



Published in final edited form as:

Landsc Urban Plan. 2021 May ; 209: . doi:10.1016/j.landurbplan.2021.104060.

Proximity to freshwater blue space and type 2 diabetes onset: the importance of historical and economic context

Melissa N. POULSEN¹, Brian S. SCHWARTZ^{1,2,3,4}, Joseph DEWALLE¹, Cara NORDBERG¹, Jonathan S. POLLAK², Jennifer SILVA⁵, Carla I. MERCADO⁶, Deborah B. ROLKA⁶, Karen Rae SIEGEL⁶, Annemarie G. HIRSCH¹

¹Department of Population Health Sciences, Geisinger, Danville, PA

²Department of Environmental Health and Engineering, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

³Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

⁴Department of Medicine, Johns Hopkins School of Medicine, Baltimore, MD

⁵Paul H. O'Neill School of Public and Environmental Affairs, Indiana University, Bloomington, IN

⁶Division of Diabetes Translation, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, GA

Abstract

Salutogenic effects of living near aquatic areas (blue space) remain underexplored, particularly in non-coastal and non-urban areas. We evaluated associations of residential proximity to inland freshwater blue space with new onset type 2 diabetes (T2D) in central and northeast Pennsylvania, USA, using medical records to conduct a nested case-control study. T2D cases (n=15,888) were identified from diabetes diagnoses, medication orders, and laboratory test results and frequency-matched on age, sex, and encounter year to diabetes-free controls (n=79,435). We calculated distance from individual residences to the nearest lake, river, tributary, or large stream, and residence within the 100-year floodplain. Logistic regression models adjusted for community socioeconomic deprivation and other confounding variables and stratified by community type (townships [rural/suburban], boroughs [small towns], city census tracts). Compared to individuals living 1.25 miles from blue space, those within 0.25 miles had 8% and 17% higher odds of T2D onset in townships and boroughs, respectively. Among city residents, T2D odds were 38-39% higher for those living 0.25 to <0.75 miles from blue space. Residing within the floodplain was associated with 16% and 14% higher T2D odds in townships and boroughs. A post-hoc analysis demonstrated patterns of lower residential property values with nearer distance to the region's predominant waterbody, suggesting unmeasured confounding by socioeconomic disadvantage. This may explain our unexpected findings of higher T2D odds with closer proximity to blue space. Our findings highlight the importance of historic and economic context and interrelated factors such as flood risk and lack of waterfront development in blue space research.

1. INTRODUCTION

Exposure to the natural environment is hypothesized to benefit health via three pathways: 1) reduction of harm such as heat, noise, and air pollution; 2) building capacity for health by promoting health-behaviors such as physical activity; and 3) restoration through stress reduction (Markevych et al., 2017). While extensive epidemiologic research demonstrates positive associations of exposure to greenspace (parks and other natural areas in a landscape) and greenness (green vegetation) with health outcomes (Fong, Hart, & James, 2018; Twohig-Bennett & Jones, 2018), the potential salutogenic effects of blue space (aquatic environments such as coasts, lakes, and rivers), remain comparatively underexplored (Gascon, Zijlema, Vert, White, & Nieuwenhuijsen, 2017; M. P. White, Elliott, Gascon, Roberts, & Fleming, 2020). Understanding the unique influence of blue space exposure on health could bolster health-promoting community development strategies.

Researchers increasingly recognize blue space as an important component of the therapeutic landscape (Völker & Kistemann, 2013) and that improving access to such spaces could play a role in preventing poor health outcomes (M. P. White et al., 2020). Since the 1980s, a trend of waterfront redevelopment in urban planning has reclaimed blue space for public access and enjoyment, and visits to such spaces evoke health-promoting experiences, including restoration, stress and anxiety reduction, physical activity and recreation (including land-based activities that occur near water, such as beach walks), and social interaction (de Bell, Graham, Jarvis, & White, 2017; Thomas, 2015; Völker & Kistemann, 2013; Völker, Matros, & Claßen, 2016; M. White et al., 2010). A 2017 review identified 35 epidemiologic studies on blue space and health, with most studies measuring proximity to blue space from individuals' residential areas using spatial data to estimate availability (e.g., percentage of blue space within a defined area) or accessibility (e.g., distance to blue space) (Gascon et al., 2017). The strength of evidence of associations for blue space varied by health outcome, with "limited" evidence supporting an association with mental health or physical activity and "inadequate" evidence for associations with obesity, cardiovascular disease, or diabetes. Additional evidence has since emerged, including a large population-based Canadian cohort study that found protective associations between living near waterbodies and non-accidental causes of death, including cardiovascular disease, diabetes, stroke, and respiratory-related causes, with no evidence of confounding by socioeconomic status, greenness, or air pollution (Crouse et al., 2018). The nascent body of epidemiologic research on blue space and health remains limited by its predominant focus on coastal exposure; a geographic concentration in Europe and urban areas; and heterogeneity in blue space exposure metrics, which limits comparability across studies (Gascon et al., 2017; Völker et al., 2018). Few health studies have evaluated inland, freshwater blue space or blue space across diverse community types, gaps we aimed to address with the current study.

As part of an epidemiological study on determinants of geographic disparities in type 2 diabetes (T2D), we evaluated associations of residential proximity to inland freshwater blue space and new onset T2D within a geographically diverse range of communities in central and northeast Pennsylvania. Two prior studies analyzed prevalent diabetes as a covariate in evaluations of residence in coastal versus non-coastal regions and other health outcomes (Bergovec, Reiner, Milicic, & Vrazic, 2008; Modesti et al., 2013), but to our knowledge

no studies have evaluated blue space and T2D onset. Diabetes is among the leading causes of mortality in the U.S., and it increases risk for other serious health conditions, including coronary heart disease, stroke, hypertension, and kidney disease (Centers for Disease Control and Prevention, 2020). Risk factors implicated in the development of T2D include physical inactivity and psychological stress (Centers for Disease Control and Prevention, 2020; Hackett & Steptoe, 2017), components of two of the primary pathways linking the natural environment and health (Markevych et al., 2017). Though limited, the strongest evidence for health-promoting effects of blue space relates to greater physical activity and improved mental health via restoration and stress reduction (Gascon et al., 2017). We therefore hypothesized that closer residential proximity to blue space would be associated with lower odds of T2D onset. Given the impact of flooding in the region and the potential implications of flooding on health (Fernandez et al., 2015; Tapsell & Tunstall, 2008), we also evaluated residence within the 100-year floodplain in association with T2D onset. Herein, we describe the findings that emerged from this analysis—which led us to explore residential property value data to help interpret the findings—as well as implications for the development of inland waterfront areas and future research on blue space.

2. METHODS

This study was conducted by Geisinger-Johns Hopkins, one of four academic research centers in the Diabetes LEAD (Location, Environmental Attributes, and Disparities) Network (<http://diabetesleadnetwork.org>), a collaboration funded by the Centers for Disease Control and Prevention dedicated to providing scientific evidence to develop targeted community-based interventions and policies to prevent incident T2D and related health outcomes across the United States (Hirsch et al., 2020). The Geisinger Institutional Review Board approved the study and waived informed consent.

2.1. Study Setting

The study region included 37 counties in central and northeast Pennsylvania comprising the primary service area of the Geisinger health system (Figure 1). The region includes a wide geographical range of communities, including cities, suburbs, small towns, and rural areas. In comparison with Pennsylvania and national averages, the study region's population is older and more socially isolated, has lower median household income, and has a higher proportion of blue-collar workers (Baker Tilly, 2018). Much of the region has a higher rate of obesity than the state or nation, and many counties have a higher prevalence of diabetes than Pennsylvania as a whole (Baker Tilly, 2018).

This study focuses solely on inland freshwater blue space. Pennsylvania has over 85,000 miles of rivers and streams and almost 100,000 acres of publicly owned lakes. The Susquehanna River, the longest commercially non-navigable river in North America, comprises the largest watershed in the state. One of the most flood prone areas in the U.S., the Susquehanna River Basin experiences a major flood on average every 15 years (Susquehanna River Basin Commission, 2020), which may increase with the changing climate (U.S. Environmental Protection Agency, 2016).

2.2. Study Population and Design

We obtained electronic health record data from the 1.6 million patients who had a medical encounter at Geisinger from January 2001 through December 2016. Geisinger's primary care population is representative of the region's general population in terms of age and sex and has high residential stability (Casey, 2016). We conducted a nested case-control study to evaluate associations of residential proximity to blue space and new onset T2D, linking individual health data via geocoded residential addresses, obtained through the electronic health record, to geospatial hydrography data.

Individuals were assigned to one of 1070 communities. Community type was defined using a sociologically-representative definition validated through prior work in the region, which combines minor civil divisions (townships, boroughs, cities) with city census tracts (M. Poulsen et al., 2019; Schwartz et al., 2011). In general, townships comprise rural and suburban areas, boroughs are walkable towns, and city census tracts represent more densely populated urban areas.

2.3. Case Ascertainment and Control Selection

We identified cases of new onset T2D that occurred between 2008 to 2016 using diagnostic codes, diabetes medication orders, and relevant laboratory test results, as previously described (M. N. Poulsen et al., 2021; Schwartz et al., 2021). We excluded prevalent cases of T2D by requiring at least one encounter with the health system two years prior to meeting the case definition and by ensuring cases did not meet T2D criteria during that two-year period. Controls were frequency-matched (5:1) to cases by age category, sex, and year of diagnosis/medical encounter. Controls did not meet any of the diabetes criteria up to the randomly selected encounter date in the year of matching to cases. Controls similarly had at least two years of contact with the health system prior to the selected encounter date. Finally, to ensure the algorithm would capture diabetes diagnosis if diabetes was present, we required at least two visits to a Geisinger primary care provider (e.g., family medicine) on different days prior to the diagnosis/encounter date.

2.4. Residential Proximity to Blue Space

Considering the lack of a common approach to blue space measurement (Gascon et al., 2017), we developed and evaluated several availability and accessibility measures (Labib, Lindley, & Huck, 2020) using the 2004 U.S. Geological Survey National Hydrography Dataset (NHD) to define waterbodies (Appendix 1). For this analysis, we retained *distance to water polygon*, defined as Euclidian distance in miles from an individual's residence to the closest polygon water feature, which included lakes, rivers, tributaries, and large streams. We selected this measure for the analysis in large part because water polygons captured prominent landscape features and waterbodies used for recreation—in contrast with linear water features that included small, less noticeable blue space. We established criteria to increase the likelihood that the water polygons included in the blue space measure were meaningful to the general public. For example, we did not include farm ponds or industrial reservoirs, as such features are unlikely to impact population health through the hypothesized mechanisms described previously. From the NHD waterbodies feature class, we included all named lakes as well as unnamed lakes larger than five acres

that were situated within or adjacent to recreational areas owned by the Pennsylvania Game Commission, Department of Conservation and Natural Resources, or Fish and Boat Commission (n = 996). Reservoirs were identified by feature name or type in the NHD and excluded unless situated within or adjacent to recreational areas.

Using ESRI ArcGIS 10.4 (ESRI Inc., Redlands, CA), we calculated distance from individuals' residences to the nearest polygon feature and feature type (lake, river, tributary, stream). Distance to the nearest water polygon was categorized roughly into quintiles as 0 to < 0.25 miles, 0.25 to < 0.50 miles, 0.50 to < 0.75 miles, 0.75 to < 1.25 miles, and 1.25 miles. We also created an indicator for residence within the 100-year floodplain (area that a 1% annual chance of flooding) according to digital flood insurance rate maps from the U.S. Federal Emergency Management Agency.

2.5. Individual and Community Covariates

We obtained hypothesized confounding variables from the electronic health record, including sex, age at diagnosis/encounter, race, ethnicity, and percent of time using Medical Assistance. Medical Assistance (Medicaid) is Pennsylvania's needs-based insurance and serves as a proxy for low household socioeconomic status (Casey et al., 2017), a risk factor for T2D (Agardh, Allebeck, Hallqvist, Moradi, & Sidorchuk, 2011). Sex, age, and Medical Assistance were also evaluated as potential modifiers of blue space and T2D associations given past evidence of more protective associations among women, older adults, and lower income groups living near waterbodies and non-accidental causes of death (Crouse et al., 2018).

We evaluated two time-varying community features, with values assigned to individuals based on the closest measure prior to the year of T2D onset/encounter: community socioeconomic deprivation (CSD) and greenness, each of which are associated with T2D onset in the study region (Schwartz et al., 2021). CSD is a risk factor for T2D (Bilal, Auchincloss, & Diez-Roux, 2018) and a potential confounder in the association between blue space and T2D. The CSD variable was derived using six sociodemographic indicators from the American Community Survey (2006-2011, 2011-2015) (Manson, Schroeder, Van Riper, & Ruggles, 2019): proportions of the population with less than high school education, unemployed, not in labor force, in poverty, receiving public assistance, and households without a car (Nau et al., 2015). Greenness, which has been associated with lower T2D risk (den Braver et al., 2018) and may be spatially related to blue space, was measured using the normalized difference vegetation index (NDVI). Peak NDVI (16-day composite images from early July) was calculated in 1250-meter (approximately 0.78 miles) by 1250-meter square buffers around each person's residence.

2.6. Statistical Analysis

The primary goal of the analysis was to evaluate residential proximity to blue space (distance to water polygon) in association with T2D onset, controlling for CSD and key individual-level confounding variables. Secondary goals included evaluation of residence within the 100-year floodplain and type of nearest blue space in association with T2D onset. Analyses controlled for confounding variables and accounted for clustering of individuals

within communities. After evaluating all univariate distributions of study variables, we examined bivariate relations among key independent variables using Spearman correlations, t-tests, and chi-square tests. Because distributions of blue space proximity and CSD across community types did not sufficiently overlap, it was necessary to stratify all analyses by community type to avoid the regression extrapolation that can occur when there is insufficient overlap in measures between communities (Oakes, 2004). First, we used logistic regression to evaluate associations of blue space proximity and T2D onset using generalized estimating equations with robust standard errors and an exchangeable correlation structure within administrative community types. We then separately evaluated residence within the 100-year floodplain (yes versus no) and type of nearest blue space (lake, river, tributary versus stream) in relation to T2D onset using the same model structure. All models were adjusted for sex, age in years (linear, quadratic, and cubic terms to allow for non-linearity), race (white versus all other racial groups due to a small sample of racial groups other than white), ethnicity (Hispanic versus non-Hispanic), percent of time using Medical Assistance (< 50% versus ≥ 50%), and CSD (quartiles). We assessed potential confounding of the blue space proximity and T2D onset associations by greenness by adding NDVI to the blue space proximity and T2D models and evaluating whether the main effect estimate changed by at least 10%. We assessed effect modification of distance to blue space and T2D associations separately by age, sex, and Medical Assistance use by including cross-products of each variable and distance to blue space categories. Results are reported as odds ratios (OR) with 95% confidence intervals (CI). Analyses were conducted using Stata version 15.1 (StataCorp LP, College Station, TX).

2.7. Property Value Evaluation

In a *post hoc* evaluation designed to explore socioeconomic patterns in relation to blue space in the study region, we used a geospatial tax parcel database obtained from Lycoming County (2015) to assess residential property values (building value by square footage) by distance to the nearest water polygon. Lycoming County was selected for evaluation because the county contains all three community types located adjacent to the Susquehanna River—the region’s predominant waterbody—and based on data availability. We evaluated mean property values by categories of distance to blue space using Welch ANOVA tests, stratified by community type.

3. RESULTS

3.1. Characteristics of Study Individuals and Communities

We identified 15,888 cases of new onset T2D from 2008 to 2016, matched to 79,435 controls (Table 1). The majority of individuals were white and non-Hispanic, reflecting the region’s racial and ethnic composition. Average use of Medical Assistance for health insurance was higher among T2D onset cases than controls. Cases were more likely to live in a city census tract or borough (versus township), within the 100-year floodplain, closest to a river or tributary (rather than a lake or stream), and in an area with higher CSD. Cases were also more likely to live closer to blue space. Average distance between individuals’ residences and the nearest blue space was 0.85 miles but differed by community type (mean [standard deviation]) (townships: 0.97 [0.78]; boroughs: 0.63 [0.57]; city census tracts: 0.63

[0.43]). The type of nearest blue space also significantly differed by community type; for example, 33% of borough residents and 35% of city residents lived nearest a river, compared to 15% of township residents. Township residents were more likely than residents of other community types to live nearest a lake or stream.

3.2. Relations Between Distance to Blue Space and Confounding Variables

Bivariate analyses revealed the main independent variable—distance to blue space—and several potential confounders were significantly related (all had categorical cross-tabulations with a chi-square p -value < 0.001). Residential distance to blue space was not significantly correlated with CSD (Spearman = -0.11); however, individuals residing closest to blue space were significantly less likely to live in CSD quartile 1 (lowest deprivation) and those residing the farthest from blue space more likely to live in CSD quartile 4 (most deprived). Distance to blue space was also related to the type of nearby blue space: on average, individuals living nearest a river or tributary had a shorter distance to blue space (mean [standard deviation]) (0.61 [0.48]; 0.73 [0.69], respectively) than did individuals living nearest a lake or stream (0.99 [0.80]; 0.97 [0.77], respectively). Those living nearest a river or tributary were also more likely to live in communities with the greatest CSD (32% and 31% were in CSD quartile 4, respectively, as compared to 10% of those living nearest a lake and 21% of those nearest a stream) and within the 100-year floodplain (6% were in CSD quartile 4 for both rivers and tributaries, as compared to 1% for those nearest a lake and 4% for those nearest a stream). Individuals living within the 100-year floodplain were more likely to live in communities with the greatest CSD (32% were in quartile 4 versus 24% of individuals not in the floodplain). Higher greenness was moderately correlated with lower CSD (Spearman = -0.33) but not significantly with distance to blue space (Spearman = -0.06).

3.3. Blue Space and New Onset T2D

In models stratified by community type and adjusted for CSD and demographic covariates, individuals living in townships and boroughs closest to blue space had significantly higher odds of T2D onset (Table 2). Compared to individuals living 1.25 miles away from blue space, those within 0.25 miles had 8% higher odds of T2D onset in townships and 17% higher odds in boroughs. Among city residents, T2D odds were highest between 0.25 to < 0.50 and 0.50 to < 0.75 miles from blue space, with 39% and 38% higher odds, respectively. In models evaluating residence within the 100-year floodplain, living in the floodplain was associated with increased odds of T2D in townships and boroughs, but not cities (Table 2). In models evaluating type of nearest blue space, living nearest a river (versus stream) was associated with higher odds of T2D among individuals in townships, with a similar though non-significant trend for tributaries (Table 2). Among individuals living in cities, living nearest a lake (versus stream) was associated with lower odds of T2D; however, this association was driven by 334 study individuals, with 96% of them living in just three cities. In boroughs, type of nearest blue space was not associated with T2D. In all models in Table 2, race, ethnicity, and Medical Assistance status were each associated with T2D onset in all community types, with ORs ranging from 1.29-1.44 for non-white, 1.33-1.52 for Hispanic, and 1.46-1.83 for Medical Assistance (data not shown). CSD was only associated with T2D

onset in cities, with ORs ranging from 0.74-0.75 in quartile 1, 0.77-0.79 in quartile 2, and 0.79-0.80 in quartile 3 (versus quartile 4).

We found no evidence of confounding by greenness in townships or boroughs (results not shown). There was an insufficient range of greenness (i.e., no communities with high greenness) to evaluate confounding in cities. We found no evidence of effect modification of the distance to blue space and T2D associations by age, sex, or Medical Assistance (results not shown).

3.4. Property Value Evaluation

In Lycoming County, we observed a notable trend of lower residential property values closer to the Susquehanna River, tributaries of the river, and the 100-year floodplain (Figure 2). Mean property values incrementally decreased with nearer distance to water (Welch ANOVA p-values < 0.001), ranging from \$537 per square foot for properties furthest from blue space to \$429 per square foot for properties closest to blue space. This pattern held when evaluated separately by community type, with one exception. In cities, where property values were lower overall, the trend of lower property values with closer distance to water ended at 0.25 miles, with a slight uptick in mean property values with distance < 0.25 miles from blue space (\$304 per square foot with distance < 0.25 miles, compared to \$291 with distance 0.25 to 0.50 miles).

4. DISCUSSION

In this study of residential proximity to inland freshwater blue space and diabetes, we hypothesized that living closer to blue space would be associated with lower risk of T2D onset. Closer proximity to blue space is associated with more frequent visits (Elliott et al., 2020)—providing nature-based opportunities for recommended behavioral approaches for preventing diabetes (e.g., physical activity, stress reduction) (Centers for Disease Control and Prevention, 2020; Hackett & Steptoe, 2017)—and prior studies suggest a beneficial effect of blue space on health (Crouse et al., 2018; Gascon et al., 2017; M. P. White et al., 2020). Contrary to our hypothesis, we observed higher odds of new onset T2D among individuals living closer to blue space in these central and northeast Pennsylvania communities. Among residents of townships and boroughs, living within 0.25 miles of blue space was associated with 8% and 17% higher odds of T2D, respectively, as compared to living more than 1.25 miles away; among city residents, odds were 38%-39% higher between 0.25 to 0.75 miles from blue space. These associations were consistent when adjusting for Medical Assistance status, an indicator of individual-level socioeconomic status, and CSD, an indicator of community-level disadvantage. We postulate that our unexpected finding is not a direct effect of blue space exposure, but rather explained by geographic patterns of socioeconomic disadvantage and a lack of human and economic capital to transform waterfronts from landscapes of risk into salutogenic spaces conducive to physical activity and restoration.

Although analyses controlled for Medical Assistance and CSD, a *post hoc* evaluation of residential property values suggested our findings may be partly explained by residual confounding by individual- or neighborhood-level socioeconomic status, known risk factors

for T2D (Agardh et al., 2011; Bilal et al., 2018). Residential property values from Lycoming County, located in the heart of the Susquehanna River basin, demonstrated a trend of decreasing mean property values with nearer distance to blue space in townships and boroughs, with the lowest mean property values closest to blue space. The exception to this pattern was observed in cities, where the lowest mean property values were 0.25 to 0.50 miles from blue space, similar to the pattern of T2D onset risk. This pattern of lower property values closer to waterbodies such as the Susquehanna River reflects historic and economic development in the region. Many towns in the study region developed along the Susquehanna River due to its historic importance as a transportation corridor—via a canal system—for raw goods through rugged and isolated terrain, and as a driver of industrial development, providing waterpower to sawmills, iron mills, and coal machinery. In addition to devaluation that may have arisen from flooding, railways and major roads that are co-located along the transportation corridor originally established by the Susquehanna and other rivers may further degrade property values, and could also reduce opportunities for physical activity and stress reduction in blue space through physical barriers and the creation of air and noise pollution. Additionally, economic decline in the region following dual waves of de-industrialization in the past half-century has reshaped communities (Keil & Keil, 2015; MacGaffey, 2013) and their relationship with the river, and left many towns lacking the human and economic capital required to develop waterfronts into desirable places to live. With many towns originally settled in close proximity to the river, newer homes were constructed away from rivers—in the outer areas of towns and on historically agricultural land, as evidenced by the loss of farmland to development in the Northeast states during the second half of the twentieth century (Hellerstein et al., 2002). The legacy of these trends appears to have created a socioeconomic gradient that remains evident in the region’s geography today, highlighting the importance of understanding historic and economic context when evaluating the health impacts of blue space.

Individuals with lower incomes may also be more likely to reside in areas subject to flooding due to price structures that constrain their residential options, further exacerbating the geographic patterns of socioeconomic disadvantage described above. The impact of flooding—a relatively common occurrence for the Susquehanna River and its tributaries—may be most detrimental to economically vulnerable populations, who are less resilient in recovering from floods and lack the financial resources to move or modify their property (Tapsell & Tunstall, 2008), particularly as the base flood elevation area designated by the Federal Emergency Management Agency has expanded. With new floodplain studies conducted in the early 2000’s, more properties have become subject to flood insurance and elevation requirements, leading to reduced property values and making properties more difficult to sell. Our findings demonstrated that living within the 100-year floodplain was associated with 16% and 14% higher odds of T2D onset in townships and boroughs, respectively, after adjusting for individual and community socioeconomic factors. Observations in townships also suggested that compared to living nearest a stream or lake, living nearest a river or tributary—the most flood-prone waterbodies—increased the odds of T2D. The lack of association in cities may reflect greater flood mitigation efforts such as the construction of levees in larger metropolitan areas and fewer homes located in the floodplain. These findings provide further evidence that the socioeconomic context,

rather than blue space exposure, may explain elevated T2D odds with closer proximity to the regions rivers and tributaries. That said, the accumulated psychosocial stress of living through floods and anticipating future floods could have detrimental effects on long-term health, including chronic conditions such as T2D (Hackett & Steptoe, 2017). Experiencing flooding has been shown to negatively impacts mental health (Fernandez et al., 2015), with even relatively small events having long-term effects on mental health and well-being (Tapsell & Tunstall, 2008).

Our findings highlight potential health implications of the planning, management, and development of freshwater areas in this region of Pennsylvania and similar regions where historical and economic trends have led to geospatial disparities in wealth and health. First, investments in sustainable flood prevention and mitigation and assisting low-income households to move out of flood-prone areas would likely have positive impacts on community health. This will be increasingly important if climate change results in more frequent major flood events. Second, purposeful waterfront development enhances the therapeutic value of blue space, facilitating experiences of restoration, contemplation, and a sense of place (Völker & Kistemann, 2013), and can provide areas for physical activity. Formative research we conducted in 2018 in our study region among key informants and recreational users of Pennsylvania's rivers suggested that the Susquehanna River may not be perceived as an asset to some surrounding communities, particularly those that have not developed their riverfront. Interviewees reported that despite its scenic beauty and many recreational uses, the river remains "invisible" as a destination to many people, with recreation hampered by insufficient access (e.g., the major road depicted in Figure 2 was specifically noted to be a barrier to visiting the river), and that some communities located along the river perceive it as a liability due to flooding risk, likely remnants of destructive floods that occurred in 1972 and 2011 (publication in process). Interviewees also described perceptions of the river as "destructive and dirty," which reflects its past and present pollution. These findings parallel research in the region that has highlighted the population's concern about health threats from contamination of streams and soil by large companies (Silva, 2019). Many towns along the river have not developed their waterfront to encourage recreation, which, according to interviewees, is due to a lack of resources and, in some cases, political will. Without such development, this region of Pennsylvania may not see health-promoting benefits of its blue space.

Our study also has implications for epidemiologic research on blue space and health more broadly. To avoid spurious conclusions, such research must consider the historic and economic context that has shaped the development of communities, and researchers should exercise caution when making claims about the generalizability of study findings. As in Pennsylvania, many communities were established along water to facilitate transportation and industrial processes, and while the past several decades have brought significant investment in waterfront redevelopment, particularly in large cities (Völker & Kistemann, 2013), such redevelopment is not universal. Not considering such context may lead to inappropriate aggregation of blue space across large areas. While we found blue space to be negatively associated with our health outcome, the reverse is also plausible in other regions; for example, historic development of coastal areas that advantage the wealthy may lead to positive associations with health. In fact, homes in many countries that are closer to blue

space tend to be more expensive (M. P. White et al., 2020). In such cases, adjustment for individual- and community-level socioeconomic status may be insufficient to control for structural confounding by socioeconomic status (Ahern et al., 2013).

Strengths of this study include evaluation of an understudied health outcome in relation to freshwater blue space across multiple community types and evaluation of greenness as a potential confounder, a shortcoming of some prior blue space and health research (Gascon et al., 2017). As is common in studies of nature and health (Hartig, Mitchell, de Vries, & Frumkin, 2014), the primary limitation is that our measure of blue space captured proximity to blue space but did not account for other exposure types (indirect, incidental, or intentional (M. P. White et al., 2020)), or quality of blue space, which likely influence the degree to which blue spaces are utilized for health-promoting experiences such physical activity and stress reduction. We did not have residential history information and so could not account for exposure duration, although our study region experiences low rates of out-migration (Casey et al., 2016). Due to the ubiquity of waterbodies in the region, the majority of the study population lived close to a waterbody and so categories of residential distance to blue space did not differ widely. We therefore may not have detected a difference in T2D onset that could occur if comparisons were made with individuals far from blue space, pointing to a need for research on blue space and health in additional regions with different landscape types. Finally, as discussed, our findings are subject to unmeasured confounding by socioeconomic status. The property value analysis provided support that distance to blue space is related to socioeconomic status in our study region, but with limited availability of tax parcel data and without the ability to link individual health data to property values, we were not able to evaluate property values in relation to T2D onset.

5. CONCLUSION

Contrary to hypothesized health-promoting effects of blue space, we found that higher odds of new onset T2D with closer proximity to freshwater blue space. This potentially spurious association appears to reflect the primary role that the Susquehanna and other rivers have historically played in shaping development in central and northeast Pennsylvania, creating a geographic gradient of socioeconomic disadvantage that emanates from their banks. Our findings indicate that relationships between proximity to blue space and health depend not only on the use of the space for healthy behaviors, but also on historic and economic context, and provide a cautionary tale for epidemiologic studies evaluating the influence of blue space on health. They also highlight the potential broad health benefits that could result from waterfront development and investment in flood-prone communities.

Acknowledgements:

The authors thank Dione G. Mercer for project management.

Funding: This publication was made possible by Cooperative Agreement Number DP006296 funded by the U.S. Centers for Disease Control and Prevention, Division of Diabetes Translation.

Disclaimer: The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Appendix

Appendix 1.

Blue space measures developed and reasons for exclusion from analysis

Blue space measure	Definition	Reason for exclusion
<i>Availability measures</i>		
Percent blue space	Percent of a water polygon ¹ contained in a community ² boundary [(polygon square miles/ community square miles) x 100]	Distribution insufficient for analysis (most communities contained < 1% blue space)
Blue line density	Length of linear water feature ³ within a community boundary (linear miles/community square miles)	Did not meet face validity
<i>Accessibility measures</i>		
Distance to water polygon	Distance in miles from individual's residence to closest water polygon	N/A – used in analysis
Distance to linear water feature	Distance in miles from individual's residence to closest linear water feature	Overly sensitive to obscure water features (e.g., small streams)
Distance to closest water feature	Distance in miles from individual's residence to closest water feature (polygon or linear)	Highly correlated with distance to linear water feature

¹Water polygons included lakes, rivers, tributaries, and large streams, as described in Methods.

²Community types included townships, boroughs, and city census tracts, as described in Methods.

³Linear water features included streams and rivers.

REFERENCES

- Agardh E, Allebeck P, Hallqvist J, Moradi T, & Sidorchuk A (2011). Type 2 diabetes incidence and socio-economic position: a systematic review and meta-analysis. *International Journal of Epidemiology*, 40(3), 804–818. doi:10.1093/ije/dyr029 [PubMed: 21335614]
- Ahern J, Cerdá M, Lippman SA, Tardiff KJ, Vlahov D, & Galea S (2013). Navigating non-positivity in neighbourhood studies: an analysis of collective efficacy and violence. *Journal of epidemiology and community health*, 67(2), 159–165. doi:10.1136/jech-2012-201317 [PubMed: 22918895]
- Tilly Baker. (2018). Geisinger community health needs assessment: July 1,2018 - June 30, 2021. Retrieved from <https://www.geisinger.org/about-geisinger/in-our-community/chna>
- Bergovec M, Reiner Z, Milicic D, & Vrazic H (2008). Differences in risk factors for coronary heart disease in patients from continental and Mediterranean regions of Croatia. *Wien Klin Wochenschr*, 120, 684–692. doi:10.1007/s00508-008-1065-7 [PubMed: 19116710]
- Bilal U, Auchincloss AH, & Diez-Roux AV (2018). Neighborhood Environments and Diabetes Risk and Control. *Current diabetes reports*, 18(9), 62–62. doi:10.1007/s11892-018-1032-2 [PubMed: 29995252]
- Casey JA, Pollak J, Glymour MM, Mayeda ER, Hirsch AG, & Schwartz BS (2017). Measures of SES for Electronic Health Record-based Research. *Am J Prev Med*. doi:10.1016/j.amepre.2017.10.004
- Casey JA, Savitz DA, Rasmussen SG, Ogburn EL, Pollak J, Mercer DG, & Schwartz BS (2016). Unconventional Natural Gas Development and Birth Outcomes in Pennsylvania, USA. *Epidemiology*, 27(2), 163–172. doi:10.1097/EDE.0000000000000387 [PubMed: 26426945]
- Centers for Disease Control and Prevention. (2020). National Diabetes Statistics Report, 2020. Retrieved from Atlanta, GA:
- Crouse DL, Balram A, Hystad P, Pinault L, van den Bosch M, Chen H, ... Villeneuve PJ (2018). Associations between Living Near Water and Risk of Mortality among Urban Canadians. *Environ Health Perspect*, 126(7), 077008. doi:10.1289/ehp3397 [PubMed: 30044232]

- de Bell S, Graham H, Jarvis S, & White P (2017). The importance of nature in mediating social and psychological benefits associated with visits to freshwater blue space. *Landscape and Urban Planning*, 167, 118–127. doi:10.1016/j.landurbplan.2017.06.003
- den Braver NR, Lakerveld J, Rutters F, Schoonmade LJ, Brug J, & Beulens JWJ (2018). Built environmental characteristics and diabetes: a systematic review and meta-analysis. *BMC medicine*, 16(1), 12–12. doi:10.1186/s12916-017-0997-z [PubMed: 29382337]
- Elliott LR, White MP, Grellier J, Garrett JK, Cirach M, Wheeler BW, ... Fleming LE (2020). Research Note: Residential distance and recreational visits to coastal and inland blue spaces in eighteen countries. *Landscape and Urban Planning*, 198, 103800. doi:10.1016/j.landurbplan.2020.103800
- Fernandez A, Black J, Jones M, Wilson L, Salvador-Carulla L, Astell-Burt T, & Black D (2015). Flooding and mental health: a systematic mapping review. *PloS one*, 10(4), e0119929–e0119929. doi:10.1371/journal.pone.0119929 [PubMed: 25860572]
- Fong KC, Hart JE, & James P (2018). A Review of Epidemiologic Studies on Greenness and Health: Updated Literature Through 2017. *Current environmental health reports*, 5(1), 77–87. doi:10.1007/s40572-018-0179-y [PubMed: 29392643]
- Gascon M, Zijlema W, Vert C, White MP, & Nieuwenhuijsen MJ (2017). Outdoor blue spaces, human health and well-being: A systematic review of quantitative studies. *Int J Hyg Environ Health*, 220(8), 1207–1221. doi:10.1016/j.ijheh.2017.08.004 [PubMed: 28843736]
- Hackett RA, & Steptoe A (2017). Type 2 diabetes mellitus and psychological stress - a modifiable risk factor. *Nature reviews. Endocrinology*, 13(9), 547–560. doi:10.1038/nrendo.2017.64
- Hartig T, Mitchell R, de Vries S, & Frumkin H (2014). Nature and health. *Annual review of public health*, 35, 207–228. doi:10.1146/annurev-publhealth-032013-182443
- Hellerstein D, Nickerson C, Cooper J, Feather P, Gadsby D, Mullarkey D, ... Barnard C (2002). Farmland protection: the role of public preferences for rural amenities. Retrieved from
- Hirsch AG, Carson AP, Lee NL, McAlexander T, Mercado C, Siegel K, ... Thorpe LE (2020). The Diabetes Location, Environmental Attributes, and Disparities Network: Protocol for Nested Case Control and Cohort Studies, Rationale, and Baseline Characteristics. *JMIR Res Protoc*, 9(10), e21377. doi:10.2196/21377 [PubMed: 33074163]
- Keil TJ, & Keil JM (2015). *Anthracite's demise and the post-coal economy of northeastern Pennsylvania*. Bethlehem, PA: Lehigh University Press.
- Labib SM, Lindley S, & Huck JJ (2020). Spatial dimensions of the influence of urban green-blue spaces on human health: A systematic review. *Environmental research*, 180, 108869. doi:10.1016/j.envres.2019.108869 [PubMed: 31722804]
- MacGaffey J (2013). *Coal dust on your feet: the rise, decline, and restoration of an anthracite mining town*. Lewisburg, PA: Bucknell University Press.
- Manson S, Schroeder J, Van Riper D, & Ruggles S (2019). IPUMS National Historical Geographic Information System.
- Markevych I, Schoierer J, Hartig T, Chudnovsky A, Hystad P, Dzhambov AM, ... Fuertes E (2017). Exploring pathways linking greenspace to health: Theoretical and methodological guidance. *Environmental Research*, 158, 301–317. doi:10.1016/j.envres.2017.06.028 [PubMed: 28672128]
- Modesti PA, Bamoshmoosh M, Rapi S, Massetti L, Al-Hidabi D, & Goshae HA (2013). Epidemiology of hypertension in Yemen: effects of urbanization and geographical area. *Hypertension Research*, 36, 711–717. doi:10.1038/hr.2013.14 [PubMed: 23486167]
- Nau C, Schwartz BS, Bandeen-Roche K, Liu A, Pollak J, Hirsch A, ... Glass TA (2015). Community socioeconomic deprivation and obesity trajectories in children using electronic health records. *Obesity (Silver Spring)*, 23(1), 207–212. doi:10.1002/oby.20903 [PubMed: 25324223]
- Oakes JM (2004). The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Social Science & Medicine*, 58(10), 1929–1952. doi:10.1016/j.socscimed.2003.08.004 [PubMed: 15020009]
- Poulsen M, Glass T, Pollak J, Bandeen-Roche K, Hirsch A, Bailey-Davis L, & Schwartz B (2019). Associations of multidimensional socioeconomic and built environment factors with body mass index trajectories among youth in geographically heterogeneous communities. *Preventive Medicine Reports*, 15, 100939. doi:10.1016/j.pmedr.2019.100939 [PubMed: 31360629]

- Poulsen MN, Schwartz BS, Nordberg C, DeWalle J, Pollak J, Imperatore G, ... Hirsch AG (2021). Association of Greenness with Blood Pressure among Individuals with Type 2 Diabetes across Rural to Urban Community Types in Pennsylvania, USA. *Int J Environ Res Public Health*, 18(2). doi:10.3390/ijerph18020614
- Schwartz BS, Pollak J, Poulsen MN, Bandeen-Roche K, Moon K, DeWalle J, ... Hirsch AG (2021). Association of community types and features in a case-control analysis of new onset type 2 diabetes across a diverse geography in Pennsylvania. *BMJ Open*, 11(1), e043528. doi:10.1136/bmjopen-2020-043528
- Schwartz BS, Stewart WF, Godby S, Pollak J, Dewalle J, Larson S, ... Glass TA (2011). Body mass index and the built and social environments in children and adolescents using electronic health records. *Am J Prev Med*, 41(4), e17–28. doi:10.1016/j.amepre.2011.06.038 [PubMed: 21961475]
- Silva J (2019). *We're still here: pain and politics in the heart of america*: Oxford University Press.
- Susquehanna River Basin Commission. (2020). History of flooding. Retrieved from <https://www.srbc.net/our-work/programs/planning-operations/flooding.html>
- Tapsell SM, & Tunstall SM (2008). "I wish I'd never heard of Banbury": The relationship between 'place' and the health impacts from flooding. *Health & Place*, 14(2), 133–154. doi:10.1016/j.healthplace.2007.05.006 [PubMed: 17616427]
- Thomas F (2015). The role of natural environments within women's everyday health and wellbeing in Copenhagen, Denmark. *Health & Place*, 35, 187–195. doi:10.1016/j.healthplace.2014.11.005 [PubMed: 25435057]
- Twohig-Bennett C, & Jones A (2018). The health benefits of the great outdoors: A systematic review and meta-analysis of greenspace exposure and health outcomes. *Environmental research*, 166, 628–637. doi:10.1016/j.envres.2018.06.030 [PubMed: 29982151]
- U.S. Environmental Protection Agency. (2016). What climate change means for Pennsylvania. Retrieved from <https://19january2017snapshot.epa.gov/sites/production/files/2016-09/documents/climate-change-pa.pdf>
- Völker S, Heiler A, Pollmann T, Claßen T, Hornberg C, & Kistemann T (2018). Do perceived walking distance to and use of urban blue spaces affect self-reported physical and mental health? *Urban Forestry & Urban Greening*, 29, 1–9. doi:10.1016/j.ufug.2017.10.014
- Völker S, & Kistemann T (2013). "I'm always entirely happy when I'm here!" Urban blue enhancing human health and well-being in Cologne and Düsseldorf, Germany. *Social science & medicine* (1982), 78, 113–124. doi:10.1016/j.socscimed.2012.09.047 [PubMed: 23273410]
- Völker S, Matros J, & Claßen T (2016). Determining urban open spaces for health-related appropriations: a qualitative analysis on the significance of blue space. *Environmental Earth Sciences*, 75(13), 1067. doi:10.1007/s12665-016-5839-3
- White M, Smith A, Humphries K, Pahl S, Snelling D, & Depledge M (2010). Blue space: The importance of water for preference, affect, and restorativeness ratings of natural and built scenes. *Journal of Environmental Psychology*, 30(4), 482–493. doi:10.1016/j.jenvp.2010.04.004
- White MP, Elliott LR, Gascon M, Roberts B, & Fleming LE (2020). Blue space, health and well-being: A narrative overview and synthesis of potential benefits. *Environmental Research*, 191, 110169. doi:10.1016/j.envres.2020.110169 [PubMed: 32971082]

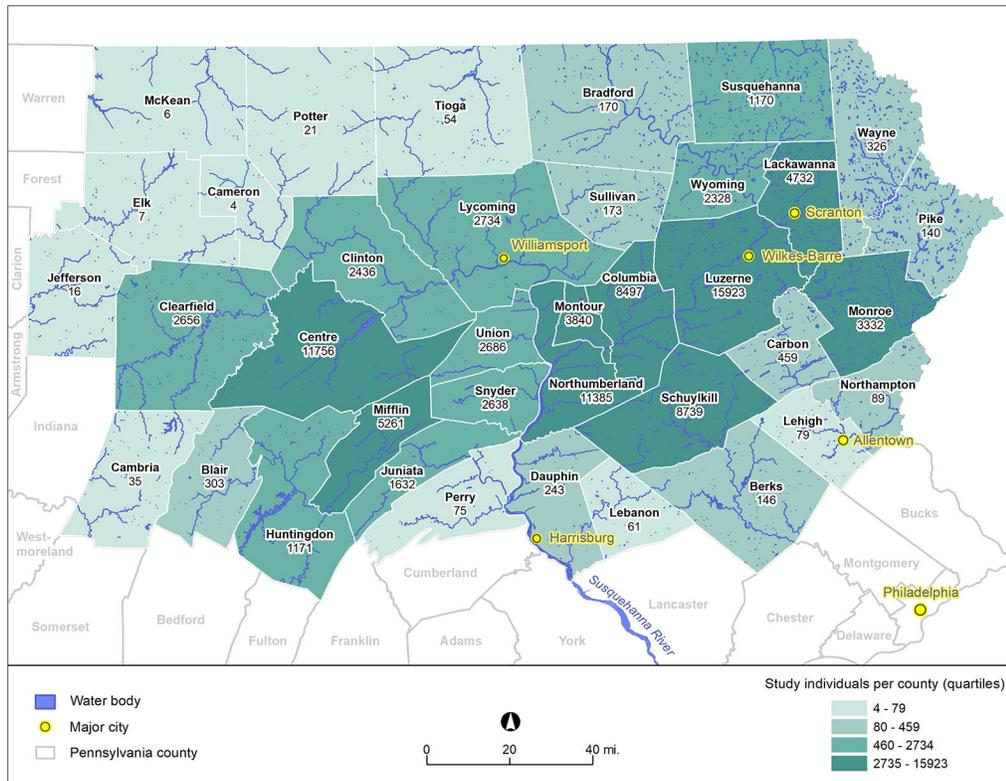


Figure 1. Distribution of individuals by county and water bodies in the study region in Pennsylvania, USA. Numbers within each county indicate the number of study individuals per county.

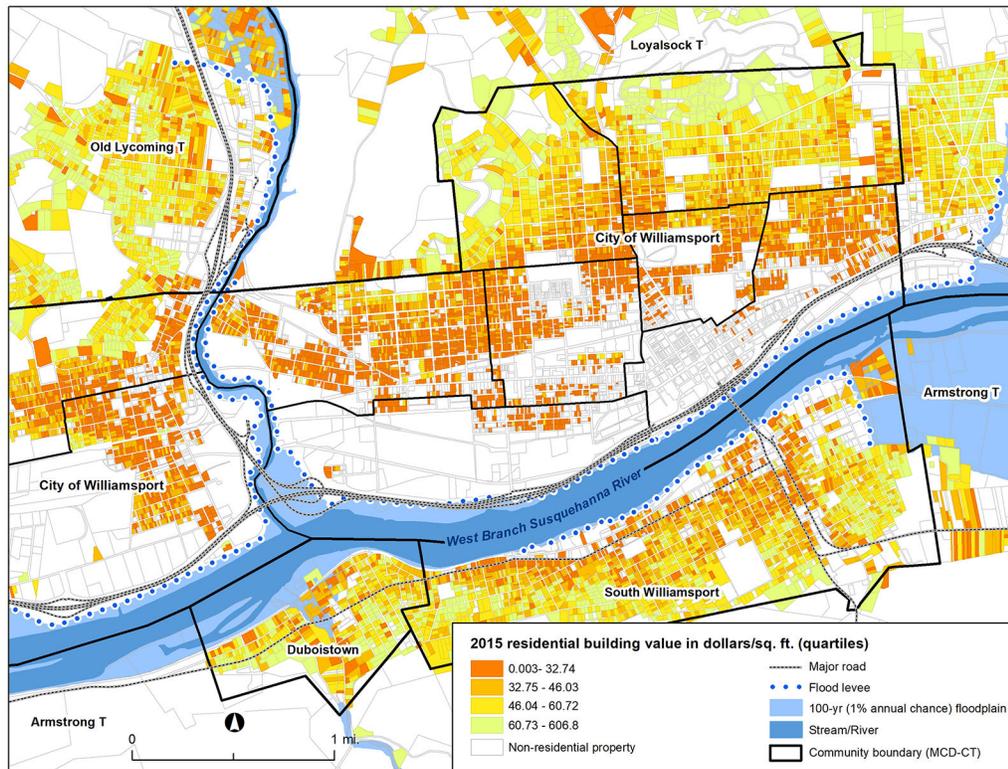


Figure 2. Property values of residential tax parcels in the Williamsport Metropolitan Statistical Area, Lycoming County, Pennsylvania, USA. Abbreviations: T, township

Table 1.

Selected characteristics of T2D cases and controls without T2D, frequency-matched (5:1) on age, sex, and year of diagnosis or encounter date

Variable	n (%) unless otherwise noted		p-value ¹
	Cases	Controls	
Number of individuals	15,888	79,435	n/a
Sex, female	7,798 (49.1)	38,988 (49.1)	matched
Age at diagnosis or medical encounter, years, mean (SD)	54.9 (15.1)	54.9 (15.3)	0.82
Age, years, categories			
10 to < 20 years	304 (1.9)	1,520 (1.9)	
20 to < 30 years	628 (4.0)	3,140 (4.0)	
30 to < 40 years	1,611 (10.1)	8,055 (10.1)	
40 to < 50 years	3,086 (19.4)	15,429 (19.4)	
50 to < 60 years	4,286 (27.0)	21,428 (27.0)	matched
60 to < 70 years	3,510 (22.1)	17,548 (22.1)	
70 to < 80 years	1,737 (10.9)	8,685 (10.9)	
80 to < 90 years	645 (4.1)	3,225 (4.1)	
90 years	81 (0.5)	405 (0.5)	
Race, white	15,429 (97.1)	77,867 (98.0)	< 0.001
Hispanic ethnicity	369 (2.3)	1,094 (1.4)	< 0.001
Medical Assistance, 50%	967 (6.1)	2,730 (3.4)	< 0.001
Community type			
Township	9,461 (59.6)	51,131 (64.4)	< 0.001
Borough	4,621 (29.1)	21,756 (27.4)	Ref
Census tract in city	1,806 (11.4)	6,548 (8.2)	< 0.001
Residential distance to blue space			
0 to < 0.25 miles	3,485 (21.9)	16,120 (20.3)	Ref
0.25 to < 0.50 miles	3,340 (21.0)	15,520 (19.5)	0.869
0.50 to < 0.75 miles	2,313 (14.6)	11,476 (14.5)	0.022
0.75 to < 1.25 miles	3,483 (21.6)	17,369 (21.9)	0.001
1.25 miles	3,312 (20.8)	18,950 (23.9)	< 0.001
Type of nearest water polygon			
Lake	1,674 (10.5)	9,450 (11.9)	Ref
River	3,744 (23.6)	16,629 (20.9)	< 0.001
Tributary	3,044 (19.2)	14,427 (18.2)	< 0.001
Stream	7,426 (46.7)	38,929 (49.0)	0.014
Within 100-year floodplain	761 (4.8)	3,301 (4.2)	< 0.001
Community socioeconomic deprivation ²			
Quartile 1 (low deprivation)	3,001 (18.9)	17,329 (21.8)	Ref
Quartile 2	4,300 (27.1)	23,172 (29.2)	0.009
Quartile 3	4,217 (26.5)	20,327 (25.6)	< 0.001

Variable	n (%) unless otherwise noted		p-value ¹
	Cases	Controls	
Quartile 4 (high deprivation)	4,369 (27.5)	18,606 (23.4)	< 0.001

¹Because controls could be in these comparisons more than once, methods were used for significance testing that accounted for this, including inverse-probability weighted regression for time-invariant characteristics, mixed-effect regression for time-varying continuous (linear), binary (logistic), and count (Poisson) characteristics, and multinomial logistic regression with robust standard errors for polytomous time-varying characteristics. In the weighted analyses, weights were the number of appearances in the analysis (implemented with a dataset having only one record per person).

²Quartile cutoffs were defined with the three time periods; the range of values for Q1, Q2, Q3, and Q4 were -18.33 to -1.96; -1.99 to -0.015; 0.005 to 2.05; and 2.11 to 12.4.

Table 2.

Adjusted¹ associations of distance to blue space, residence in the 100-year floodplain, and type of nearest blue space with type 2 diabetes status, stratified by community type

Variable	Townships OR (95% CI)	Boroughs OR (95% CI)	Cities OR (95% CI)
<i>Distance to blue space models</i>			
Distance to blue space			
0 to < 0.25 miles	1.08 (1.00, 1.17)	1.17 (1.01, 1.37)	1.23 (0.99, 1.53)
0.25 to < 0.50 miles	1.07 (0.99, 1.16)	1.14 (0.97, 1.34)	1.39 (1.14, 1.70)
0.50 to < 0.75 miles	1.04 (0.96, 1.13)	0.97 (0.83, 1.14)	1.38 (1.10, 1.72)
0.75 to < 1.25 miles	1.04 (0.97, 1.10)	1.03 (0.87, 1.21)	1.16 (0.93, 1.44)
1.25 miles	1.0	1.0	1.0
<i>Residence in the 100-year floodplain models</i>			
Residence in floodplain			
Yes	1.16 (1.02, 1.31)	1.14 (1.00, 1.29)	0.72 (0.50, 1.04)
No	1.0	1.0	1.0
<i>Type of nearest blue space models</i>			
Type of nearest blue space			
River	1.08 (1.00, 1.16)	1.03 (0.94, 1.13)	0.99 (0.86, 1.14)
Tributary	1.07 (0.99, 1.15)	1.01 (0.90, 1.13)	1.19 (0.99, 1.42)
Lake	1.00 (0.93, 1.08)	0.90 (0.76, 1.06)	0.68 (0.54, 0.85)
Stream	1.0	1.0	1.0

¹Models were adjusted for sex, age, age², age³, race, ethnicity, Medical Assistance, and community socioeconomic deprivation. One exception was the type of nearest blue space model in cities, which could not be adjusted for CSD due to low cell counts.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript