The Relationship Between Trees and Human Health
Evidence from the Spread of the Emerald Ash Borer

Geoffrey H. Donovan, PhD, David T. Butry, PhD, Yvonne L. Michael, ScD,
Jeffrey P. Prestemon, PhD, Andrew M. Liebhold, PhD,
Demetrios Gatziolis, PhD, Megan Y. Mao

**Background:** Several recent studies have identified a relationship between the natural environment and improved health outcomes. However, for practical reasons, most have been observational, cross-sectional studies.

**Purpose:** A natural experiment, which provides stronger evidence of causality, was used to test whether a major change to the natural environment—the loss of 100 million trees to the emerald ash borer, an invasive forest pest—has influenced mortality related to cardiovascular and lower-respiratory diseases.

**Methods:** Two fixed-effects regression models were used to estimate the relationship between emerald ash borer presence and county-level mortality from 1990 to 2007 in 15 U.S. states, while controlling for a wide range of demographic covariates. Data were collected from 1990 to 2007, and the analyses were conducted in 2011 and 2012.

**Results:** There was an increase in mortality related to cardiovascular and lower-respiratory-tract illness in counties infested with the emerald ash borer. The magnitude of this effect was greater as infestation progressed and in counties with above-average median household income. Across the 15 states in the study area, the borer was associated with an additional 6113 deaths related to illness of the lower respiratory system, and 15,080 cardiovascular-related deaths.

**Conclusions:** Results suggest that loss of trees to the emerald ash borer increased mortality related to cardiovascular and lower-respiratory-tract illness. This finding adds to the growing evidence that the natural environment provides major public health benefits.


Introduction

There is increasing evidence from multiple scientific fields that exposure to the natural environment can improve human health. However, existing research has often been hampered by cross-sectional study design and a failure to adequately address confounding factors. Quasi-experimental designs, such as the pioneering work by Ulrich, who showed that patients recovered faster from surgery in a room with a view of a natural scene than those without such a view, are needed to provide stronger evidence of a causal link between the natural environment and health.

To address this gap in the literature, a longitudinal study design was used to quantify the public health effects of an introduced forest pest, the emerald ash borer, which has killed tens of millions of ash trees since it was first detected in the U.S. in 2002. The spread of the borer is a unique natural experiment allowing the evaluation of the effect of changes in the natural environment on public health. The goal of the study was not to track the borer per se, but to use it as a proxy for tree loss.
Natural experiments approximate RCTs, as the mechanism determining exposure is independent of the outcome, and therefore, common prior causes of exposure and outcome are equally distributed between those exposed and those unexposed. The borer spreads directly from county to county, but it is spread also by accidental transport—typically on firewood—which results in satellite populations (Figure 1). This accidental spread adds an important random element to the current natural experiment. Nonetheless, natural experiments remain observational studies and cannot prove causality.

This study examined whether the spread of the emerald ash borer is associated with increased mortality related to cardiovascular and lower-respiratory-tract illness. These two types of health issues were chosen because they are the first and third most common causes of death in the U.S., and there are plausible mechanisms linking these types of deaths with trees. Specifically, the natural environment has been shown to decrease stress, increase physical activity, and improve air quality. In turn, stress, lack of physical activity, and poor air quality have been linked with cardiovascular and lower-respiratory-tract disease.

The pioneering work in the field by Ulrich found that patients recovering from gall bladder—removal surgery in a room with a view of a natural scene recovered faster and took fewer pain medications than patients in a room with a view of a brick wall. However, extending Ulrich’s work has been problematic, because most health outcomes of interest have causes that long precede the short surgical recovery period, and most people spend little time in environments as controlled as a hospital room.

Observational studies of the relationship between the natural environment and health have examined a range of health outcomes. Mitchell and Popham found that, after controlling for SES, “greenness” was negatively associated with overall mortality in England. This relationship was particularly strong for cardiovascular-related mortality. Takona and colleagues studied the 5-year survival rate of 3144 senior citizens living in Tokyo. They found a positive association between survival rate and access to walkable green space.

In Holland, Maas et al. reported a positive association between greenness and self-reported health. In a later study, Maas et al. analyzed the health records of 345,000 people. They found that those living in greener areas were less likely to be diagnosed with 15 of the 24 health outcomes examined. Results were particularly strong for anxiety and depression and for children and those with lower SES. Park and colleagues showed that walking in a forest reduced heart rate and cortisol levels. Finally, in New York City, Lovasi et al. found that children who lived in areas with more street trees were less likely to have asthma.

Two studies have examined the relationship between the natural environment and birth outcomes. Donovan et al. found that mothers living in Portland OR with more tree canopy within 50 meters of their homes, or who lived closer to open space, were less likely to have a baby...
that was small for gestational age. Dadvand and colleagues\(^1\) conducted a similar study in Spain. They found that women with more greenness within 100 meters of their homes, or who lived within 500 meters of a major green space, gave birth to heavier babies, although results only held for women with the lowest level of education.

**Emerald Ash Borer**

The emerald ash borer, *Agrilus planipennis*, is a phloem-feeding borer native to East Asia. It was discovered in North America in 2002, when it was identified as the cause of widespread ash mortality (*Fraxinus* spp.) in Detroit MI and nearby Windsor, Ontario.\(^22\) By 2012, this borer had killed approximately 100 million trees in the U.S. (D. McCullough, Michigan State University, personal communication, 2012). However, its potential impact is much larger, as there are 7.5 billion ash trees in the country.\(^22\) In addition, the borer kills all 22 species of North American ash and virtually all infested trees, so it is a good proxy for ash tree death. For more information about this borer, see the video in Appendix A (available online at www.ajpmonline.org).

**Methods**

**Study Area and Data**

Data were collected from 1990 to 2007, and the analyses were conducted in 2011 and 2012. The study sample consists of the 15 states that had at least one confirmed case of the borer in 2010. Data were observed annually at the county level (1296 counties), from 1990 through 2007 (maximum number of observations=22,032, but because of missing data, the actual number of observations=21,080).

This sample allowed observation of mortality before and after 2002, when the emerald ash borer was initially discovered in the U.S. By 2007, the U.S. Department of Agriculture (USDA) Animal and Plant Inspection Service had detected the borer in 244 counties (once a county is infected, it remains infected; Figure 1). Two variables were used to describe the presence of the borer: a dummy variable, which takes on a value of 1 in infested counties, and a continuous variable (0–6) that denotes the number of years it has been present in a county.

Mortality data were obtained from the National Center for Health Statistics\(^23,24\) and stratified by age (<18 years and ≥18 years) and cause of death (major cardiovascular disease [ICD-10 Codes I00-I78]; chronic lower-respiratory-tract disease [ICD-10 codes J40-47]; accidental death [ICD-10 Codes V01-X59, Y85-86]). Demographic covariates were chosen based on neighborhood determinants of cardiovascular and lower-respiratory-tract mortality.\(^25\) Demographic data were obtained from the 1990 and 2000 censuses and the 2009 American Community Survey. The authors estimated census variables for all other years by interpolation.

The impact of the borer—on tree mortality and public health—depends on the number and distribution of ash trees in a county. Unfortunately, comprehensive data on ash abundance are not available. Therefore, a two-stage process was used to estimate ash-canopy cover, which is the area occupied by a tree’s crown when viewed from above.

First, total tree canopy was estimated in a county using National Land Cover Data (NLCD) raster maps from 1992, 2001, and 2006. Tree canopy for all other years was estimated by linear interpolation or extrapolation from these 3 years. The NLCD maps were generated by processing 30-meter-resolution satellite imagery using class definitions that have remained consistent through time.\(^26\)

Second, total tree canopy in a county was weighted by the proportion of ash in a state (ash as a percentage of total tree canopy varied from a low of 1.5% in Virginia to a high of 7.9% in New York). For example, if a county had 40% tree canopy and 5% of tree canopy was ash, then ash canopy was 2%. Data from the USDA Forest Service’s Forest Inventory and Analysis (FIA) program were used to estimate statewide ash abundance. State-level data, as opposed to county-level, were used to estimate ash-canopy cover, as some counties have little forestland and therefore have few plots.

**Data Analysis**

Two regression models were estimated relating the presence of the borer with rates of adult mortality related to cardiovascular and lower-respiratory-tract illness. Models of the following general form can be fit to longitudinal data (where \(i\) denotes county and \(t\) denotes time):

\[
Y_{i,t} = \beta + \beta X_{i,t} + v_i + e_{i,t},
\]

where \(Y_{i,t}\) is the mortality rate (per 100,000 adults); \(X_{i,t}\) is a vector of independent variables; \(v_i\) is an i.i.d. error term uncorrelated with the county-specific residual \(v_i\); and \(\beta\) is coefficients that are estimated in the regression step. Typically, linear models of this form are estimated using either fixed-effects or random-effects estimators. Fixed-effect estimators were used, as a Hausman specification test showed that the assumptions underlying the random-effects estimators were not met.\(^27\)

Heteroskedasticity is a common problem in panel-data analysis. It can arise when observational units vary greatly in scale. In this analysis, mortality rates were used rather than number of deaths, which addresses the issue of scaling. However, in less-populous counties, mortality counts are low, which means mortality rates are sensitive to small changes in mortality counts (counties with low mortality counts were not dropped from the analysis). Therefore, a priori, error-term variance is expected to be higher in low-population counties. For this reason, model coefficients were estimated with heteroskedasticity-robust fixed-effects estimators.\(^28\)

Variables were selected for inclusion in the final model using iterative backwards selection. Progressively lower significance thresholds were used with a final threshold of 0.1. A variance–covariance matrix was used to avoid including highly collinear combinations of variables in the model. When similar demographic variables were collinear (those describing income, for example), the variable from a group that had the lowest \(p\)-value when individually regressed against mortality was used.

Each variable that was dropped from the model was checked to determine whether it varied significantly between counties that were infested and those that were not. If a variable did vary, then it was reintroduced and retained, if it caused the coefficients on the two borer variables to change by more than 10%.\(^29\) None of the reintroduced variables met this threshold.

In addition to controlling for potential confounders, all models included a linear time-trend variable (1–18 years) to account for
broad trends in mortality—improved medical technology, for example—that would not be captured by demographic covariates. In addition, a 1-year lag of mortality rate was included to address temporal autocorrelation (AR(1) correction). Finally, a variable denoting the amount of ash-canopy cover in a county was included, because if the borer does have a negative public health effect, then one would expect ash to have a positive effect, especially in counties not yet infected.

Interaction terms between demographic covariates and both presence of the borer in a county and amount of ash also were included. This was done because past research has shown that access to greenness varies among demographic groups. Thus, the borer’s impact would be expected to depend on the demographic makeup of a county.

Natural experiments provide stronger evidence of causation than observational studies, but it is still possible that the results were influenced by an omitted variable that is correlated with the spread of the emerald ash borer. Therefore, to provide an additional safeguard, a model was estimated with accidental death as the dependent variable, because this is a type of death that the borer could not plausibly affect (the same model selection criteria were used as those used for the cardiovascular and lower-respiratory-tract models).

### Results

#### Respiratory-Related Mortality

Regression results for the respiratory-related mortality model are shown in Table 1. The negative coefficient on the time trend confirms that overall respiratory-related mortality declined over the 18-year study period. The presence of the borer was significant by itself and in interaction with years of infestation and median income (dichotomized at the median split).

The positive coefficient on the income interaction term suggests that the borer has a bigger effect on mortality in wealthier counties. This is consistent with previous research showing a positive correlation between tree cover and income in urban areas. This result was mirrored by the effect of ash on mortality related to lower-respiratory-tract illness, as counties with more ash trees had lower rates of this type of mortality. In addition, the effect of the interaction term between ash trees and median income is negative, which suggests that the health benefits of ash are greater in wealthier counties.

Given the significance of the interaction terms, the coefficients on these terms cannot be interpreted in isolation. Therefore, the delta method was used to calculate the net marginal effect of the borer on respiratory-related mortality taking into account the direct effect and the effect through interaction terms. The presence of the borer in a county is associated with 6.8 additional deaths per year per 100,000 adults (95% CI = 4.8, 8.7).

To determine how the effect changes over time, separate marginal effects were calculated for each year of infestation (Table 2). Results show that as infestation in a county progresses, the magnitude of the marginal effect also increases. Applying these marginal effects to the appropriate infested counties shows that the emerald ash borer was associated with 6113 excess deaths between 2002 and 2007.

The delayed effect of the borer may be due to the 2–5 years it takes an ash tree to die after initial infestation. In addition, any effect on human mortality would be expected to lag behind the borer’s effect on tree mortality. Indeed, it may be surprising that the borer has any effect on human mortality in the first year of infestation. However, once the borer is detected in a county, often many healthy trees are cut down

<table>
<thead>
<tr>
<th>Variable</th>
<th>Beta coefficienta (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time trend</td>
<td>$-2.98 (-3.23, -2.72)$</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>1-year mortality-rate lag</td>
<td>0.31 (0.303, 0.310)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Percentage non-Hispanic white</td>
<td>9.40 (6.40, 12.40)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Percentage Native Hawaiian and other Pacific Islander</td>
<td>2.14 (0.32, 3.97)</td>
<td>0.022</td>
</tr>
<tr>
<td>High median income</td>
<td>13.95 (6.50, 21.39)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Aged ≥25 years with no high school diploma, %</td>
<td>1.22 (0.92, 1.52)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Aged ≥25 years with college degree, %</td>
<td>$-0.33 (-0.70, 0.03)$</td>
<td>0.077</td>
</tr>
<tr>
<td>Population below 100% of poverty line, %</td>
<td>2.24 (1.89, 2.58)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Percentage of county covered by ash canopy</td>
<td>$-5.22 (-7.79, -2.64)$</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Emerald ash borer</td>
<td>$-4.24 (-8.10, -0.39)$</td>
<td>0.031</td>
</tr>
<tr>
<td>Emerald ash borer X high median income</td>
<td>6.23 (2.23, 10.22)</td>
<td>0.002</td>
</tr>
<tr>
<td>Years of infestation</td>
<td>1.44 (0.95, 1.92)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Ash canopy X high median income</td>
<td>$-0.85 (-1.30, -0.41)$</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.609</td>
<td></td>
</tr>
<tr>
<td>Within counties</td>
<td>0.187</td>
<td></td>
</tr>
<tr>
<td>Between counties</td>
<td>0.352</td>
<td></td>
</tr>
</tbody>
</table>

*Mortality rate per 100,000 adults

---

to prevent its spread. This practice was particularly common during the early years of its spread.

In addition, there is often extensive media coverage when the borer is first detected in a county, which may cause some of the same stressful responses as tree mortality. For example, within the first year of discovery in Michigan, the *Detroit Free Press* ran 39 stories on the borer including four on the front page. Similarly, the *Chicago Tribune* printed 53 stories relating to the borer in the first year of infestation in Illinois, including 16 on the front page. For the current paper, stories were identified using *emerald ash borer* as a search term. Each story was checked to make sure the emerald ash borer was in fact the subject.

**Cardiovascular-Related Mortality**

Results for the cardiovascular model are shown in Table 3. The significance and magnitude of the coefficients on the borer and borer interaction terms are similar to those for the respiratory model. The effects of other covariates are generally consistent, although ash is not associated with cardiovascular-related mortality except in interaction with median income.

The marginal effect of the borer on cardiovascular-related mortality is 16.7 additional deaths per year per 100,000 adults (95% CI=5.7, 27.7) for a total of 15,080 excess deaths from 2002 to 2007 (Table 4).

In the accidental-mortality model, no effect of the borer was found. Specifically, using a Wald test, the null hypothesis that the coefficients on the borer, the interaction of the borer with median income, and years of infestation were jointly zero was not rejected \( (p=0.22) \). Although the accidental-death model is not a control in a formal sense, it is encouraging that the model found the borer to be uncorrelated with a cause of death it could not plausibly affect.

**Discussion**

Results suggest that the widespread death of ash trees from the emerald ash borer lead to an increase in mortality related to cardiovascular and lower-respiratory-tract illness. These results are consistent with previous research that has identified a correlation between the natural environment and health. They also provide stronger support for a causal relationship.

The borer had a greater effect in counties whose median household income was above average. There are a number of possible interpretations for these results. People in wealthier counties may have greater access to ash trees, so the death of these trees has a greater impact on
The possible mechanisms linking the natural environment,14 moderating temperature,36 and buffering stressful life events.37 Future research could fruitfully investigate them. In particular, urban areas within wealthier counties may have more trees, or better maintain them. Indeed, past studies have found that within a city, wealthier neighborhoods have more tree-canopy cover.32

It also is possible that trees provide different benefits in wealthier areas. For example, Troy and Grove34 found that proximity to urban parks increased the sales price of homes in wealthier neighborhoods, whereas, in poor neighborhoods, houses close to parks sold for less. The authors suggest that parks may attract criminal behavior in poorer neighborhoods, so residents are not able to benefit from the park as much as people living in a wealthier neighborhood. In addition, risk factors such as air quality, which trees can mediate, may be different in wealthier counties.

Results do not provide any direct insight into how trees might improve mortality rates related to cardiovascular and lower-respiratory-tract illness. However, there are several plausible mechanisms including improving air quality,11,35 reducing stress,13 increasing physical activity,14 moderating temperature,36 and buffering stressful life events.37 Future research could fruitfully investigate the possible mechanisms linking the natural environment and health.

Limitations
This study has several limitations. It is possible that despite the natural-experiment design, the results are an artifact of an omitted risk factor that is correlated with the borer or residual confounding. The authors believe that this possibility is unlikely for three reasons. First, a wide range of covariates that have been shown to influence mortality related to cardiovascular and lower-respiratory system illness were included in the model. Second, a confounder would have to be strongly correlated with the borer across both space and time. Third, an omitted variable would need to influence these types of mortality but not accidental death.

Nonetheless, it is re-emphasized that this is an observational study, and the results await confirmation. In addition, this is an ecologic study, so the overall results do not necessarily apply to a particular county or group of counties. Finally, even well-controlled ecologic studies can be subject to ecologic bias.

The variables used to denote the borer and ash are another potential source of error. Specifically, data availability forced use of a simple dummy variable to denote the presence of the borer, and past research has shown that modeling a continuous process with a binary variable can result in coefficients that are biased upward.38 In contrast, ash abundance was measured continuously. However, the ash coverage variable is a composite of county-level canopy-cover data and state-level data on ash abundance and is, therefore, a coarse approximation of the true amount of ash in a county. However, when models were estimated in which ash-canopy cover was replaced with canopy cover from all tree species, there was little change in the coefficients of borer-related variables. Therefore, the choice of ash variable does not affect the conclusions about the relationship between the borer and mortality.

Conclusion
Tree loss from the spread of the emerald ash borer is associated with increased mortality related to the cardiovascular and lower-respiratory systems. This relationship is particularly strong in counties with above-average median household income.

No financial disclosures were reported by the authors of this paper.

References

Table 4. Estimated marginal effects of the emerald ash borer on cardiovascular-related mortality by years of infestation

<table>
<thead>
<tr>
<th>Years of infestation</th>
<th>Number of observations</th>
<th>Marginal effect (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>126</td>
<td>9.68 (1.79, 17.57)</td>
<td>0.006</td>
</tr>
<tr>
<td>2</td>
<td>102</td>
<td>14.68 (7.09, 22.28)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>3</td>
<td>72</td>
<td>19.90 (11.4, 28.4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>24.61 (14.27, 34.96)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>5</td>
<td>22</td>
<td>33.58 (21.14, 46.02)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>38.59 (23.62, 53.56)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Across all years</td>
<td></td>
<td>16.70 (5.73, 27.67)</td>
<td>0.001</td>
</tr>
<tr>
<td>Wald test: all variables²=0</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

²Emerald ash borer, years of infestation, and interaction terms