Vanadium in groundwater aquifers increases the risk of MAC pulmonary infection in O'ahu, Hawai'i.

Ettie M. Lipner, Ph.D., M.P.H.[1,2], Joshua P. French, Ph.D.[3§], Stephen Nelson, Ph.D.[4§], Joseph O. Falkinham III, Ph.D.[5], Rachel A. Mercaldo, Ph.D.[1], Rebekah A. Blakney, M.S.[1], Yihe G. Daida, Ph.D.[6], Timothy B. Frankland, M.A.[6], Kyle P. Messier, Ph.D.[7], Jennifer R. Honda, Ph.D.[2], Stacey Honda, M.D., Ph.D.[6], D. Rebecca Prevots, Ph.D., M.P.H. [1]

[§]Equal contribution of authors

[1]Epidemiology and Population Studies Unit, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA, [2]Center for Genes, Environment and Health, National Jewish Health, Denver, CO, USA (Former affiliation for EML), [3]Department of Mathematical and Statistical Sciences, University of Colorado Denver, Denver, CO, USA; [4]Department of Geological Sciences, Brigham Young University, Provo, UT, USA [5]Virginia Tech, Blacksburg, VA, USA; [6]Center for Integrated Health Care Research, Kaiser Permanente Hawaii, Honolulu, Hawaii, USA, [7]National Institute of Environmental Health Sciences, National Institutes of Health, Durham, NC, USA.

Conflicts of interest: None declared.

Running title: Vanadium in groundwater increases MAC risk in Hawai'i.

Source of Funding: This work was supported by NSF award 1915277 for JPF; the Padosi Foundation for JRH, YGD, TBF, SH; the Division of Intramural Research, NIAID for EML, RAM, RAB, DRP; intramural funds from NIEHS/DNTP and NIMHD for KPM. Authors receive funding to support open access publishing.

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: Hawai'i has the highest prevalence of nontuberculous mycobacterial (NTM) pulmonary disease in the United States. Previous studies indicate that certain trace metals in surface water increase the risk of NTM infection.

Objective: To identify whether trace metals influence the risk of NTM infection in O'ahu, Hawai'i.

Methods: A population-based ecologic cohort study was conducted using NTM infection incidence data from patients enrolled at Kaiser Permanente Hawai'i (KPH) during 2005-2019. We obtained sociodemographic, microbiologic and geocoded residential data for all KPH beneficiaries. To estimate the risk of NTM pulmonary infection from exposure to groundwater constituents, we obtained groundwater data from three data sources: 1) Water Quality Portal, 2) the Hawai'i Department of Health, and 3) Brigham Young University, Department of Geological Science faculty. Data were aggregated by aquifer and were associated with corresponding beneficiary aquifer of residence. We used Poisson regression models with backward elimination to generate models for NTM infection risk as a function of groundwater constituents. We modeled two outcomes: *Mycobacterium avium* complex (MAC) species and *Mycobacterium abscessus* group species.

Results: For every 1-unit increase in the log concentration of vanadium in groundwater at the aquifer level, infection risk increased by 22% among MAC patients. We did not observe significant associations between water-quality constituents and infection risk among *M. abscessus* patients.

Conclusions: Concentrations of vanadium in groundwater were associated with MAC pulmonary infection in O'ahu, Hawai'i. These findings provide evidence that naturally occurring trace metals influence the presence of NTM in water sources that supply municipal water systems.

Abstract word count: 250

Manuscript word count: 4,000

What this study adds

This is the first study to examine the relationship between water-quality constituents in groundwater aquifers and NTM pulmonary infection among Hawai'i residents. Our findings suggest that specific characteristics of groundwater, which supplies the municipal drinking water systems, may influence NTM growth, thereby increasing the risk of NTM exposure and disease. Identification of environmental factors may provide insight into the influence of NTM growth and infection and is critical for the development of prevention strategies to minimize exposure and infection in high-risk regions.

1. INTRODUCTION

Nontuberculous mycobacterial (NTM) pulmonary disease continues to increase in the United States (U.S.) [1] The disease is characterized by non-specific symptoms including chronic cough and fatigue; the incubation period is unknown, and many patients experience delays of up to 5 years between onset of symptoms and disease diagnosis. [2] The NTM represent a diverse group of organisms, but recent studies indicate that 85% of clinical isolates are MAC or *M. abscessus*/*M. chelonae*. NTM are ubiquitous in the environment, [3,4] although the distribution of NTM pulmonary infection (NTM PI) and disease (NTM PD) varies substantially by geographic region. [5-7] Host-related factors, e.g., cystic fibrosis and demographic characteristics, contribute to an individual's susceptibility to NTM infection, but they do not explain the geographic variation of NTM infection, as a much higher prevalence of infection is still observed in certain geographic areas even after adjusting for these factors. [5] This geographic variation could be due to local environmental differences [6] placing humans at increased exposure, and in turn, increased risk. Our previous research demonstrated that increased vanadium concentrations in surface water in Oregon was associated with a higher risk of *Mycobacterium avium* complex (MAC) infections, and that increasing molybdenum concentrations in surface water in Colorado and Oregon was associated with a higher risk of *Mycobacterium abscessus* infection. [8,9] We hypothesize that in the environment, these trace metals are required for mycobacterial metabolism, as evidence exists that molybdenum serves as a cofactor for nitrate reductase.[10-12] Moreover, a significant correlation has been found between mycobacterial abundance and disease prevalence in the CF and non CF population. [13] Here, we investigated the relationship between water-quality constituents in groundwater and the risk of NTM infection in Hawai'i (HI), the state with the highest prevalence of NTM disease in the U.S. [14] We associated water-quality data from a combined dataset of three data sources on O'ahu with MAC and *M. abscessus* infection data from Kaiser Permanente beneficiaries residing in O'ahu, HI.

2. METHODS

2.1 Data Collection

2.1.1 NTM Patient & Beneficiary Data

This retrospective population-based cohort study utilized de-identified patient data from Kaiser Permanente Hawaii (KPH). Approval was obtained from the KPH Institutional Review Board (#00000402) and the National Jewish Health Institutional Review Board (HS-3479). We included KPH beneficiaries for the study period January 1, 2005 through December 31 2019, aged ≥18 years at study entry who resided in HI for at least one year prior to the year of sputum culture collection and who were residents of HI during the year in which the culture was collected. We extracted residential latitude and longitude coordinates rounded to two decimal places (±750 m precision) for the KPH beneficiary population, as well as age, sex, self-identified ethnicity, and neighborhood deprivation index (NDI) assigned to a person's census tract. We calculated mean age, proportion male/female, proportion of White only, Asian only and Native Hawaiian or Pacific Islander (NHOPI) only for all beneficiaries within each aquifer. Groups were termed "only" when a participant identified with either Asian, White, or NHOPI groups exclusively. KPH assigned each beneficiary an NDI value for the census tract in which each beneficiary resided. We calculated the mean NDI value for each aquifer based on the beneficiaries who resided there; higher values indicated greater neighborhood deprivation, which ranged from -0.99 to 1.73. Using residential latitude and longitude coordinates, we geocoded each beneficiary to an aquifer, and aggregated the KPH beneficiary population by aquifer. This beneficiary population was used as the at-risk population (log offset) in our regression models.

We selected KPH patients with incident NTM pulmonary infections, based on ATS microbiologic criteria. [15] We combined probable and confirmed cases because we found their characteristics to be similar (Blakney et al. EID In Press). We defined incident cases as any patient over age 18 without any prior respiratory cultures positive for NTM in the year prior to at least one NTM-positive pulmonary culture. We included only cultures from beneficiaries who were residents of HI for at least one year prior to the year of culture collection and were residents of HI during the year in which the culture was collected. We included only patients with either MAC or *M. abscessus* infection.

2.1.2 Environmental Exposure Data

Aquifer boundaries were obtained from the Hawaii Geospatial Data Portal. [16] These aquifers are defined for both administrative purposes and for conceptualizing aquifer compartments which may have different properties or physical boundaries. A basal aquifer system extends across the entire island of O'ahu, where water flows from one "aquifer" to another; [17,18] however, the large number of defined aquifers for O'ahu provides a fine-scale subdivision of the groundwater system useful for our statistical approach.

Groundwater data were extracted from the Water Quality Portal (WQP) [19] for the period January 1, 1995 through December 31, 2019. Appendix 1 details the groundwater data extraction procedure and data cleaning steps that were taken to go from the raw data to the cleaned dataset used in the analysis. HI groundwater data also came from water sampling datasets collected through Dr. Stephen Nelson's laboratory at Brigham Young University (BYU) between May

2007 and October 2008. Finally, we obtained a statewide water sampling dataset from the HI Department of Health (HI DOH) comprising raw groundwater and treated water measurements from January 1, 2010 to November 18, 2020. In this dataset, well water data were separated from treated water data. HI DOH well water data were combined with the WQP and the BYU datasets.

2.2 Data Compilation

2.2.1 Water Data

We focused our analyses on the island of O'ahu due to insufficient water sampling on the remaining islands. Data were analyzed using various R packages (see Online Supplement). We excluded aquifers from the dataset for which no water samples were collected. Our water-quality constituent dataset contained water sampling data from 32 out of 77 aquifers on O'ahu (Supplementary Figure 1) that represent most of the population and a majority of the land area on O'ahu. A large fraction of the 45 remaining aquifers underlies mountainous areas or lands set aside from development as state, federal and private lands.

On O'ahu, a total of 3,541 total water samples were tabulated from 128 unique sampling sites. We calculated aquifer-specific median values for each constituent from the combined water dataset (WQP, BYU, HI DOH) (see Online Supplement). Individual water-quality constituents were eliminated if data were not available for more than 50% of the 32 aquifers or if the median values for each aquifer were all zero for a given constituent (Supplementary Table 1, Supplementary Table 2). Following data curation steps (see Online Supplement), 20 water-quality constituents remained for analysis (Supplementary Table 2 & Supplementary Figure 3). The aquifer-median values for these 20 constituents were first natural log transformed and then standardized to have a mean of 0 and standard deviation of 1. Finally, we imputed the median value for aquifers with missing values (see Online Supplement).

2.2.2 Principal Components Analysis (PCA)

Principal Component Analysis (PCA) [20] was performed on the aquifer-level (Supplementary Figure 2). We retained the top three principal components, [21,22] which explained 78.5% of the data variability. We used the fviz_contrib function in the **factoextra** package to identify the most important variables in explaining variability of the top three principal components [22] (Online Supplement). Any constituent with a contribution above the reference line (red dashed line in Supplementary Figure 3) was considered important in contributing to the principal components. [23,24]

2.3 Statistical Analysis

2.3.1 Poisson Regression models with Individual Water-Quality Constituents from Top 3 Principal Components.

We used Poisson regression models with backward elimination to generate models for NTM PI risk as a function of water-quality constituents in groundwater. NTM infection risk, i.e., NTM incidence, was defined as the number of incident cases in a given aquifer during our study period as the numerator, and the number of KPH beneficiaries living in that aquifer as a denominator (modeled in R using the log of the beneficiary population as the offset). Because incidence is synonymous with risk, we use the term risk to express the exposure-disease association. Aquifer-level median values of each water-quality constituent as well as age, sex, ethnic group, and NDI were included as predictor variables. We estimated the risk of NTM PI given exposure to water-quality constituents in groundwater sources, with statistical significance assessed at $p<0.05$. We modeled two separate outcomes variables: NTM PI associated with MAC species, and *M.*

abscessus species, as a function of water-quality constituents and other covariates (Supplementary Table 1). We calculated the variance inflation factor (VIF) for each water-quality constituent in each model and included only those water-quality constituents with VIFs less than 10 in a p-value-based backward elimination procedure to mitigate the potential effect of collinear predictors (Model 1). We then constructed separate single-constituent Poisson regression models (Models 2 & 3) for the metals that demonstrated statistical significance from Model 1 ($p<0.05$). To create Figure 1, we used the best-fit estimates of the aquifer-specific risks from the single-constituent Poisson model for vanadium and risk of MAC infection (Model 2). This approach has been described in our prior studies. [8,9,25] Additionally, we used a spatial variable selection (Spatial VS) approach proposed by Xie et al. [26] to estimate a spatial Poisson regression model that simultaneously selects important covariates while adjusting for possible spatial correlation and multicollinearity among the explanatory variables to further evaluate our results from the GLM with backward elimination (Online Supplement).

3. RESULTS

3.1 Study Population Characteristics

Our at-risk population comprised 193,284 KPH beneficiaries who lived in O'ahu for at least 2 years at any point during the years 2005 through 2019 and did not have any pulmonary cultures positive for pathogenic NTM infection. Among our cases, we had 402 patients with incident MAC infection and 136 patients with incident *M. abscessus* infection. Demographic characteristics of MAC and *M. abscessus* cases and beneficiaries are shown in Table 1. Overall, our groundwater dataset contained water samples collected from 32 out of 77 aquifers (Supplementary Figure 1), and our beneficiary population resided in 58 out of 77 aquifers. Only the 32 aquifers with water data were included in our regression analyses.

This restriction left 348 MAC patients, 122 *M. abscessus* patients, and 163,560 beneficiaries available for analysis. Thus, 13% of all MAC patients and 10% of all *M. abscessus* patients on O'ahu were excluded from the analysis (see Online Supplement, section 3.1, for greater detail).

3.2 Regression Models with Individual Water-Quality Constituents from Top 3 Principal Components

We identified 14 out of 20 constituents that were important contributors to the top three principal components (see Online Supplement): aluminum, bromide, cadmium, calcium, chloride, chromium, lead, magnesium, nitrate, potassium, sodium, strontium, sulfate and vanadium (Supplementary Figure 3). We modeled the risk of NTM PI as a function of these 14 water-quality constituents in a Poisson regression model and selected water-quality constituents whose variance inflation factors (VIF) were below 10 to mitigate the potential impact of collinear covariate constituents. In each model, we sequentially removed the constituent with the highest VIF and reran the model until all constituents had VIFs less than 10. Bromide, calcium, lead, magnesium, and sodium had VIF values greater than 10, and were thus omitted from the final model (Table 2; Model 1). The correlation matrix for water-quality constituents is available in Supplementary Table 3.

For MAC, we found that vanadium and sulfate were significantly positively associated with infection risk (Table 2; Model 1). For *M. abscessus*, chloride, and sulfate were significantly positively associated with infection risk, whereas potassium was negatively associated with infection risk (Table 2; Model 1). We then modeled the risk of MAC infection (Table 3; Model 2) as a function of each significant metal from Model 1 in separate single-constituent models ($p < 0.05$); vanadium and sulfate remained statistically significant, although the strength of association was much stronger for vanadium than for sulfate. For every 1-log unit increase in sulfate and vanadium concentrations in groundwater at the aquifer-level, the risk of MAC infection increased by 12% and 22%, respectively. The effect of increasing vanadium concentration (ranging from 10.9 µg/L to 50.5 µg/L) on infection risk remained statistically significant (p = 0.0003) after controlling for multiple comparisons using the Bonferroni method (7 models; adjusted significance level is p = 0.05/7 = 0.007) but increasing sulfate concentration did not remain significant after controlling for multiple testing (p = 0.043). When modeling *M. abscessus* infection as a function of each significant metal from Model 1 in single-constituent models ($p < 0.05$), none of the water-quality constituents remained statistically significant (Table 3; Model 3).

In a Poisson model with backward elimination, we also performed a sensitivity analysis by including only aquifers with beneficiary populations ≥ 150. Here sulfate was no longer statistically significant, while the significant effect of vanadium on NTM risk for MAC infection became greater (Supplementary Table 4). Spatial models presented in the Online Supplement confirmed our results for vanadium. Other constituents were also deemed significant by this approach, therefore we conducted post hoc GLM analyses on these individual constituents. Only aluminum and boron remained statistically significant, although these estimates were not consistent across the spatial and GLM approaches (Online Supplement).

3.3 MAC Risk by Aquifer

We estimated risk of NTM PI by aquifer for MAC using a Poisson model adjusted for sex. Figure 1 shows the risk estimates per aquifer based on an adjusted Poisson model for MAC infection with vanadium included as an independent predictor. We observed the lowest risk estimates to the north and the south of the island, whereas the highest risk estimates were observed in the center of the island. Supplementary Figures 4 & 5 show MAC rates and vanadium concentrations per aquifer on O'ahu, respectively.

4. DISCUSSION

We found that increasing concentrations of vanadium in groundwater were significantly associated with an increased risk of MAC infection on O'ahu. As groundwater is the sole source of HI's public drinking water supplies, [27-29] we used groundwater samples from within aquifer boundaries as the main source of water exposure in this study.

Vanadium is a naturally occurring element; [30] it is found in soil, water, and air, as well as in over 150 different minerals, iron ores and crude petroleum deposits. [30,31] The primary industrial use of vanadium is in the steel industry where it is used as an alloy to strengthen steel. [30] Vanadium is released into the environment from both natural sources and anthropogenic activity. The natural release of vanadium into water and soil occurs primarily from the weathering of rocks and soil erosion. The concentration of vanadium in soil depends on the type of soil found within a geographic region. [30] The Hawaiian Islands are predominantly composed of basalt. [32,33] Basalts generally contain high concentrations of vanadium compared with other common rock types and the process of weathering determines the distribution of vanadium throughout the soil on the islands. [33] Vanadium will largely reside in magnetite and ilmenite in fresh volcanic rock. [34] As previously reported, [9] we observed a strong significant association between vanadium and MAC infection in Oregon; in the Pacific Northwest, where basalt is also a common rock type, including the Columbia River Basalts of eastern Washington and Oregon, [35] as well as younger and older mafic (i.e., Fe- and Mg-rich) rocks distributed throughout these states. [36-38]

The concentration of vanadium found in groundwater depends primarily on geographical location and the type of surrounding soil in the region. [39] In a U.S. Geological Survey (USGS) assessment of water-quality on O'ahu, the presence of vanadium in streams and groundwater was due to natural abundance in volcanic rocks and soil and varied little by land use or because of human activities. [29] In HI, the presence of vanadium in drinking water is so well known that the Hawai'i State Department of Health issued a statement in 2014 indicating that vanadium is naturally occurring in its drinking water and is not considered to be harmful. [40] The EPA is examining the prevalence and concentrations of vanadium in the U.S. drinking water supplies. [41] Currently, no federal drinking water regulations for vanadium exists in the U.S. [30]

Vanadium could influence either host susceptibility or mycobacterial metabolism; we have no evidence related to host susceptibility, rather we postulate that vanadium is influencing NTM abundance or metabolism (e.g., virulence [10,42]) in natural water sources. In regions where vanadium concentrations are high, NTM abundance (or more pathogenic species) is likely increased in the water supply, thereby increasing overall risk of exposure and infection. In a previous study in HI, Nelson et al. [43] examined possible routes that NTM may take from the environment into groundwater, subsequently into the public water supplies, and ultimately into homes. The authors demonstrated through aquifer flow models that NTM may survive within the water supply over the course of months via interconnected fractured basalt networks before entering domestic water supplies. Specific aquifer networks may have higher vanadium concentrations, which may influence greater NTM growth as the flow of water moves through the network of aquifers and water wells.

While there is a considerable amount of vanadium in Hawaiian soils, these weathered materials are not likely to be the primary source of this element in groundwater. As a trace metal, vanadium will reside primarily in titanium-rich minerals in basalt,[34] including magnetite, ilmenite, and clinopyroxene. However, magnetite and ilmenite do not weather readily,[44,45] so vanadium is retained in refractory minerals in soil and within fresh basalt. Clinopyroxene, on the other hand, weathers readily[46] and over the last ~2 Ma (million years) this mineral has leached to form Fe-oxides/hydroxides and clay. Although largely eliminated from soil, clinopyroxene will persist in the relatively fresh basalts that comprise the aquifers. In fact, Hawaiian clinopyroxenes may contain ~1 wt. % TiO_2 and ~200 ppm vanadium.[47] Thus, water-clinopyroxene reactions within the aquifer are the likely source of vanadium in water supplies.

Water-rock contact time is also a primary factor in determining the solute content of waters in HI[46] as older waters have had more time to overcome kinetic barriers and leach solutes from host basalt. Patterns of vanadium concentration as a function of water table elevation show some interesting patterns (Fig. 2). Waters with low to intermediate abundances tend to occur in wells adjacent to the base of the Koolau and Waianae Mountains where recharge occurs, consistent with short flow paths and residence times. Wells with intermediate to high vanadium concentrations tend to occur where longer residence times are produced by longer flow paths, supporting the hypothesis that a primary control on vanadium is the length of time that groundwater is in contact with basalt in the aquifer.

To test this hypothesis, we correlated mean groundwater elevation with median vanadium concentrations by aquifer and found a statistically significant negative correlation ($\rho = -0.44$, 95% CI [-0.68, -0.10], $p = 0.014$). Concentrations of vanadium were lower at high water table elevations, whereas vanadium concentrations were higher at low water table elevations. This suggests that groundwater age may be an effect modifier in the relationship between vanadium in groundwater and MAC infection. The longer the duration that groundwater travels within the basalt aquifers, vanadium concentrations will increase in groundwater, resulting in higher growth of MAC, and ultimately increasing risk of infection.

Vanadium in domestic water may be just one risk factor for opportunistic NTM infection. For example, previous work has shown a relationship between different types of soil and NTM presence and growth. Glickman et al. [48] demonstrated that *M. avium* growth was observed when the bacteria were incubated with various iron-rich minerals, including hematite, maghemite and magnetite. According to a USGS vanadium report, [30] these iron-rich minerals contain elevated concentrations of vanadium. It is possible, however, that the contact between vanadium and NTM is not a result of soil directly, but rather from the fractured basalt where the magnetite and ilmenite minerals are already weathered. In addition, NTM have been isolated from volcanic ash samples in HI; [49] it should be noted basaltic volcanic ash is inherently vanadium-rich. [33]

We hypothesize that these metals stimulate NTM growth, metabolism, and numbers. The waters sampled on O'ahu may be nitrogen-deficient and scavenging for nitrogen is required for survival and growth of the microbial flora; some NTM species, notably *M. avium* can grow on nitrate or nitrite as a sole source of nitrogen. [12] The enzyme responsible for growth on nitrate or nitrite as a sole nitrogen source is nitrate reductase. [12,50] Nitrate reductases contain either molybdenum or vanadium as essential co-factors. [42,51,52] We note the correlation between nitrate and vanadium concentrations (Supplementary Table 3) to support our hypothesis that vanadium is functioning as a cofactor for nitrate utilization. Experiments to document the role of trace metals in microbial growth are difficult, as increases or decreases would be expected to have minimal impacts. Thus, to overcome any limited response of NTM to trace metals, the role of molybdenum and vanadium as co-factors for the enzyme nitrate reductase in supporting nitrate- or nitrite-dependent growth of *M. avium* can be measured.

This study has confirmed our previous finding of the role of vanadium [9] in NTM infection, using an independent patient population in a separate geographic region of the U.S. We previously reported that increasing vanadium concentrations in surface water were associated with increased risk of MAC infection in Oregon. [9] We also found that molybdenum in surface water was significantly associated with increased risk of *M. abscessus* infection in Colorado and Oregon. [8,9,25] In this study, we could not assess the effect of molybdenum in groundwater on *M. abscessus* infection risk because of insufficient molybdenum measurements across aquifers on O'ahu, However, we did not observe any significant findings among other water-quality constituents in our *M. abscessus* models (Model 3). Further studies in other geographic regions are necessary to confirm the molybdenum and *M. abscessus* infection hypothesis.

Our study has several limitations. The water sampling dataset included measurements from only 42% of the total number of aquifers on O'ahu, although these aquifers cover a majority of the land area on the island and its population (Supplementary Figure 1). In addition, the frequency of water sampling was not extensive. For all constituents in the dataset, often only one sample was collected per aquifer, which may limit the generalizability of water exposure from that particular constituent. Finally using a separate spatial approach, we also identified that aluminum and boron were significantly associated with MAC infection (Online Supplement). However, their estimates were not consistent across spatial and GLM approaches, and aluminum did not remain significant after adjusting for multiple testing. Additionally, boron is rarely found in basalt; it is probably derived from sea spray aerosols as boron occurs in seawater in concentrations of 5-6 mg/L. [53]

Despite limitations inherent to our water dataset, we have observed positive robust associations between vanadium and MAC infection in both Oregon [9] and Hawai'i. This study's main question relates to whether naturally occurring trace metals influence NTM abundance or metabolism in natural water sources. In the case of O'ahu, increased vanadium content is likely associated with longer water-basalt contact times at low water table elevations. Elevated vanadium concentrations may promote NTM growth. These natural water sources supply water to municipal water systems, which in turn may increase the risk of exposure and infection for MAC or *M. abscessus* for the populations served by those water utilities.

5. CONCLUSION

We anticipate that these findings will generate new research questions. For example: Do trace metals influence NTM growth *in vitro*? Is there a dose-response relationship between certain trace metals and NTM growth? Do these trace metals influence host susceptibility to infection and/or disease progression? While many in the NTM patient community inquire about exposure-reduction strategies to prevent reinfection, little consensus and scant evidence exists in the medical and scientific communities on how to prevent infection and reinfection, leaving clinicians with only anecdotal strategies to impart to patients. We hope that this line of research will establish a more comprehensive and systematic framework of the environmental factors that predict NTM exposure and infection, with the ultimate goal of informing prevention strategies.

ACKNOWLEDGEMENTS

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

References

1. Winthrop KL, Marras TK, Adjemian J, Zhang H, Wang P, Zhang Q. Incidence and Prevalence of Nontuberculous Mycobacterial Lung Disease in a Large U.S. Managed Care Health Plan, 2008-2015. *Ann Am Thorac Soc.* Feb 2020;17(2):178-185.
2. Kim RD, Greenberg DE, Ehrmantraut ME, et al. Pulmonary nontuberculous mycobacterial disease: prospective study of a distinct preexisting syndrome. *Am J Respir Crit Care Med.* Nov 15 2008;178(10):1066-1074.
3. Falkinham JO, 3rd. Surrounded by mycobacteria: nontuberculous mycobacteria in the human environment. *J Appl Microbiol.* Aug 2009;107(2):356-367.
4. Falkinham JO, 3rd. Environmental sources of nontuberculous mycobacteria. *Clin Chest Med.* Mar 2015;36(1):35-41.
5. Adjemian J, Olivier KN, Prevots DR. Nontuberculous mycobacteria among patients with cystic fibrosis in the United States: screening practices and environmental risk. *Am J Respir Crit Care Med.* Sep 1 2014;190(5):581-586.
6. 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.
7. 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.
8. 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.
9. 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.
10. Zhong Q, Kobe B, Kappler U. Molybdenum Enzymes and How They Support Virulence in Pathogenic Bacteria. *Front Microbiol.* 2020;11:615860.
11. Shi T, Xie J. Molybdenum enzymes and molybdenum cofactor in mycobacteria. *J Cell Biochem.* Oct 2011;112(10):2721-2728.
12. Butala NS, Falkinham JO. Nitrate and nitrite reductase activities of Mycobacterium avium. *Int J Mycobacteriol.* Oct-Dec 2018;7(4):328-331.
13. 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).
14. 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.
15. Daley CL, Iaccarino JM, Lange C, et al. Treatment of Nontuberculous Mycobacterial Pulmonary Disease: An Official ATS/ERS/ESCMID/IDSA Clinical Practice Guideline. *Clin Infect Dis.* Aug 14 2020;71(4):e1-e36.

16. Hawaii Statewide GIS Program GDP. DOH Aquifers (Polygons). 2017; https://files.hawaii.gov/dbedt/op/gis/data/aquifers_doh.pdf. Accessed February 12, 2021.

17. Gingerich SB, & Oki, D. S. . Ground water in Hawai'i. U.S. Geological Survey Fact Sheet FS. In: U.S. Geological Survey USotI, ed2000:6.

18. Whittier RB, Rotzoll, K., Dhal, S., El-Kadi, A.I., Ray, C., Chang, D. Groundwater source assessment program for the state of Hawaii, USA: methodology and example application. . *Hydrogeology Journal.* 2010;18(3):711-723.

19. U.S. Gelogical Survey NWQMC. Water Quality Portal. 2012; https://www.waterqualitydata.us/portal/. Accessed August 8, 2019.

20. King JR JD. VARIABLE SELECTION IN LARGE ENVIRONMENTAL DATA SETS USING PRINCIPAL COMPONENTS ANALYSIS. *Environmetrics.* 1999;10:67-77.

21. Kaiser HF. The application of electronic computers to factor analysis. *Educational and Psychological Measurement.* 1960;20:141-151.

22. Cattell RB. The scree test for the number of factors. *Multivariate Behavioral Research.* 1966;1:245-276.

23. A. K. *Practical Guide To Principal Component Methods in R: PCA, M(CA), FAMD, MFA, HCPC, factoextra.* Vol 2: STHDA (http://www.sthda.com); 2017.

24. analysis SStfh-tpd. Articles - Principal Component Methods in R: Practical Guide. CA - Correspondence Analysis in R: Essentials. http://www.sthda.com/english/articles/31-principal-component-methods-in-r-practical-guide/113-ca-correspondence-analysis-in-r-essentials/.

25. 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).

26. Xie Y, Xu, L., Li, J., Deng, X., Hong, Y., Kolivras, K. and Gaines, D.N. Spatial variable selection and an application to Virginia Lyme disease emergence. *Journal of the American Statistical Association.* 2019;114(528):1466-1480.

27. Survey USG. Ground Water in Hawaii. In: Interior USDot, ed2016.

28. Supply BoW. Oahu Groundwater Levels. *Data & Statistics* 2021.

29. Anthony SS, Hunt, Jr., Charles D., Brasher, Anne M., Miller, Lisa D., and Tomlinson, Michael S. Water quality on the island of Oahu, Hawaii: Reston, Va., U.S. Geological Survey Circular 12392004:37.

30. Kelley KD SC, Polyak DE, Kimball BE. Vanadium, Chapter U of Schulz, K.J., DeYoung, J.H., Jr., Seal, R.R., II, and Bradley, D.C., eds., Critical mineral resources of the United States—Economic and environmental geology and prospects for future supply: U.S. Geological Survey Professional Paper 1802. In: U.S. Department of the Interior USGS, ed2017:U1-U36.

31. (ATSDR) AfTSaDR. Toxicological profile for Vanadium. In: Department of Health and Human Services PHS, Centers for Disease Control and Prevention, ed. Atlanta, GA2012.

32. Volcanology HCf. The Formation of the Hawaiian Islands. 2018; https://www.soest.hawaii.edu/GG/HCV/haw_formation.html.

33. GD NMS. The Vanadium Content of Hawaiian Island Soils. *Hawaii Agricultural Experiment Station.* 1961;Technical Bulletin No. 45.

34. Best MG. *Igneous and metamorphic petrology.* . 2nd ed: John Wiley & Sons.; 2013.
35. Camp VE, Reidel, S.P., Ross, M.E., Brown, R.J., and Self, S. Field-trip guide to the vents, dikes, stratigraphy, and structure of the Columbia River Basalt Group, eastern Oregon and southeastern Washington: U.S. Geological Survey Scientific Investigations Report 2017–5022–N. 2017:88.
36. Walker GW, MacLeod, N.S. Geologic Map of Oregon. In: U.S. Geological Survey USotI, ed. Reston, VA1991.
37. Raines GL, Sawatzky, D.L., Connors, K.A. Great Basin geoscience data base: Digital Data Series DDS-41. In: U.S. Geological Survey USotI, ed. Washington, D.C.19996.
38. Schuster JE. Geologic map of Washington state. Olympia, WA. In: Resources WSDoN, ed2005.
39. Naval Facilities Engineering Command HN. Record of Decision for Building 6. JOINT BASE PEARL HARBOR-HICKAM, PEARL HARBOR NAVAL SHIPYARD AND INTERMEDIATE MAINTENANCE FACILITY, OAHU, HAWAII. In: Navy Dot, ed2012.
40. Health HSDo. Statement regarding vanadium in drinking water2014.
41. Water Oo. The Third Unregulated Contaminant Monitoring Rule (UCMR 3). In: Agency USEP, ed2016.
42. 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.
43. Nelson ST, Robinson S, Rey K, et al. Exposure Pathways of Nontuberculous Mycobacteria Through Soil, Streams, and Groundwater, Hawai'i, USA. *Geohealth.* Apr 2021;5(4):e2020GH000350.
44. Cornu S, Lucas, Y., Lebon, E., Ambrosi, J.P., Luizão, F., Rouiller, J., Bonnay, M. and Neal, C. Evidence of titanium mobility in soil profiles, Manaus, central Amazonia. *Geoderma.* 1999;91(3-4):281-295.
45. Kierczak J, Neel, C., Bril, H., Puziewicz, J. Effect of mineralogy and pedoclimatic variations on Ni and Cr distribution in serpentine soils under temperate climate. *Geoderma.* 2007;91(3-4):281-295.
46. 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.
47. Karner J, Papike, J.J. and Shearer, C.K. Comparative planetary mineralogy: Pyroxene major-and minor-element chemistry and partitioning of vanadium between pyroxene and melt in planetary basalts. *American Mineralogist.* 2006;91(10):1574-1582.
48. Glickman CM, Virdi R, Hasan NA, et al. Assessment of soil features on the growth of environmental nontuberculous mycobacterial isolates from Hawai'i. *Appl Environ Microbiol.* Aug 28 2020.
49. Honda JR, Dawrs, S.N., Virdi, R., Norton, G.J., Elias, T., Epperson, E., Hasan, N.A., Glickman, C.M. Crooks, J.L., Chan, E.D., Strong, M., Damby, D.E. Nontuberculous Mycobacteria and Volcanic Ash from the Kilauea Volcano, Hawai'i. *American Thoracic Society International Conference*. Dallas, TX2019.

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

51. Nicholas DJ, Nason A. Molybdenum and nitrate reductase. II. Molybdenum as a constituent of nitrate reductase. *J Biol Chem.* Mar 1954;207(1):353-360.

52. Darnajoux R, Magain N, Renaudin M, Lutzoni F, Bellenger JP, Zhang X. Molybdenum threshold for ecosystem scale alternative vanadium nitrogenase activity in boreal forests. *Proc Natl Acad Sci U S A.* Dec 3 2019;116(49):24682-24688.

53. Kabay N, Güler, E., Bryjak, M. Boron in seawater and methods for its separation—a review. *Desalination.* 2010;261(3):212-217.

54. Rotzoll KaE-K, A.I. . *Numerical ground-water flow simulation for Red Hill Fuel Storage Facilities.* prepared for TEC Inc., Oahu, Hawaii.2007.

Figure Legend.

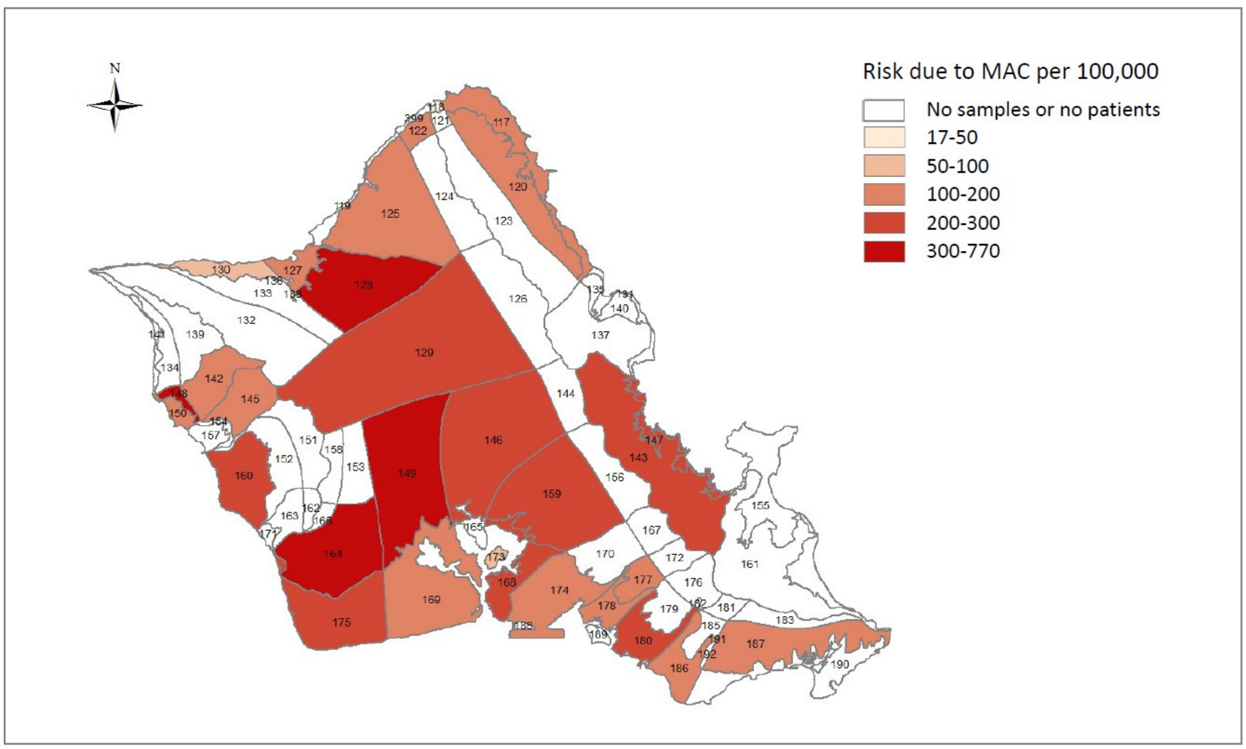


Figure 1. MAC infection risk estimates for aquifers where MAC patients resided based on a vanadium regression model (Model 2; Table 3). Grey lines represent aquifer boundaries in O'ahu, HI.

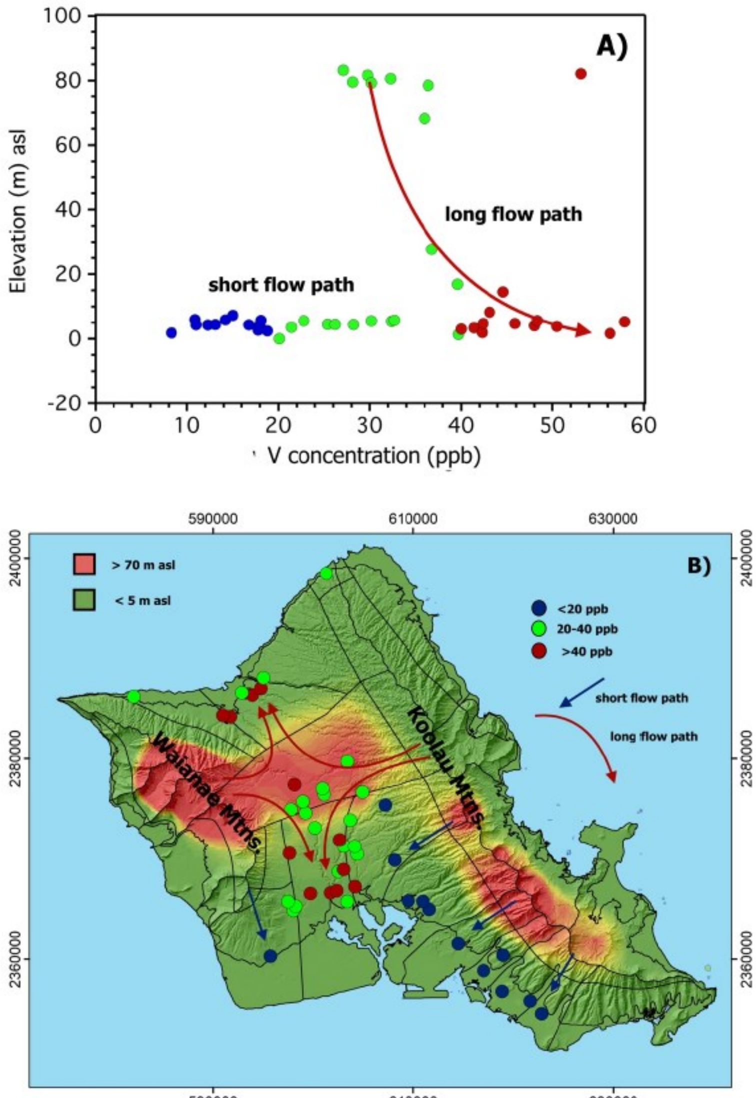


Figure 2. A) Graph of modeled water table elevation above sea level (asl). B) Hillshade map of O'ahu with the modeled elevation of the water table (control points from Rotzoll and El Kadi, 2007 [54]) superimposed. Circles indicate the concentration of vanadium in parts per billion (ppb). Also included are aquifer boundaries as shown in Figure 1. Waters with short flow paths down the flanks of the Koolau and Waianae Mountains to sampling wells tend to have lower vanadium concentrations due to shorter residence times. The low apparent water table elevations beneath the northern Koolau and southern Waianae Mountains are due to a lack of control points for these areas in the model. UTM coordinates, NAD83, Zone 4.

Supplementary Information

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** **(5.9×10^{-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.

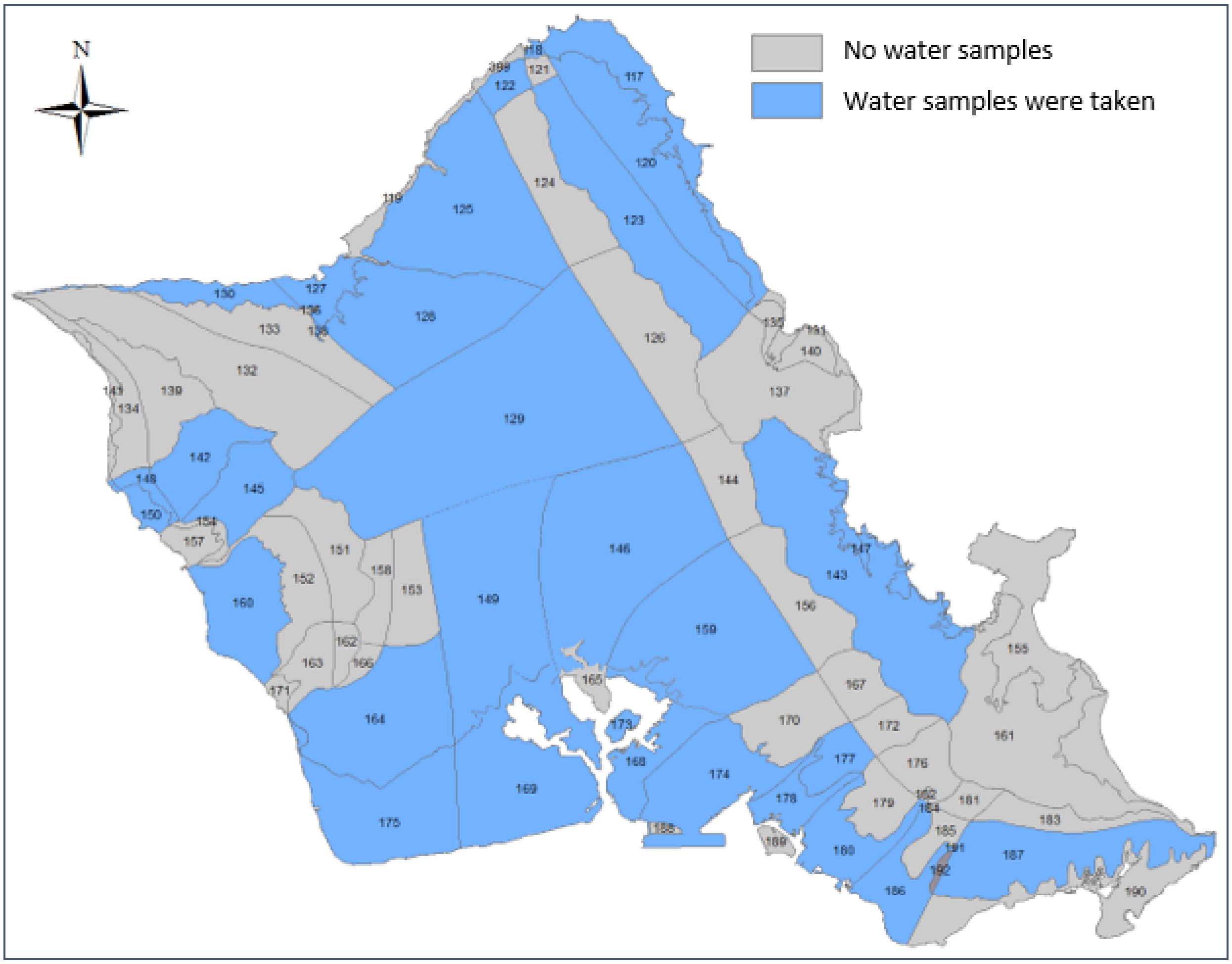
Post hoc Poisson regression models examining water-quality constituents (significant in the Spatial VS methods) associated with NTM infection risk in Oahu, HI.

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.

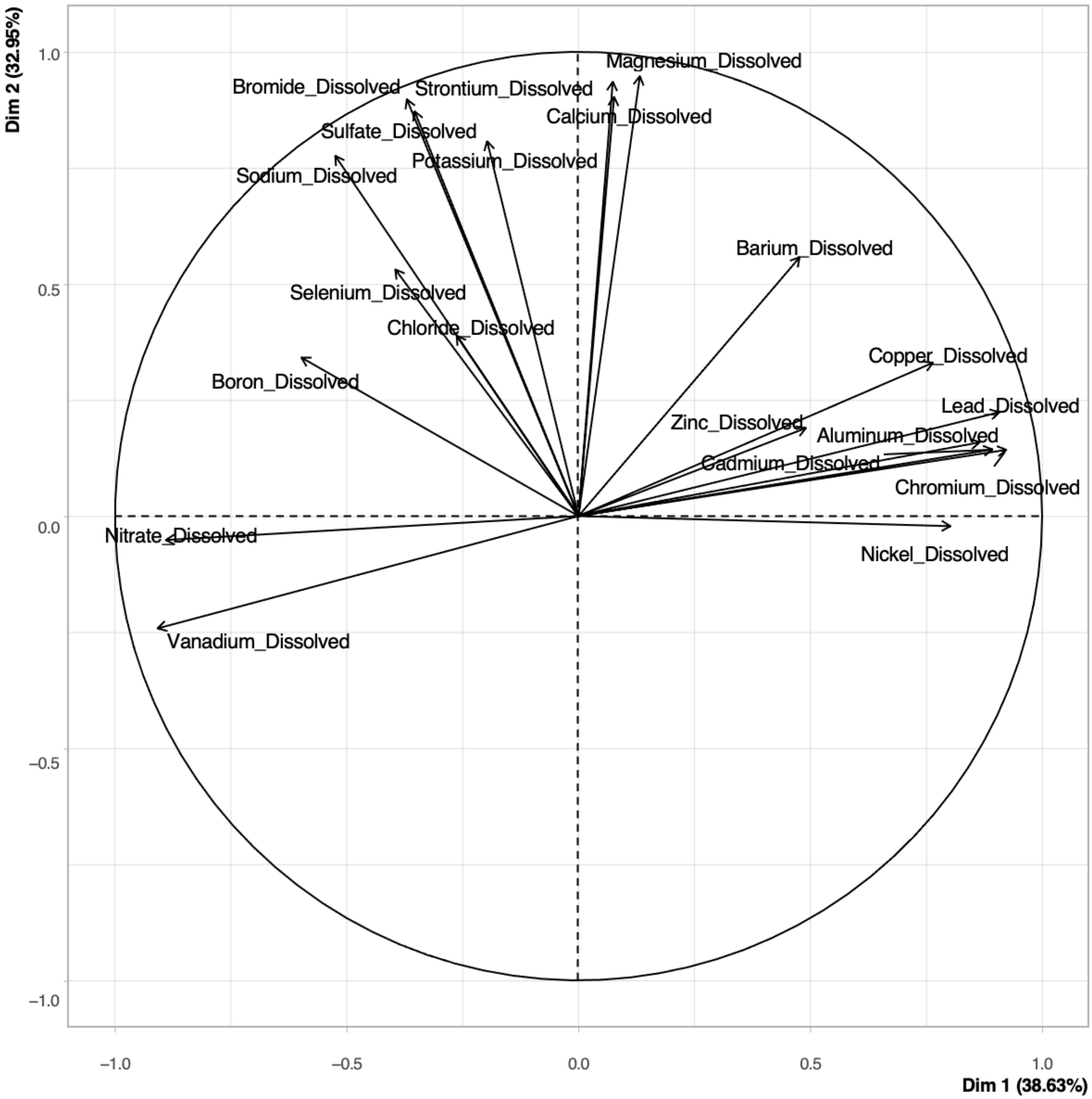
Supplementary Figure 1. Aquifers* with and without water samples in Oahu, HI.



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

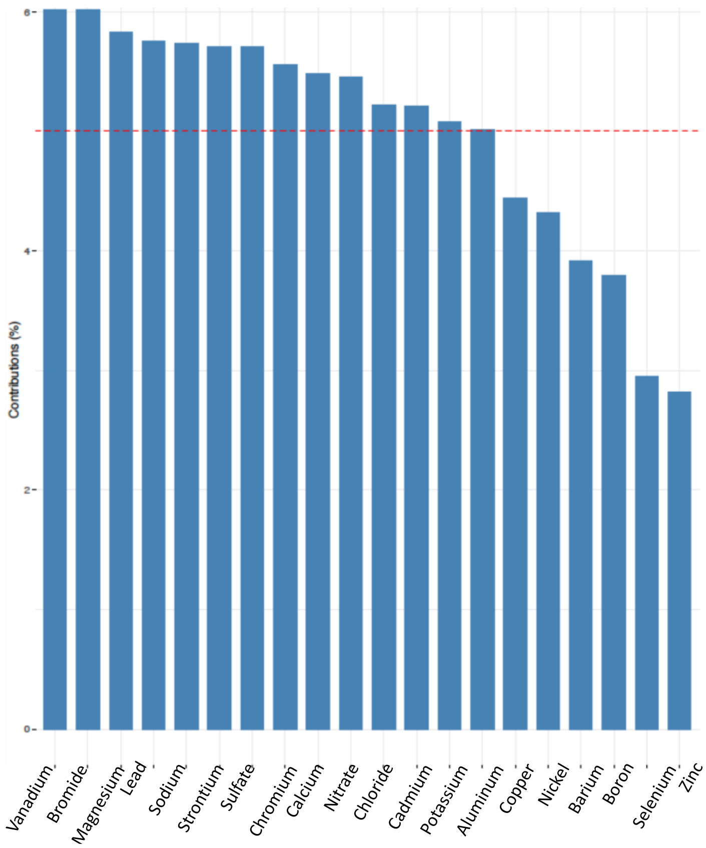
Supplementary Figure 2.

Principal Components Analysis (PCA) graph of variables



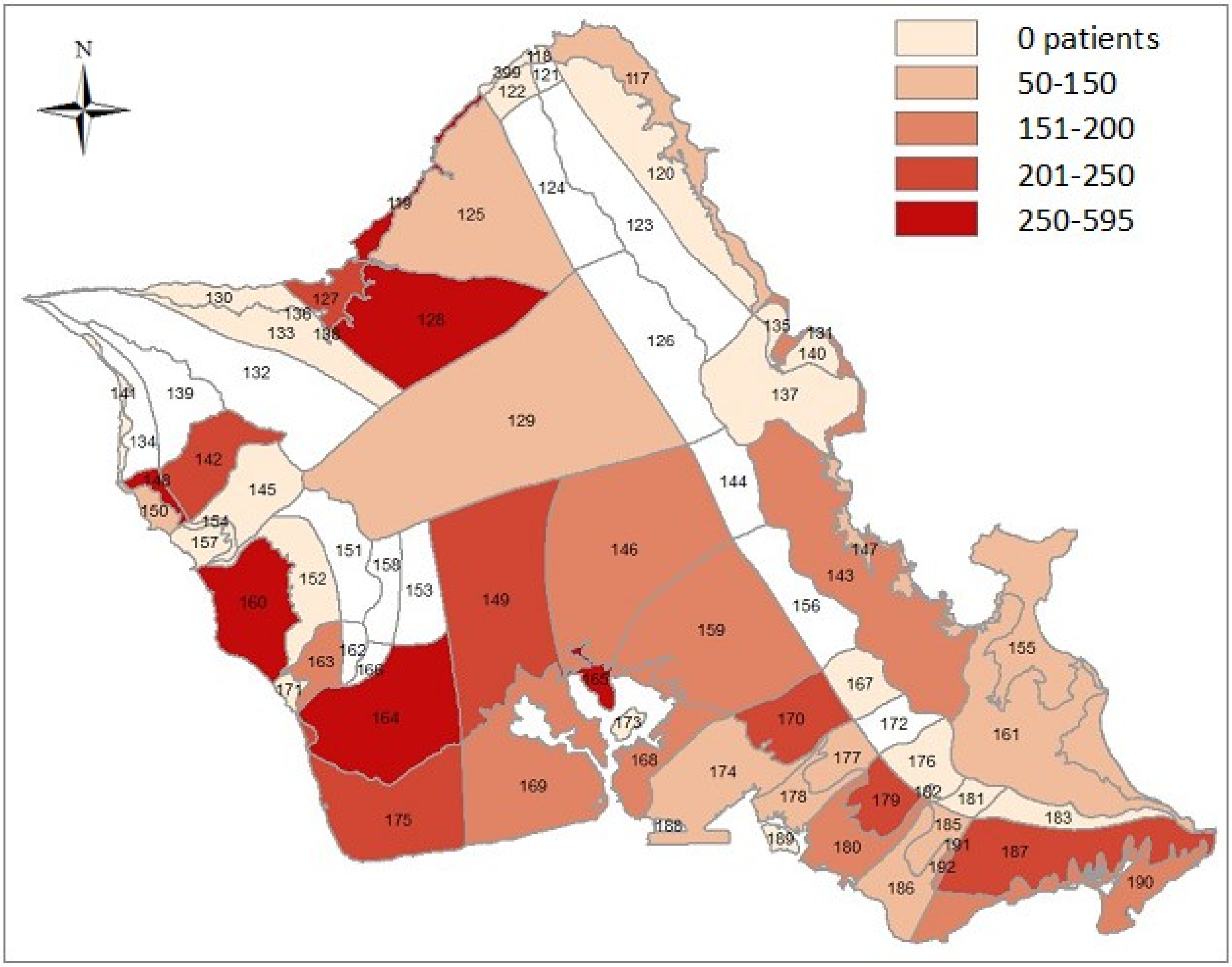
Supplementary Figure 3.

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



Supplementary Figure 4.

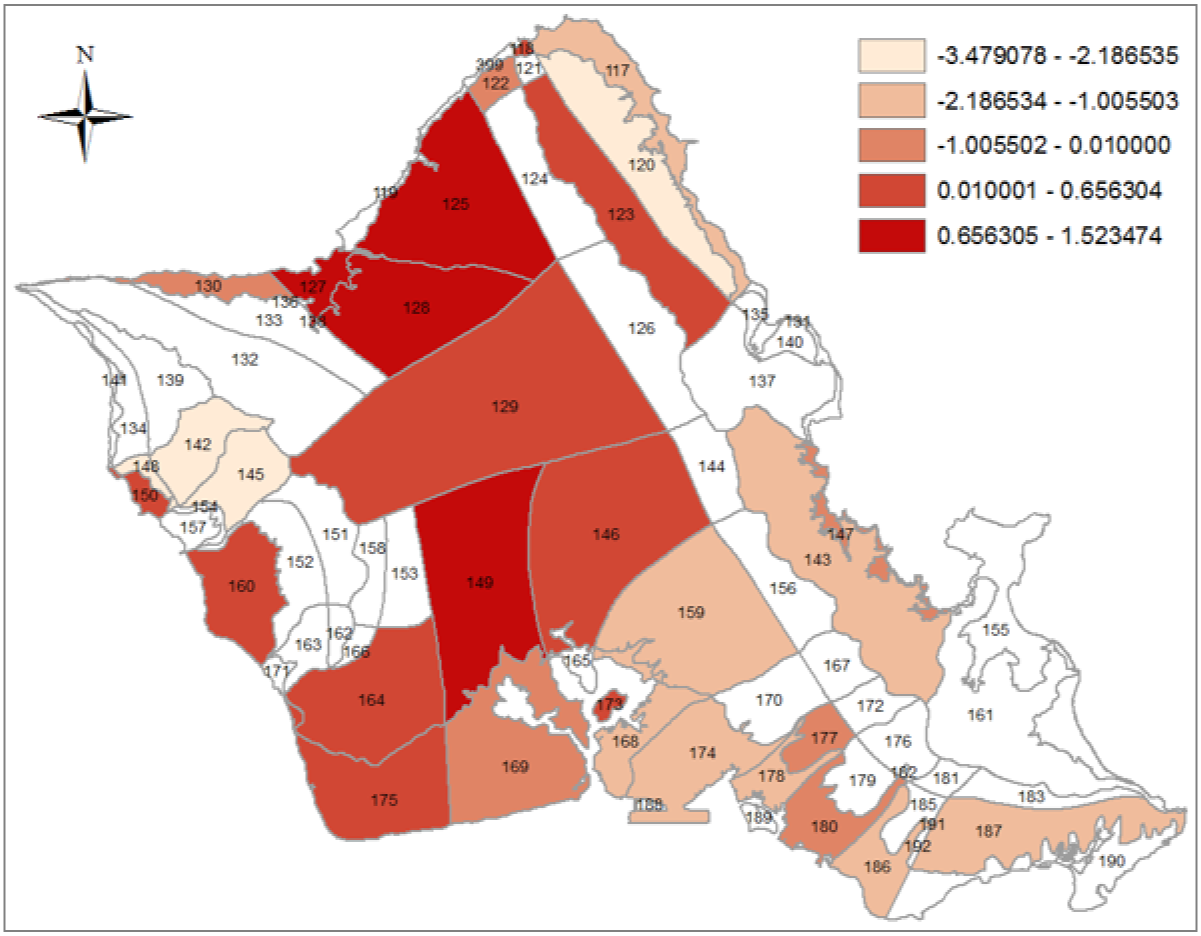
MAC incidence/100,000 per aquifer* in Oahu, HI.



*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/).