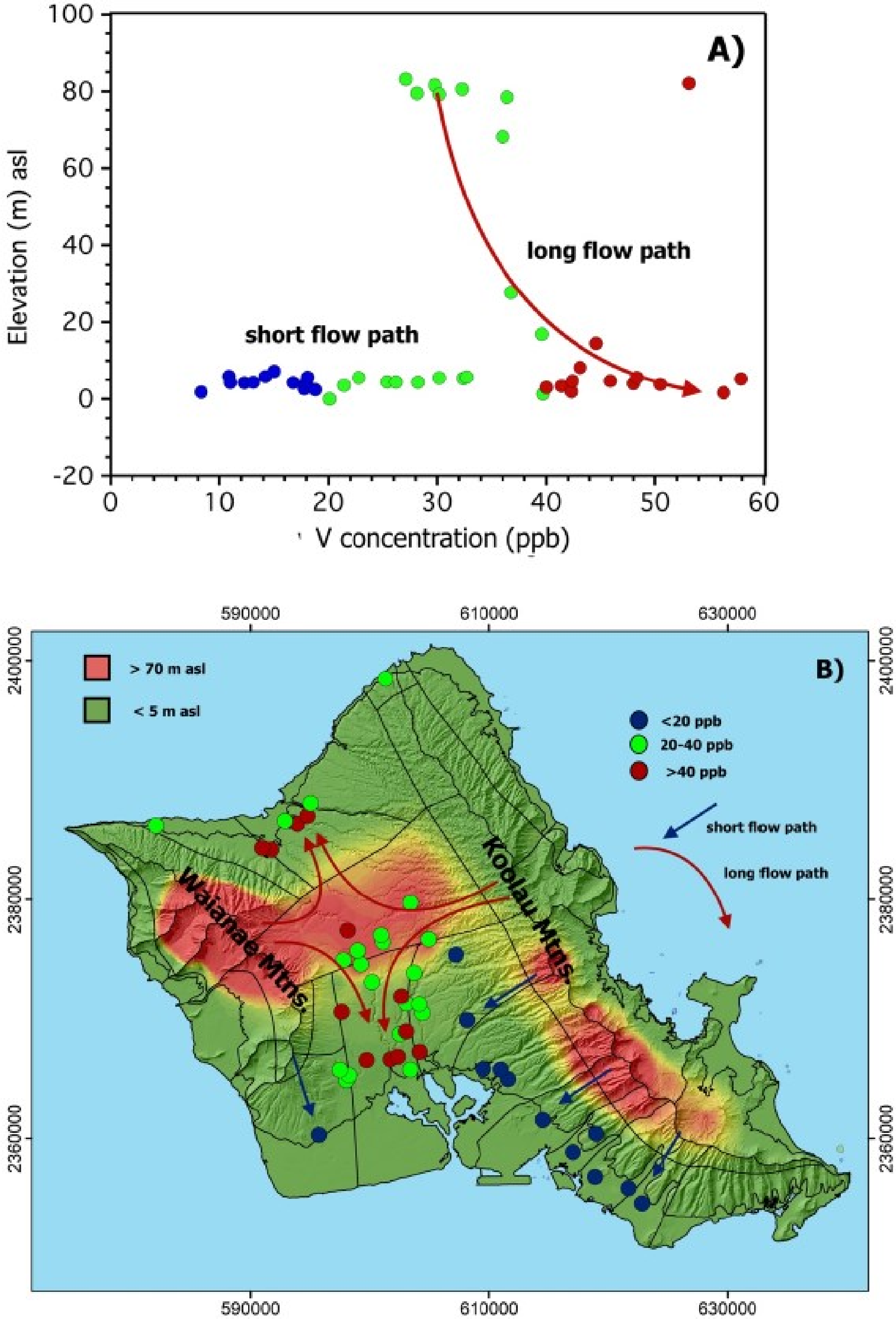


Figure 2. Predicted probability of *M. abscessus* infection for counties where pwCF resided. Orange lines indicated state line boundaries. Gray lines represent county line boundaries in the U.S. State names are printed in *black*.

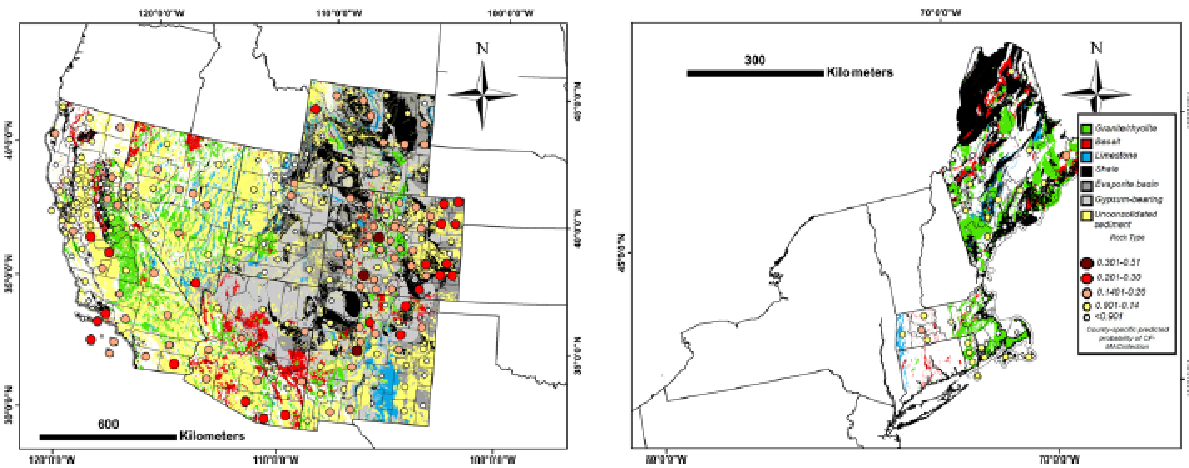


Figure 3. Simplified geologic map of the states discussed, with emphasis on rocks that impart vanadium, molybdenum and sulfate to surface and groundwaters. The light gray pattern represents areas where gypsum is present at the surface and subsurface, whereas the dark gray pattern represents evaporite basins containing gypsum (Weary and Doctor, 2014). Black represents the location of shales, mudstsones, clay deposits, and their metamorphic equivalents (pelitic schists). Red represents rocks of basaltic composition, including intrusive and metamorphic lithologies. Green represents granitic/rhyolitic rocks, also including their intrusive and metamorphic counterparts. Blue represents carbonate rocks, chiefly limestone and dolomite. Yellow represents unconsolidated Quaternary deposits (Horton, 2017).

Table 1. Demographic characteristics of cases and beneficiaries in Kaiser Permanente Hawaii residing in Oahu.

Characteristic	MAC cases (n=402)	*M. abscessus* cases (n=136)	Beneficiaries (n=193,284)
Age, yr, mean±SD	61.8±12.2	62.2±12.3	43.5±16.1
Female sex, (%)	53.0	58.1	49.4
Ethnicity:			
White (%)	24.1	23.5	19.9
Asian (%)	47.8	58.1	33.1
NHOPI (%)	4.7	4.4	9.6
Neighborhood deprivation index	-0.22±0.69	-0.24±0.74	-0.18±0.77

Table 2. Model 1. Poisson regression model with backward elimination examining water-quality constituents (with VIF values less than 10) associated with NTM infection risk in Oahu, HI. Bolded estimates are statistically significant ($p < 0.05$). CI = Confidence Interval.

MAC species		*M. abcessus* group species	
Variable	Relative Risk (95% CI) p-value	Variable	Relative Risk (95% CI) p-value
Sex: Female	**1.20 1.06, 1.34 (0.0002)**	Age: (1 Year)	1.10 0.98, 1.24 (0.104)
Chloride (1-log unit)	1.11 0.98, 1.27 (0.103)	Chloride (1-log unit)	**1.34 1.05, 1.72 (0.020)**
Potassium (1-log unit)	0.81 0.62, 1.05 (0.114)	Nitrate (1-log unit)	1.24 0.98, 1.56 (0.069)
Sulfate (1-log unit)	**1.23 1.03, 1.47 (0.023)**	Potassium (1-log unit)	**0.49 0.29, 0.83 (0.008)**
Vanadium (1-log unit)	**1.18 1.06, 1.33 (0.004)**	Sulfate (1-log unit)	**1.52 1.09, 2.14 (0.015)**

Table 3. Model 2. Single-constituent Poisson regression model examining significant water-quality constituents from Model 1 and other covariates significantly associated with NTM infection risk for MAC species in Oahu, HI. Bolded estimates are statistically significant ($p < 0.05$). CI = Confidence Interval

Characteristic	Relative Risk 95% CI p-value	Characteristic	Relative Risk 95% CI p-value
Sex: Female	1.28 (1.10, 1.36) (0.0003)	Sex: Female	1.28 (1.14, 1.44) 4.6×10^{-5}
Sulfate (1-log unit)	**1.12 1.004, 1.24 (0.043)**	Vanadium (1-log unit)	**1.22 1.10, 1.36 (0.0003)**

Table 4. Model 3. Single-constituent Poisson regression model examining significant water-quality constituents from Model 1 and other covariates significantly associated with NTM infection risk for *M. abscessus* group species in Oahu, HI. Bolded estimates are statistically significant ($p < 0.05$). CI = Confidence Interval

Characteristic	Relative Risk 95% CI p-value	Characteristic	Relative Risk 95% CI p-value	Characteristic	Relative Risk 95% CI p-value
Age: (1 Year)	1.00 (0.91, 1.11) 0.935	Age: (1 Year)	1.01 (0.92, 1.12) 0.775	Age: (1 Year)	1.01 (0.92, 1.11) 0.825
Chloride (1-log unit)	1.03 0.91, 1.11 (0.662)	Potassium (1-log unit)	0.95 0.78, 1.17 (0.644)	Sulfate (1-log unit)	1.17 0.97, 1.41 (0.113)

ACKNOWLEDGEMENTS

The authors thank U.S. Geologic Survey (USGS) scientists, W. Scott McBride for detailed assistance in curating our water quality dataset, as well as Dr. Katherine Walton-Day for additional assistance with selecting water source types obtained from the Water Quality Portal. The authors would also like to thank the Cystic Fibrosis Foundation for the use of CF Foundation Patient Registry data to conduct this study. Additionally, we would like to thank the patients, care providers, and clinic coordinators at CF Centers throughout the United States for their contributions to the CF Foundation Patient Registry.

This work was supported [in part] by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases.

REFERENCES

1. Prevots DR, Marras TK. Epidemiology of human pulmonary infection with nontuberculous mycobacteria: a review. *Clin Chest Med.* Mar 2015;36(1):13-34.
2. Floto RA, Olivier KN, Saiman L, et al. US Cystic Fibrosis Foundation and European Cystic Fibrosis Society consensus recommendations for the management of non-tuberculous mycobacteria in individuals with cystic fibrosis. *Thorax.* Jan 2016;71 Suppl 1:i1-22.
3. Gomez-Smith CK, LaPara TM, Hozalski RM. Sulfate Reducing Bacteria and Mycobacteria Dominate the Biofilm Communities in a Chloraminated Drinking Water Distribution System. *Environ Sci Technol.* Jul 21 2015;49(14):8432-8440.
4. Falkinham JO, 3rd. Surrounded by mycobacteria: nontuberculous mycobacteria in the human environment. *J Appl Microbiol.* Aug 2009;107(2):356-367.
5. Gerdes ME, Miko S, Kunz JM, et al. Estimating Waterborne Infectious Disease Burden by Exposure Route, United States, 2014. *Emerg Infect Dis.* Jul 2023;29(7):1357-1366.
6. Falkinham JO, 3rd. Ecology of Nontuberculous Mycobacteria. *Microorganisms.* Oct 30 2021;9(11).
7. George KL, Parker BC, Gruft H, Falkinham JO, 3rd. Epidemiology of infection by nontuberculous mycobacteria. II. Growth and survival in natural waters. *Am Rev Respir Dis.* Jul 1980;122(1):89-94.
8. Kirschner RA, Jr., Parker BC, Falkinham JO, 3rd. Epidemiology of infection by nontuberculous mycobacteria. Mycobacterium avium, Mycobacterium intracellulare, and Mycobacterium scrofulaceum in acid, brown-water swamps of the southeastern United States and their association with environmental variables. *Am Rev Respir Dis.* Feb 1992;145(2 Pt 1):271-275.
9. Strollo SE, Adjemian J, Adjemian MK, Prevots DR. The Burden of Pulmonary Nontuberculous Mycobacterial Disease in the United States. *Ann Am Thorac Soc.* Oct 2015;12(10):1458-1464.
10. Gebert MJ, Delgado-Baquerizo M, Oliverio AM, et al. Ecological Analyses of Mycobacteria in Showerhead Biofilms and Their Relevance to Human Health. *mBio.* Oct 30 2018;9(5).
11. Adjemian J, Olivier KN, Seitz AE, Holland SM, Prevots DR. Prevalence of nontuberculous mycobacterial lung disease in U.S. Medicare beneficiaries. *Am J Respir Crit Care Med.* Apr 15 2012;185(8):881-886.
12. Adjemian J, Olivier KN, Seitz AE, Falkinham JO, 3rd, Holland SM, Prevots DR. Spatial clusters of nontuberculous mycobacterial lung disease in the United States. *Am J Respir Crit Care Med.* Sep 15 2012;186(6):553-558.
13. Lewis AH, Falkinham JO, 3rd. Microaerobic growth and anaerobic survival of Mycobacterium avium, Mycobacterium intracellulare and Mycobacterium scrofulaceum. *Int J Mycobacteriol.* Mar 2015;4(1):25-30.
14. Thorel MF, Falkinham JO, 3rd, Moreau RG. Environmental mycobacteria from alpine and subalpine habitats. *FEMS Microbiol Ecol.* Sep 1 2004;49(3):343-347.
15. Guenette S, Williams MD, Falkinham JO, 3rd. Growth Temperature, Trehalose, and Susceptibility to Heat in Mycobacterium avium. *Pathogens.* Aug 15 2020;9(8).

16. Foote SL, Lipner EM, Prevots DR, Ricotta EE. Environmental predictors of pulmonary nontuberculous mycobacteria (NTM) sputum positivity among persons with cystic fibrosis in the state of Florida. *PLoS One.* 2021;16(12):e0259964.
17. Lipner EM, French J, Bern CR, et al. Nontuberculous Mycobacterial Disease and Molybdenum in Colorado Watersheds. *Int J Environ Res Public Health.* May 29 2020;17(11).
18. Lipner EM, Crooks JL, French J, Strong M, Nick JA, Prevots DR. Nontuberculous mycobacterial infection and environmental molybdenum in persons with cystic fibrosis: a case–control study in Colorado. *Journal of Exposure Science & Environmental Epidemiology.* 2021/07/03 2021.
19. Lipner EM, French JP, Falkinham JO, 3rd, et al. NTM Infection Risk and Trace Metals in Surface Water: A Population-Based Ecologic Epidemiologic Study in Oregon. *Ann Am Thorac Soc.* Sep 28 2021.
20. Lipner EM, French JP, Nelson S, et al. Vanadium in groundwater aquifers increases the risk of MAC pulmonary infection in O'ahu, Hawai'i. *Environ Epidemiol.* Oct 2022;6(5):e220.
21. Levillain F, Poquet Y, Mallet L, et al. Horizontal acquisition of a hypoxia-responsive molybdenum cofactor biosynthesis pathway contributed to Mycobacterium tuberculosis pathoadaptation. *PLoS Pathog.* Nov 2017;13(11):e1006752.
22. Williams MJ, Kana BD, Mizrahi V. Functional analysis of molybdopterin biosynthesis in mycobacteria identifies a fused molybdopterin synthase in Mycobacterium tuberculosis. *J Bacteriol.* Jan 2011;193(1):98-106.
23. Williams M, Mizrahi V, Kana BD. Molybdenum cofactor: a key component of Mycobacterium tuberculosis pathogenesis? *Crit Rev Microbiol.* Feb 2014;40(1):18-29.
24. McGuire AM, Weiner B, Park ST, et al. Comparative analysis of Mycobacterium and related Actinomycetes yields insight into the evolution of Mycobacterium tuberculosis pathogenesis. *BMC Genomics.* Mar 28 2012;13:120.
25. Knapp EA FA, Goss CH, Sewall A, Ostrenga J, Dowd C, Elbert A, Petren KM, Marshall BC. The Cystic Fibrosis Foundation Patient Registry: Design and Methods of a National Observational Disease Registry. . *Annals of the American Thoracic Society.* 2016;3(7):1173-1179.
26. U.S. Geological Survey NWQMC. Water Quality Portal. 2012; https://www.waterqualitydata.us/portal/. Accessed September 19, 2022.
27. Nelson ST, Tingey, D.G., Selck, B. The denudation of ocean islands by ground and surface waters: The effects of climate, soil thickness, and water contact times on Oahu, Hawaii. *Geochimica et Cosmochimica Acta.* 2013;103(15):276-294.
28. Cerrato JM, Falkinham JO, 3rd, Dietrich AM, Knocke WR, McKinney CW, Pruden A. Manganese-oxidizing and -reducing microorganisms isolated from biofilms in chlorinated drinking water systems. *Water Res.* Jul 2010;44(13):3935-3945.
29. City of Tucson, Drinking Water Distribution System. 2022; https://www.tucsonaz.gov/water/distribution-system. Accessed December 6, 2022.
30. Harkness JS, Darrah, T.H., Moore, M.T., Whyte, C.J., Mathewson, P.D., Cook, T., Vengosh, A. Naturally occurring versus anthropogenic sources of elevated molybdenum

in groundwater: Evidence for Geogenic contamination from Southeast Wisconsin, United States. *Environmental Science & Technology, .* 2017;51(21):12190-12199.

31. Johannesson KH, Lyons, W.B., Graham, E.Y. and Welch, K.A. Oxyanion concentrations in eastern Sierra Nevada rivers–3. Boron, molybdenum, vanadium, and tungsten. *Aquatic Geochemistry.* 2000;6(1):19-46.
32. Wright MT, Belitz K. Factors controlling the regional distribution of vanadium in groundwater. *Groundwater.* 2010;48(4):515-525.
33. Smedley PL, Cooper DM, Lapworth DJ. Molybdenum distributions and variability in drinking water from England and Wales. *Environ Monit Assess.* Oct 2014;186(10):6403-6416.
34. Smedley PL, Cooper DM, Lapworth DJ. Molybdenum in natural waters: A review of occurrence, distributions and controls. *Applied Geochemistry.* 2017;84:387-432.
35. Hodge VF, Johannesson, K.H. and Stetzenbach, K.J. Rhenium, molybdenum, and uranium in groundwater from the southern Great Basin, USA: evidence for conservative behavior. *Geochimica et Cosmochimica Acta.* 1996;60(17):3197-3214.
36. Hem JD. Study and interpretation of the chemical characteristics of natural water In: Department of the Interior UGS, ed. Vol 22541985.
37. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* Jan 1 2000;28(1):27-30.
38. U.S. Census Bureau. Clark County, Nevada https://www.census.gov/quickfacts/clarkcountynevada. Accessed 26 June, 2023.
39. United States - National Weather Service. National Weather Service. 1999; https://www.weather.gov/vef/1991-2020Normals. Accessed 26 June 2023.
40. LLVVW (Las Vegas Valley Water District). 2022; https://www.lvvwd.com/water-system/where-your-water-comes-from/index.html. . Accessed December 2, 2022.
41. Cassio Barreto de Oliveira M, Balan A. The ATP-Binding Cassette (ABC) Transport Systems in Mycobacterium tuberculosis: Structure, Function, and Possible Targets for Therapeutics. *Biology (Basel).* Dec 4 2020;9(12).
42. Soni DK, Dubey SK, Bhatnagar R. ATP-binding cassette (ABC) import systems of Mycobacterium tuberculosis: target for drug and vaccine development. *Emerg Microbes Infect.* 2020;9(1):207-220.
43. Ingale P, Lad B, Kabra R, Singh S. Dissecting druggability of ABC transporter proteins in Mycobacterium species through network modeling. *J Biomol Struct Dyn.* Nov 2022;40(18):8365-8374.
44. Zhong Q, Kobe B, Kappler U. Molybdenum Enzymes and How They Support Virulence in Pathogenic Bacteria. *Front Microbiol.* 2020;11:615860.
45. Camacho LR, Ensergueix D, Perez E, Gicquel B, Guilhot C. Identification of a virulence gene cluster of Mycobacterium tuberculosis by signature-tagged transposon mutagenesis. *Mol Microbiol.* Oct 1999;34(2):257-267.
46. Hatzios SK, Bertozzi CR. The regulation of sulfur metabolism in Mycobacterium tuberculosis. *PLoS Pathog.* Jul 2011;7(7):e1002036.
47. Butala NS, Falkinham JO. Nitrate and nitrite reductase activities of Mycobacterium avium. *Int J Mycobacteriol.* Oct-Dec 2018;7(4):328-331.

48. McCarthy CM. Utilization of nitrate or nitrite as single nitrogen source by Mycobacterium avium. *J Clin Microbiol.* Feb 1987;25(2):263-267.

49. Tzou CL, Dirac MA, Becker AL, et al. Association between Mycobacterium avium Complex Pulmonary Disease and Mycobacteria in Home Water and Soil. *Ann Am Thorac Soc.* Jan 2020;17(1):57-62.

Supplementary Materials

Supplementary Table 1. Water-quality constituents extracted from the Water Quality Portal (WQP) included in the cleaned dataset [12]. https://www.waterqualitydata.us/portal/

Metals	Nonmetals
Aluminum	Alkalinity
Barium	Antimony
Beryllium	Arsenic
Bromide	Bicarbonate
Cadmium	*Boron*
Calcium	Carbonate
Chromium	*Chloride*
Cobalt	Fluoride
Copper	Fluorine
Cyanide	*Nitrate*
Iron	Nitrite
Lead	*Selenium*
Magnesium	*Sulfate*
Manganese	
Mercury	
Molybdenum	
Nickel	
Potassium	
Sodium	
Strontium	
Thallium	
Vanadium	
Zinc	

*Italicized variables were included in principal component analysis.

Supplementary Table 2. Median and standard deviation (SD) values of water-quality constituents obtained from 3 datasets, the Water Quality Portal (WQP), the HI Department of Health (HI DOH), and Brigham Young University (BYU), used in principal component analysis (PCA) (µg/L = micrograms per liter).

Water-quality constituents	Median ± IQR (µg/L)
Aluminum	9.85 ± 28.4
Barium	2.36 ± 3.6
Boron	44 ± 36.3
Bromide	190 ± 185.5
Cadmium	0 ± 0
Calcium	10300 ± 7710
Chloride	49800 ± 99400
Chromium	10 ± 6.9
Copper	2.75 ± 5.9
Lead	14 ± 43.4
Magnesium	9440 ± 7930
Nickel	0 ± 0.33
Nitrate	4100 ± 2075
Potassium	2100 ± 1800
Selenium	0 ± 0.3
Sodium	29400 ± 26300
Strontium	72.8 ± 67.8
Sulfate	11500 ± 9980
Vanadium	30.2 ± 22.6
Zinc	7.8 ± 9.5

Supplementary Table 3. Correlation matrix (Pearson's Correlation Coefficient, ρ) for the water-quality constituents contributing to Principal Components 1 - 3.

Constituent	Al	Br	Cd	Ca	Cl	Cr	Pb	Mg	NO_{3^-}	K	Na	Sr	SO_4^{2-}	V
Aluminum (Al)	1.00													
Bromide (Br)	-0.19	1.00												
Cadmium (Cd)	**0.75**	-0.20	1.00											
Calcium (Ca)	0.23	**0.78**	0.22	1.00										
Chloride (Cl)	-0.20	**0.42**	-0.17	0.16	1.00									
Chromium (Cr)	**0.85**	-0.23	**0.83**	0.15	-0.13	1.00								
Lead (Pb)	**0.87**	-0.15	**0.84**	0.21	0.01	**0.95**	1.00							
Magnesium (Mg)	0.19	**0.81**	0.27	**0.88**	0.30	0.24	0.28	1.00						
Nitrate (NO_{3^-})	**-0.73**	0.25	**-0.77**	-0.04	0.08	**-0.82**	**-0.84**	-0.16	1.00					
Potassium (K)	-0.08	**0.79**	-0.02	**0.59**	**0.60**	-0.10	0.07	**0.71**	0.05	1.00				
Sodium (Na)	**-0.36**	**0.91**	**-0.39**	**0.56**	**0.55**	-0.31	-0.25	**0.64**	**0.37**	**0.76**	1.00			
Strontium (Sr)	0.23	**0.82**	0.19	**0.96**	0.21	0.13	0.21	**0.89**	-0.09	**0.69**	**0.63**	1.00		
Sulfate (SO_4^{2-})	-0.07	**-0.93**	-0.23	**0.75**	0.34	-0.16	-0.11	**0.75**	0.27	**0.69**	**0.89**	**0.80**	1.00	
Vanadium (V)	**-0.80**	0.08	**-0.83**	-0.27	0.00	**-0.84**	**-0.87**	**-0.37**	**0.88**	-0.08	0.27	-0.29	0.13	1.00

Bolded estimates are statistically significant ($p<0.05$)

Supplementary Table 4. Poisson regression model with backward elimination examining water-quality constituents (with VIF values less than 10) associated with NTM infection risk for aquifer populations greater than 150 in Oahu, HI.
Bolded estimates are statistically significant ($p < 0.05$). CI = Confidence Interval

MAC species	
Variable	Relative Risk (95% CI) p-value
Sex: Female	**1.38 1.18, 1.62 (5.9x10^{-5})**
Mean Neighborhood deprivation index	1.27 0.96, 1.66 (0.090)
Sulfate (1-log unit)	1.10 0.99, 1.23 (0.075)
Vanadium (1-log unit)	**1.22 1.10, 1.36 (0.0003)**

Supplementary Table 5.
Penalized Quasi-Likelihood method (PQL).
Bolded estimates are statistically significant.

Characteristic	Coefficient*	Bias-corrected 95% CI*
Age	-0.301	-10.8, 1.47
Gender: Female	0.137	-1.52, 15.4
Ethnicity: White	0.00	-3.74, 9.92
Ethnicity: Native Hawaiian or Pacific Islander	0.00	-3.68, 21.2
Neighborhood Deprivation	0.00	-2.39, 6.39
Aluminum	**3.75**	**0.425, 11.2**
Barium	-1.0	-14.3, 7.65
Boron	-0.663	-9.95, 0.364
Bromide	**-4.10**	**-9.85, -3.00**
Cadmium	0.651	-0.545, 10.6
Calcium	**-3.02**	**-9.19, -0.688**
Chloride	-0.298	-8.60, 5.34
Chromium	-0.191	-7.52, 6.18
Copper	0.00	-2.64, 8.26
Lead	**-1.86**	**-16.1, 0.00**
Magnesium	**3.79**	**1.43, 17.7**
Nickel	-0.481	-26.4, 1.55
Nitrate	-0.620	-10.7, 1.03
Potassium	0.00	-3.03, 8.02
Selenium	-0.209	-11.1, 6.17
Sodium	**3.39**	**2.26, 11.0**
Strontium	**2.91**	**0.0256, 12.6**
Sulfate	-1.01	-18.8, 5.35
Vanadium	**2.47**	**0.0314, 10.5**
Zinc	**-1.71**	**-8.97, -0.205**

*Estimates are in the log scale

Supplementary Table 6.
Approximate penalized Loglikelihood method (APL).
Bolded estimates are statistically significant.

Characteristic	Coefficient*	Bias-corrected 95% CI*
Age	-0.236	-5.48, 0.00
Gender: Female	0.085	-0.032, 3.28
Ethnicity: White	0.00	-0.836, 19.2
Ethnicity: Native Hawaiian or Pacific Islander	0.00	-0.887, 14.1
Neighborhood Deprivation	0.00	-0.370, 9.39
Aluminum	**0.396**	**0.00, 7.64**
Barium	0.00	-0.409, 7.81
Boron	**-0.772**	**-9.61, -0.015**
Bromide	**-0.782**	**-9.41, -0.010**
Cadmium	0.00	-0.467, 11.5
Calcium	**-0.169**	**-7.47, 0.00**
Chloride	0.289	-0.224, 11.6
Chromium	0.00	-0.388, 4.09
Copper	0.00	-0.712, 12.5
Lead	0.00	-0.479, 11.7
Magnesium	**0.423**	**0.00, 9.56**
Nickel	**0.234**	**0.00, 4.47**
Nitrate	0.102	-0.043, 7.96
Potassium	0.00	-0.363, 11.8
Selenium	**-0.224**	**-16.4, 0.00**
Sodium	0.00	-0.363, 11.8
Strontium	0.00	-0.323, 30.0
Sulfate	**1.17**	**0.983, 3.74**
Vanadium	**0.653**	**0.00, 10.4**
Zinc	0.00	-0.268, 15.5

*Estimates are in the log scale

Supplementary Table 7.

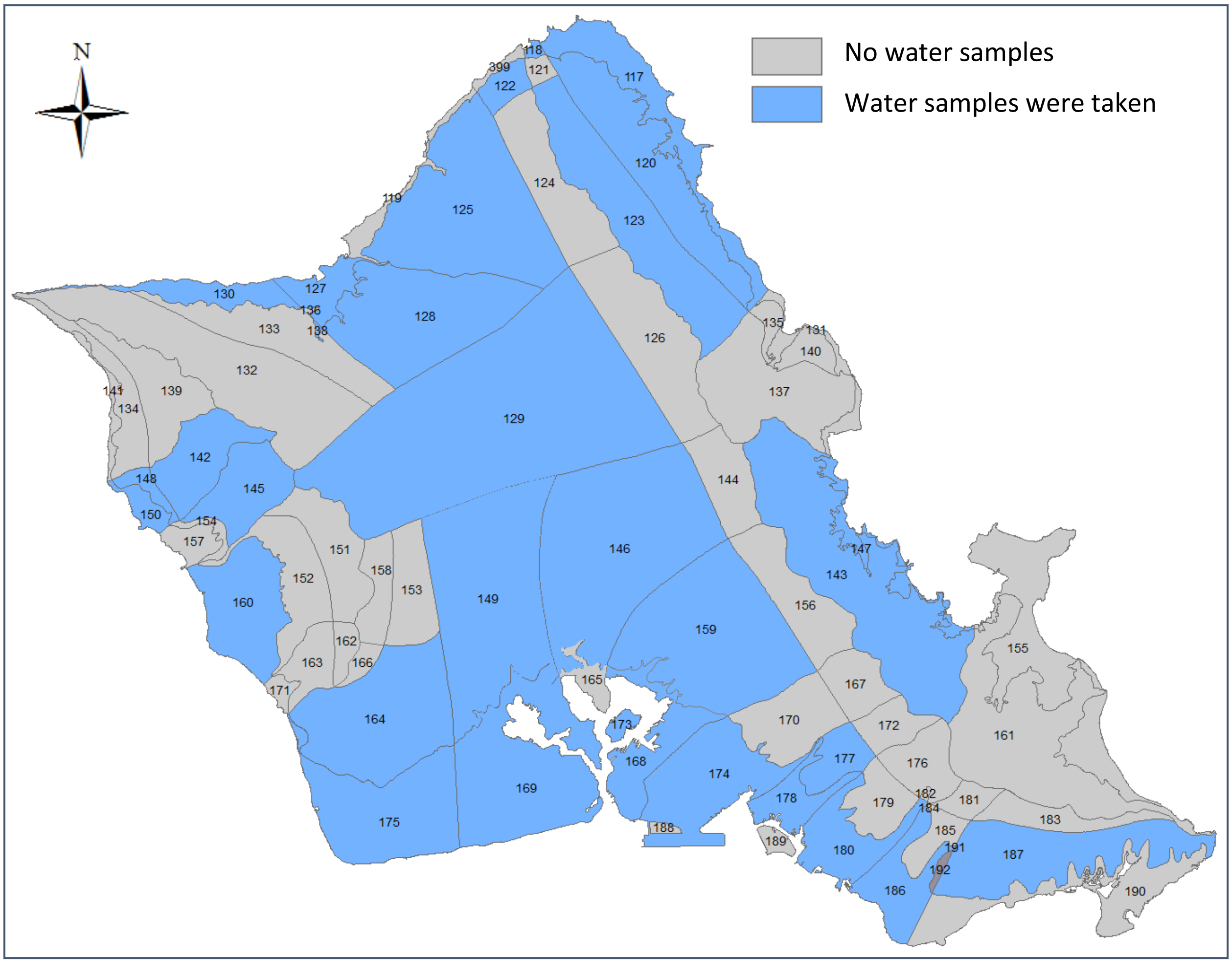
Post hoc Poisson regression models examining water-quality constituents (significant in the Spatial VS methods) associated with NTM infection risk in Oahu, HI.

Bolded estimates are statistically significant ($p < 0.05$). CI = Confidence Interval

MAC species	
Variable*	Relative Risk (95% CI) p-value
Aluminum (1-log unit)	**0.85** **0.73, 0.99** **(0.033)**
Boron (1-log unit)	**1.22** **1.09, 1.36** **(0.0004)**
Bromide (1-log unit)	1.08 0.98, 1.19 (0.104)
Calcium (1-log unit)	1.06 0.97, 1.17 (0.176)
Magnesium (1-log unit)	1.06 0.95, 1.18 (0.299)
Sodium (1-log unit)	1.11 0.99, 1.26 (0.262)
Strontium (1-log unit)	1.06 0.96, 1.16 (0.081)
Zinc (1-log unit)	0.98 0.87, 1.12 (0.807)

*Each row indicates a separate single-constituent Poisson model. All models are controlled for sex.

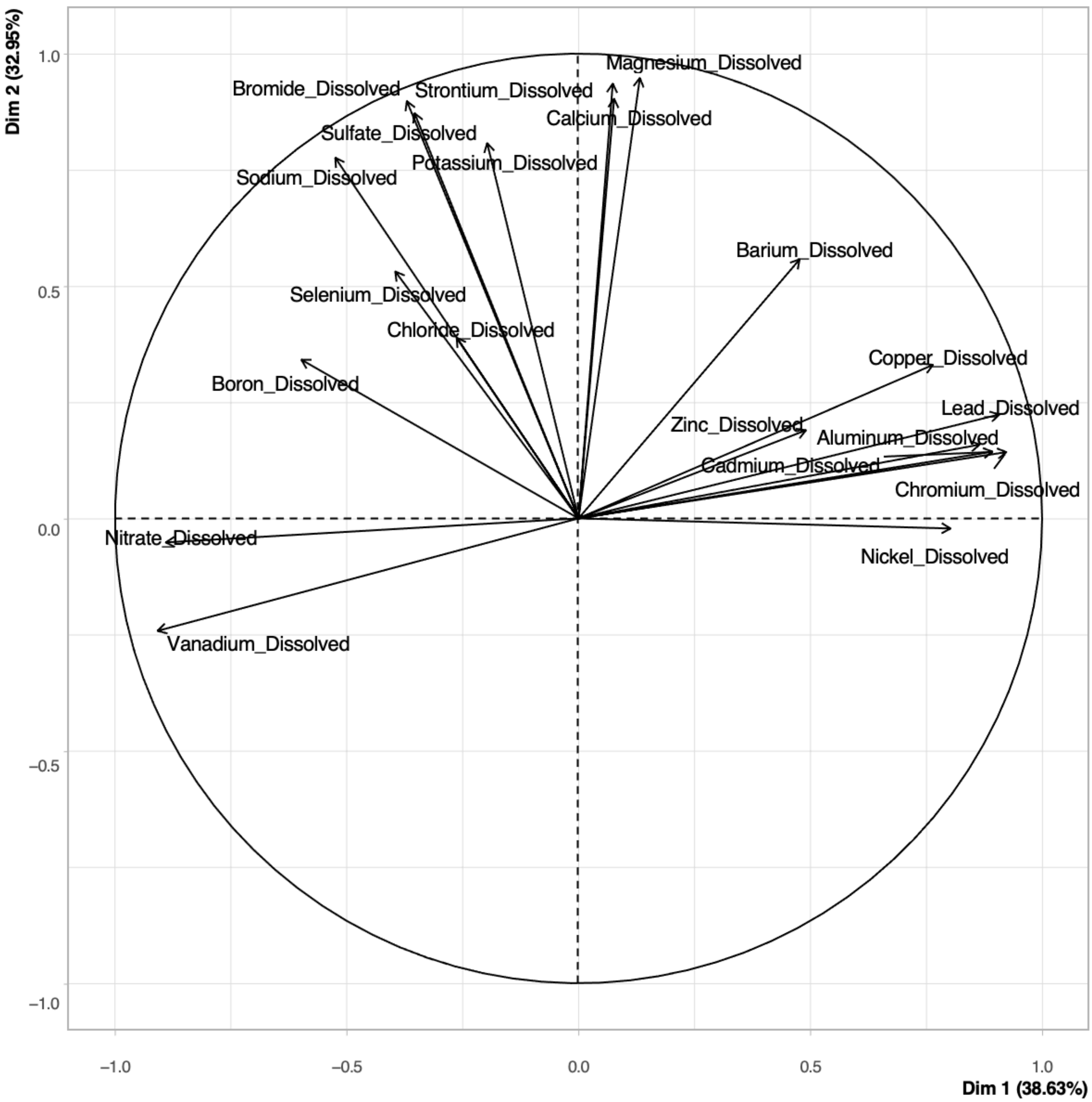
Supplementary Figure 1. Aquifers* with and without water samples in Oahu, HI.



Numbers refer to “objectid” for each aquifer (https://geoportal.hawaii.gov/datasets/doh-aquifers-polygons/).

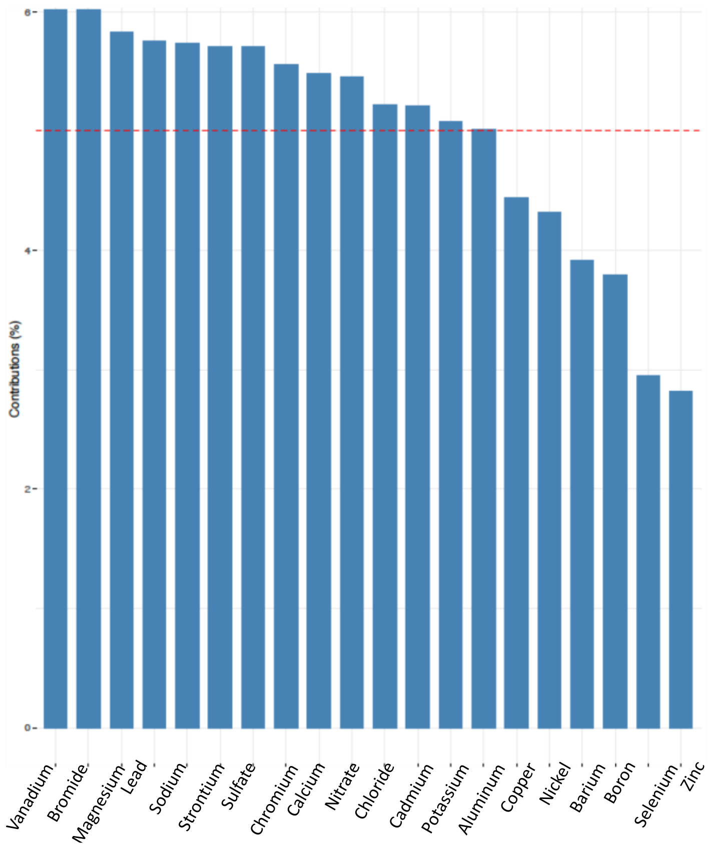
Supplementary Figure 2.

Principal Components Analysis (PCA) graph of variables



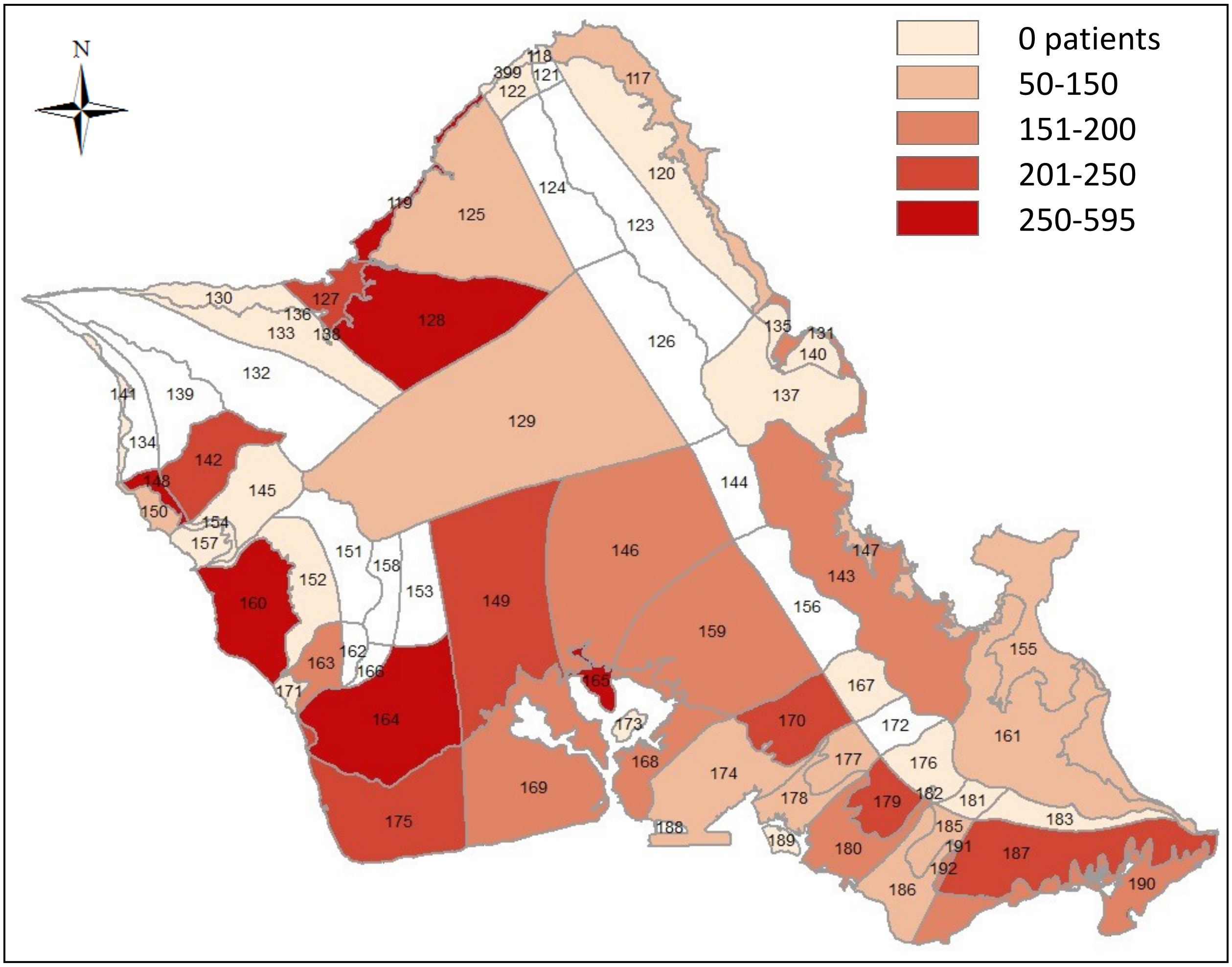
Supplementary Figure 3.

Contribution of water-quality constituents to principal components 1-3.



Supplementary Figure 4.

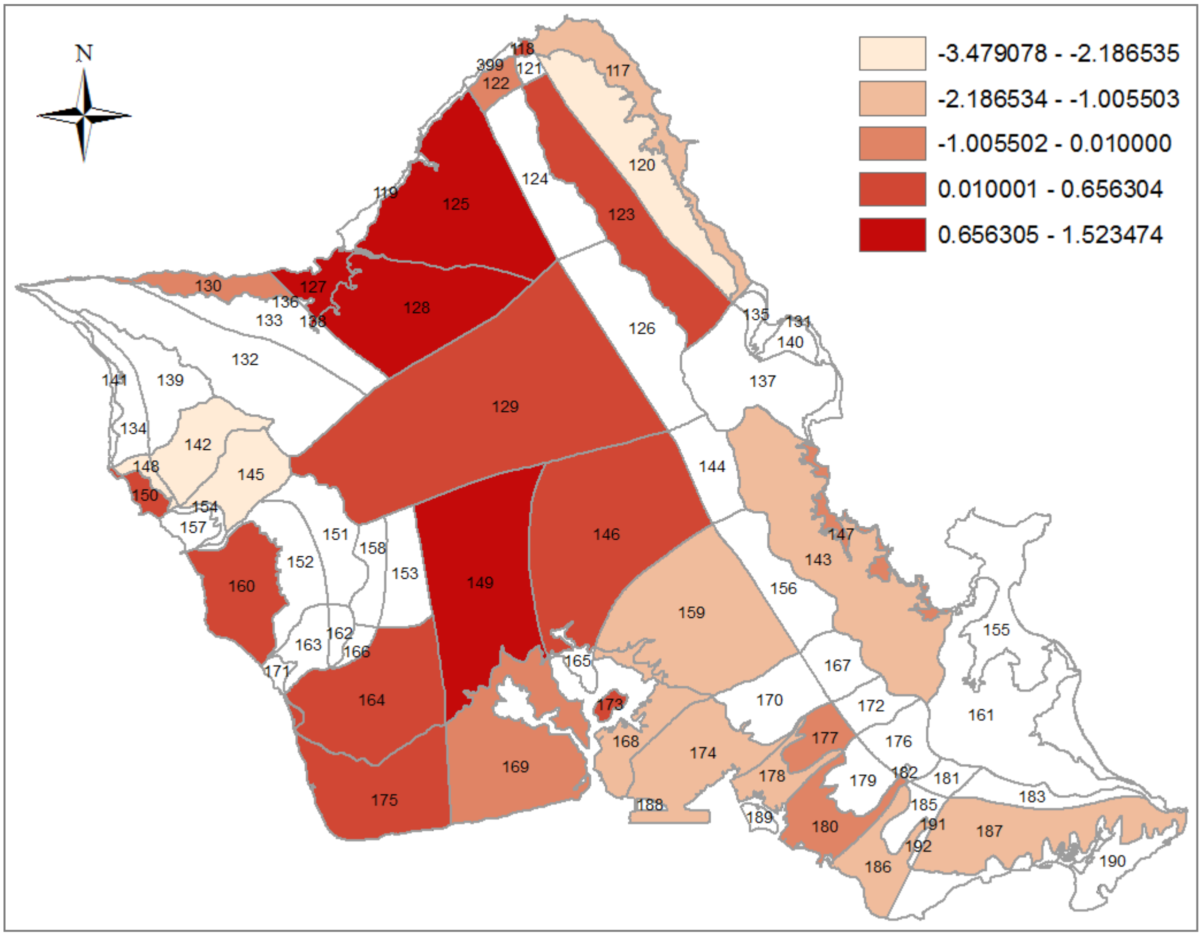
MAC incidence/100,000 per aquifer* in Oahu, HI.



*Numbers refer to "objectid" for each aquifer (https://geoportal.hawaii.gov/datasets/doh-aquifers-polygons/).

Supplementary Figure 5.

Vanadium concentrations* per aquifer# in Oahu, HI.



*Median concentrations are natural log transformed, standardized, and imputed (if missing).
#Numbers refer to "objectid" for each aquifer (https://geoportal.hawaii.gov/datasets/doh-aquifers-polygons/).