ARTICLES
https://doi.org/10.1038/s41477-018-0151-8
nature plants

Vegetation diversity protects against childhood asthma: results from a large New Zealand birth cohort

Geoffrey H. Donovan[1,2*], Demetrios Gatziolis[2], Ian Longley[3] and Jeroen Douwes[1]

We assessed the association between the natural environment and asthma in 49,956 New Zealand children born in 1998 and followed up until 2016 using routinely collected data. Children who lived in greener areas, as measured by the normalized difference vegetation index, were less likely to be asthmatic: a 1 s.d. increase in normalized difference vegetation index was associated with a 6.0% (95% CI 1.9–9.9%) lower risk of asthma. Vegetation diversity was also protective: a 1 s.d. increase in the number of natural land-cover types in a child's residential meshblock was associated with a 6.7% (95% CI 1.5–11.5%) lower risk. However, not all land-cover types were protective. A 1 s.d. increase in the area covered by gorse (*Ulex europaeus*) or exotic conifers, both non-native, low-biodiversity land-cover types, was associated with a 3.2% (95% CI 0.0–6.0%) and 4.2% (95% CI 0.9–7.5%) increased risk of asthma, respectively. The results suggest that exposure to greenness and vegetation diversity may be protective of asthma.

Since the pioneering work of Roger Ulrich[1]—who found that patients with a view of a natural scene recovered more quickly from surgery—research has shown that exposure to the natural environment is associated with a wide range of beneficial health outcomes, including lower mortality[2,3], higher birth weight[2–6], reduced risk of cardiovascular disease[7] and improved mental health[8,9]. However, most studies have used a cross-sectional study design that measured the natural environment and health outcomes at the same point in time. This is a major limitation, as the health consequences of environmental exposures often do not manifest immediately[10], and timing of exposure across the life course may strongly affect the outcome, with the first few years of life believed to be particularly critical[11]. In addition, past studies have relied heavily on a single measure of overall greenness called the normalized difference vegetation index (NDVI), which is typically derived from satellite imagery. However, markedly different landscapes can have similar NDVI values, so the use of NDVI has limited researchers' capacity to determine which aspects of the natural environment provide the greatest public-health benefits.

Several recent studies suggest that exposure to the natural environment may be protective of asthma, although the evidence is mixed[12]. Asthma is a chronic health condition that affects 334 million people worldwide and its prevalence is increasing[13]. Although asthma treatment has improved, it is not effective for all asthmatics, and there is no cure. In addition, there are limited public-health interventions that can reduce asthma risk, and these interventions are often difficult to implement. Therefore, there is considerable interest in understanding how exposure to the natural environment may protect against asthma and in identifying what specific aspects of the natural environment afford the most protection.

An ecological study in 42 health catchment areas in New York City[14] showed that areas with more street trees within 1 square kilometre had lower rates of asthma among 4–5 year olds, but there was no relationship between tree density and hospitalizations for asthma. A follow-up study[15], using individual-level data from 549 African American and Dominican children, who lived in socially deprived neighbourhoods in New York City, found that tree cover was not protective of asthma. Furthermore, more trees within 250 m of the prenatal address was a risk factor for allergic sensitization.

In a 10-year follow-up study of a large ($n=65{,}000$) birth cohort in Vancouver, British Columbia[16], it was found that greenness was protective of early-life asthma, but not of school-age asthma. The authors measured greenness using NDVI. The results held after controlling for air pollution and proximity to major roads. A study[17] on greenness (measured using NDVI) and asthma rates in 14 metropolitan statistical areas in Texas found no association, although this may be an artefact of the ecological study design and small sample size.

A study[18] in 3,178 9–12 year olds in Sabadell, Spain, found no relationship between greenness around a child's residential address (NDVI in buffers of 100 m to 1,000 m) and asthma. However, children who lived closer to a park were at greater risk, which the authors speculate may be because parks in Catalonia have more non-native species than other green areas.

A recent cross-sectional study[19] in Australia, not focused exclusively on asthma, examined the relationship between the natural environment and rates of hospital admission for respiratory disease in local government areas. They found that multiple measures of the natural environment were protective of a respiratory disease.

One reason for the observed inconsistent association between the natural environment and asthma may be that NDVI is a relatively coarse measure of greenness: different plants can have similar spectral profiles. In addition, most studies only measured greenness at one point in time and a single measure is unlikely to accurately represent a child's lifetime exposure. Indeed, using a one-point-in-time measure may result in systematic measurement error as demonstrated previously by Brokamp et al.[20]. The authors tracked all the residential addresses (from birth to age 7) of a cohort of 613 children in Cincinnati, OH, and showed that, across time, children moved to less socially deprived, greener neighbourhoods, with

[1]Center for Public Health Research, Massey University, Wellington, New Zealand. [2]USDA Forest Service, PNW Research Station, Portland, OR, USA. [3]National Institute of Water and Atmospheric Research, Auckland, New Zealand. *e-mail: gdonovan@fs.fed.us

Table 1 | Sample descriptive statistics ($n = 49,956$)*

Variable	Mean	s.d.
Ethnicity: European white (%)	55.8	49.7
Ethnicity: Māori (%)	28.3	45.1
Ethnicity: Pacific (%)	8.4	27.8
Ethnicity: Asian (%)	5.3	22.4
Ethnicity: MELAA† (%)	0.38	6.2
Ethnicity: other (%)	0.56	7.8
Father's age in 1998	32	6.4
Mother's age in 1998	29	5.7
Father no high-school qualifications (%)	27.6	44.7
Mother no high-school qualifications (%)	22.4	41.7
Father never smoked (%)	53.3	49.9
Mother never smoked (%)	50.6	50.0
Number of antibiotic scripts (2005–2016)	6.2	7.0
Born before 37 weeks gestational age (%)	6.0	23.8
Number of siblings	1.5	1.5
Always lived in a rural meshblock (%)	6.8	25.3
Mean lifetime land cover: built (%)	29.3	13.5
Mean lifetime land cover: agricultural (%)	10.8	12.8
Mean lifetime land cover: forest (%)	3.8	5.2
Mean lifetime land cover: water (%)	0.89	2.3
Mean lifetime land cover: non-native conifer (%)	1.3	4.4
Mean lifetime land cover: gorse (%)	0.37	2.5
Mean lifetime NDVI	0.549	0.091

*Following IDI protocols, all sample sizes have been rounded to the nearest multiple of three.
†Aggregate ethnicity category used in New Zealand denoting Middle Eastern, Latin American or African (MELAA) ethnicity.

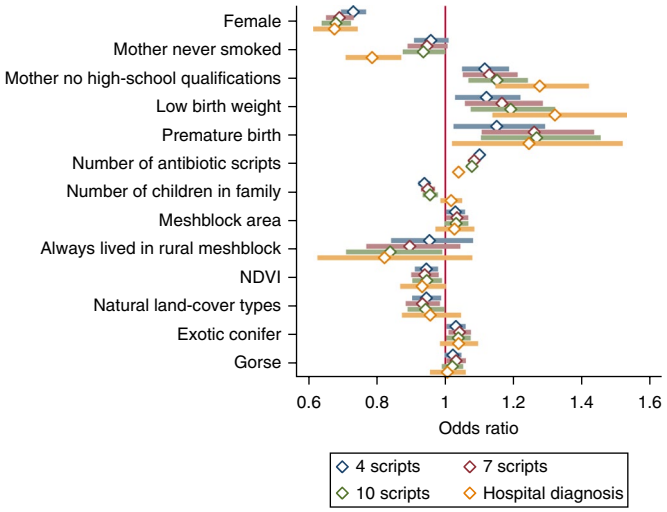

Fig. 1 | Odds ratio plots of asthma prevalence among children born in New Zealand in 1998 ($n = 39,108$). Four different definitions of asthma: four or more asthma scripts or hospital diagnosis; seven or more asthma scripts or hospital diagnosis; ten or more asthma scripts or hospital diagnosis; hospital diagnosis only. Following IDI protocols, all sample sizes have been rounded to the nearest multiple of three.

less traffic-related air pollution. In addition, they found that residential greenness affected the probability that a child would move. Therefore, a child's birth address, for example, would systematically underestimate lifetime greenness exposure. The problem of using a one-point-in-time greenness measure may be compounded by differential responses to the natural environment across a child's life course. Other methodological limitations that may have contributed to inconsistent results include small sample size, ecological study design and inadequate control for other risk factors.

The mechanisms linking greenness and asthma remain unclear, but may include improved air quality[21], reduced stress[22] and increased physical activity[23]. In addition, the relationship between greenness and asthma may be partly explained by the hygiene hypothesis: exposure to the natural environment may increase microbial exposure, resulting in improved immune function, and subsequent decreased risk of immunoglobulin E-mediated allergic diseases, such as allergic asthma[24].

We address the limitations and inconsistencies of the available literature by assessing the association between asthma and the natural environment in a large ($n = 49,956$) New Zealand birth cohort followed for 18 years. We make use of a uniquely rich source of longitudinal, individual-level data: the integrated data infrastructure (IDI) maintained by Statistics New Zealand[25]. The IDI consists of multiple datasets describing drug prescriptions and hospitalizations, education and benefits, criminal justice, social deprivation, population (births, deaths and immigration), income and work, and residential history, all linked by a common ID number for each individual for the majority of the New Zealand population.

Results

Table 1 provides descriptive statistics for selected exposures and covariates.

Table 2 shows associations between exposure to the natural environment and asthma prevalence. Model 1 accounts for overall greenness; model 2 accounts for overall greenness plus vegetation diversity (number of land cover types); model 3 accounts for overall greenness, vegetation diversity and exposure to specific land-cover types.

Mean greenness across a child's life course (prenatal to 18) was protective of asthma: a 1 s.d. increase in mean lifetime NDVI in a child's residential meshblock was associated with a 6.0% (95% CI 1.9–9.9%) lower risk of asthma. Vegetation diversity was also protective of asthma: a 1 s.d. increase in the mean lifetime number of natural land-cover types in a child's residential meshblock was associated with a 6.7% (95% CI 1.5–11.5%) reduction in asthma risk. In contrast, two land-cover types were a risk factor for asthma. A 1 s.d. increase in mean lifetime exposure to gorse (*Ulex europaeus*) in a child's residential meshblock was associated with a 3.2% (95% CI 0.0–6.0%) increase in asthma risk, and a 1 s.d. increase in mean lifetime exposure to exotic conifers in a child's residential meshblock was associated with a 4.2% (95% CI 0.9–7.5%) increased risk. The percentage risk reduction was calculated by subtracting the hazard ratio from 1 and multiplying by 100 (if the hazard ratio is greater than 1, then this translates into an increase in risk). For example, the hazard ratio on mean lifetime NDVI in model 3 is 0.940, so the percentage risk reduction is 6%.

The protective effect of lifetime NDVI exposure was attenuated by controlling for the number of natural land-cover types (model 2) and further attenuated by controlling for exposure to gorse and exotic conifers (model 3). In contrast, controlling for gorse and exotic conifers increased the protective effect of the number of natural land-cover types (model 3). Meshblock size was included in all models, so number of land-cover types is not simply a proxy for larger meshblocks.

Consistent with previous research, we found that girls are less likely to have asthma than boys. Asian and Māori children had higher rates of asthma than Europeans, whereas Pacific Island children had lower rates. Underweight or premature birth, the number

Table 2 | Logit regression results of asthma prevalence among children born in New Zealand in 1998 ($n = 39{,}108$)[†,‡]

Variables	Model 1 (NDVI)		Model 2 (NDVI plus number of land cover types)		Model 3 (NDVI plus number and type of land cover)	
	Odds ratios	95% CI	Odds ratios	95% CI	Odds ratios	95% CI
Female	0.692***	0.653–0.734	0.693***	0.653–0.734	0.691***	0.652–0.733
Ethnicity (European white excluded)						
Asian	1.386***	1.215–1.582	1.381***	1.210–1.576	1.381***	1.210–1.576
Māori	1.233***	1.147–1.324	1.231***	1.146–1.322	1.232***	1.147–1.323
MELAA[§]	1.119	0.603–2.077	1.116	0.601–2.070	1.122	0.605–2.082
Pacific	1.01	0.680–1.500	1.011	0.681–1.502	1.012	0.681–1.504
Other	0.863**	0.762–0.976	0.857**	0.757–0.970	0.856**	0.756–0.969
Mother never smoked	0.947*	0.890–1.007	0.947*	0.890–1.008	0.947*	0.890–1.007
Mother no high-school qualification	1.129***	1.052–1.212	1.127***	1.050–1.210	1.128***	1.051–1.211
Low birth weight	1.167***	1.059–1.287	1.167***	1.059–1.286	1.165***	1.057–1.285
Premature birth	1.260***	1.107–1.435	1.260***	1.107–1.435	1.259***	1.106–1.434
Number of antibiotic scripts	1.084***	1.080–1.088	1.084***	1.080–1.088	1.084***	1.080–1.089
Number of children in family	0.949***	0.929–0.971	0.949***	0.928–0.970	0.949***	0.929–0.971
Always lived in a rural meshblock	0.853**	0.738–0.985	0.886	0.761–1.032	0.896	0.770–1.044
Mean lifetime NDVI (standardized)	0.919***	0.888–0.951	0.938***	0.899–0.978	0.940***	0.901–0.981
Meshblock area (standardized)	1.021	0.992–1.051	1.032**	1.000–1.064	1.034**	1.002–1.067
Number of natural land-cover types (standardized)			0.96	0.913–1.009	0.933**	0.885–0.985
Exotic conifer land cover (standardized)					1.042**	1.009–1.075
Gorse land cover (standardized)					1.032**	1.004–1.060

***$P < 0.01$, ** $P < 0.05$, * $P < 0.1$. [†]Following IDI protocols, all sample sizes have been rounded to the nearest multiple of three. [‡]The sample size in our regression model was lower than the full sample because of missing data. Variables derived from the 2013 census (education and smoking status) accounted for the majority of missing values. [§]Aggregate ethnicity category used in New Zealand denoting Middle Eastern, Latin American or African ethnicity.

of antibiotic prescriptions and having a mother who did not receive any high-school qualifications also increased the risk of asthma. In contrast, a mother who never smoked and more siblings were protective of asthma. Always having lived in a rural meshblock was also protective of asthma although not at conventional significance levels (Table 2). Finally, neither road density nor mean annual nitrogen dioxide concentration were significantly associated with asthma.

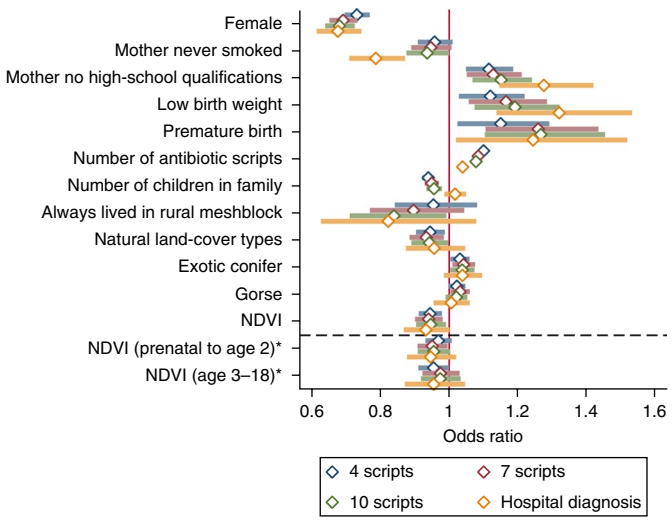

Fig. 2 | Odds ratio plots of asthma prevalence among children born in New Zealand in 1998 (39,108). Exposure to greenness represented as either mean lifetime NDVI (above dotted line) or split into early- and late-life exposure (below the dotted line). Following IDI protocols, all sample sizes have been rounded to the nearest multiple of three.*For clarity, only the coefficients on NDVI are shown.

To ensure that the results were not an artefact of our asthma definition, we repeated the analysis using three other definitions (Fig. 1; for scaling reasons, we excluded ethnicity from the figure). With the exception of the hospital-only definition of asthma, the associations between asthma, NDVI, number of land-cover types, gorse and exotic conifers were not sensitive to asthma definition.

Splitting greenness exposure into early (prenatal to age 2) and later-life exposure (age 2+) reduced the magnitude and significance of the association between NDVI and asthma, but both NDVI measures remained significant suggesting that greenness may be protective of asthma throughout a child's life course (Fig. 2). The association between asthma and number of land-cover types, gorse and exotic conifers became statistically insignificant when stratified into early- and late-life exposure, most likely because of a reduced number of children in each stratum and the collinearity of early- and late-life exposures.

Figure 3 shows associations among asthma, mean lifetime NDVI and mean lifetime number of natural land-cover types for different strata of the sample (asthma was defined as seven or more prescriptions or hospital diagnosis). Both NDVI and the number of land-cover types were more protective for children who had always lived in an urban area. Note that children who always lived in a rural area had significantly higher lifetime exposure to greenness (Fig. 4). NDVI was also more protective when the analysis was restricted to children living in more socially deprived neighbourhoods (mean

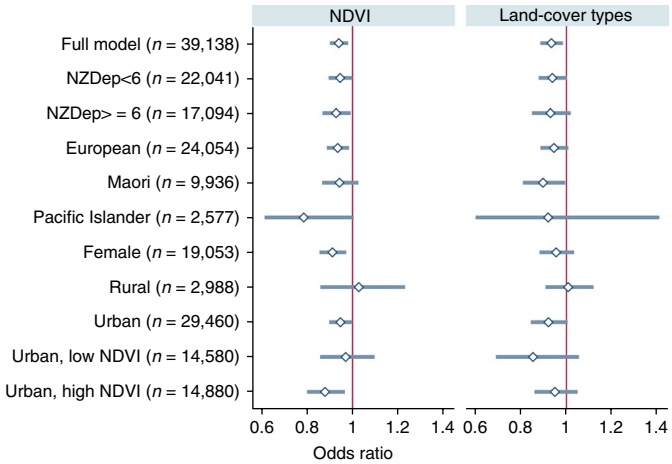

Fig. 3 | Odds ratio plot of asthma prevalence in different strata of children born in New Zealand in 1998. Asthma defined as seven or more asthma scripts or hospital diagnosis. Following IDI protocols, all sample sizes have been rounded to the nearest multiple of three.

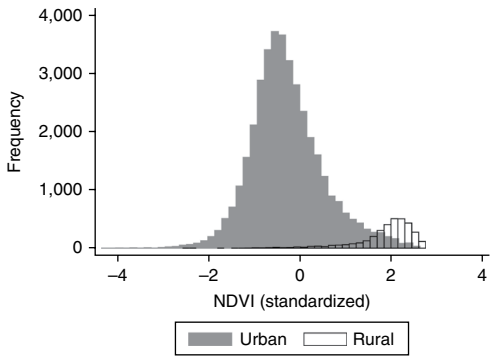

Fig. 4 | Standardized, mean lifetime NDVI for children who always lived in a rural meshblock ($n = 3{,}420$) versus children who always lived in an urban meshblock ($n = 38{,}229$). Following IDI protocols, all sample sizes have been rounded to the nearest multiple of three.

lifetime NZDep score ≥6). The protective effect of NDVI was stronger in the high-NDVI group than the low-NDVI group. In contrast, the protective effect of number of land-cover types was higher in the low-NDVI group (Fig. 3).

Discussion

In a large cohort of children born in 1998, we found that exposure to greenness was associated with a significantly lower risk of developing asthma and exposure to a greater number of natural land-cover types provided an additional increment of protection. To our knowledge, this is the first longitudinal study showing that vegetation diversity can protect against a specific adverse health outcome, and supports recent findings of an ecological study in Australia[19] that biodiversity may protect against respiratory disease. In contrast, exposure to gorse and exotic conifers was a risk factor for asthma. The effects on asthma of gender, ethnicity, education, smoking and adverse birth outcomes were consistent with previous research.

The reasons for the observed protective effects of exposure to greenness and vegetation diversity are unclear, but we found no evidence that the relationship is mediated by reduced air pollution. Instead, we hypothesize that the natural environment may protect against asthma through greater and more diverse microbial exposure. Past research supports a mediating role for microbial biodiversity. Specifically, children living in more biodiverse areas have skin bacteria with greater genetic diversity and a reduced prevalence of atopic sensitization[26]. This is consistent with the hygiene hypothesis, which posits that reduced early-childhood exposure to microbes increases children's susceptibility to allergic diseases such as asthma, through a biased T-helper 2 immune response[27]. Our findings are also consistent with research suggesting that environmental microbial diversity in farm environments may underlie the lower prevalence of asthma in children growing up on a farm[28]. Alternatively, ecological diversity (including plant and microbial species) may itself be protective through currently unknown biological mechanisms.

Number of land-cover types and plant biodiversity are positively correlated but not in a linear or systematic fashion, as plant diversity varies considerably across land-cover types. Nonetheless, our results are suggestive of a link between plant biodiversity and asthma. However, it is unlikely that all plants provide the same protection. Indeed, several plants species are known to provoke asthma in susceptible people. Therefore, future research could valuably focus on identifying which plants provide the greatest protection against childhood asthma.

We found that early- and late-life exposure to greenness were equally protective of asthma, which suggests that the relationship between asthma and the natural environment is not solely mediated by the hygiene hypothesis (the hygiene hypothesis is often associated with early-life exposure). However, although early life is a critical period for many environmental exposures (including those associated with the hygiene hypothesis)[29], later-life exposures can also protect against asthma[30]. Other factors that may mediate the relationship between the natural environment and asthma include reduced stress[31] and increased physical activity[32], although this could not be assessed from our data. In addition, increased greenness and vegetation diversity may encourage children to spend more time outside, which would increase microbial exposure and also reduce exposure to indoor allergens. Also, it is possible that a more complex outdoor ecosystem may increase the diversity of the indoor microbiota, which in turn may protect against childhood asthma.

Other results support a role for the hygiene hypothesis in mediating the relationship between the natural environment and childhood asthma. Specifically, we found that having more siblings was protective of asthma and past research has shown that more siblings increase a child's microbial exposure[33]. Similarly, we found that antibiotic use was a risk factor for asthma, and past research has shown that antibiotics reduce gut microbial diversity, which has been shown to be associated with allergic sensitization and asthma[34]. However, the association between antibiotic use and asthma may also be due to the infections (particularly respiratory infections) that prompted antibiotic use[35].

Exposure to gorse and exotic conifers was a risk factor for childhood asthma, which may be due to the low biodiversity of these land-cover types. Radiata pine (*Pinus radiata*), the most common exotic conifer in New Zealand, is the country's dominant commercial tree species. Plantations are typically managed as monocultures on short rotations and herbicides are often used to manage competing vegetation, which further reduces biodiversity. Gorse (*Ulex europaeus*), in contrast, is not a commercially valuable species. Indeed, gorse can significantly reduce the productivity of New Zealand agricultural land, because it is highly evolved to aggressively colonize recently disturbed landscapes. Native New Zealand plants did not evolve in competition with gorse, so it has rapidly spread often reducing the biodiversity of affected landscapes[36]. The negative effect of gorse on biodiversity is consistent with the effect of other successful invasive plants on the biodiversity of native landscapes[37].

There may be other reasons why both species are risk factors for asthma. In particular, radiata pine pollens, as well as several other species in the *Pinus* genus, are a well-documented source of allergen exposure in New Zealand that can provoke respiratory symptoms in sensitized individuals (less is known about the allergic potential of gorse)[38]. Results suggest that the allergic potential of these two low-biodiversity land-cover types outweigh any protective effect they may offer. This may be because the protective effects of these landscapes are unusually low or, alternatively, gorse and radiata pine may provoke a more potent allergic response than other land-cover types. It is notable that other low-biodiversity natural landscapes (exotic grassland, for example) were not risk factors for asthma (results not shown).

Previous studies have also found that non-native vegetation may be a risk factor for asthma. For example, children who lived close to a park in Sabadell, Spain, were at greater risk of developing childhood asthma[18]. The authors note that parks in Catalonia have more exotic plants than native landscapes and that these exotic plants may be responsible for the higher rates of asthma near parks. Perhaps the particular exotic plant species used in parks in Catalonia are more allergenic. Alternatively, it is possible that some of the parks in Catalonia—sports playing fields, for example—are less biodiverse, and this lower biodiversity is at least partly responsible for the increased asthma risk rather than the higher proportion of exotic plants, per se. However, low biodiversity is unlikely to be the sole reason for the observed increased asthma risk, as some parks have higher biodiversity than surrounding urban areas[39].

In our stratified analyses, NDVI and number of land-cover types were generally protective of asthma, although there was some (inevitable) loss of significance. Both were more protective, when we restricted our analysis to children who had always lived in an urban area, which suggests that greenness and vegetation diversity may provide greater protection to urban children, and more greenness is not simply a proxy for living in a rural area. NDVI was also more protective for children who lived in more socially deprived neighbourhoods. This finding is consistent with past research that found a stronger association between greenness and health among people with lower socio-economic status[4,40]. In addition, NDVI was more protective in meshblocks with above average NDVI, which is consistent with previous research showing that greenness was only protective of health above a minimum threshold[6]. In contrast, number of land-cover types was most protective in areas with below average NDVI. This suggests that, in low-greenness areas, vegetation diversity may act as a substitute for greenness. Therefore, in densely populated areas, where there are limited options for increasing overall greenness, it may be possible to increase the protective effect of the natural environment by increasing vegetation diversity.

This is a longitudinal study of the relationship between the natural environment and asthma. Significant strengths include the large number of study subjects followed from birth, the multiple measures of exposure to the natural environment used and the ability to assess this relationship at different life stages. Previous (cross-sectional) studies were often relatively small and used only one or two coarse measures of exposure to the natural environment, which may not adequately capture the complex relationship between the natural environment and public-health outcomes. This is especially true, if some elements of the natural environment protect against a health outcome, whereas others may be risk factors. For example, we found that always living in a rural area was protective of asthma when we did not control for vegetation diversity. However, the association lost significance when we included the number of natural land-cover types in the model, which suggests that rural living is, at least in part, a proxy for vegetation diversity. Similarly, the association between asthma and the number of land-cover types increased in magnitude and significance when we controlled for gorse and exotic conifers. The use of coarse exposure measures may, therefore, partially explain why previous studies have not found a consistent relationship between the natural environment and asthma. Our study also has several limitations. This is an observational study, so we were not able to show a causal link between greenness and asthma. In addition, our exposures were based on residential meshblocks rather than children's addresses. For small urban meshblocks, meshblock-level greenness may be a reasonable approximation of residential exposure, but in rural areas, our use of meshblock-level exposure is likely to have introduced non-trivial measurement errors. Finally, our definitions of asthma were based on prescriptions and hospital diagnoses and, therefore, may be influenced by access to healthcare and cultural issues.

In conclusion, this large longitudinal birth cohort study has shown, for the first time, that exposure to greenness and vegetation diversity may be protective of asthma. If true, and when biological mechanisms are better understood, this would help inform the design of public-health interventions to reduce the global burden of childhood asthma.

Methods

Study population. We identified our study population and obtained data on outcomes and covariates from Statistics New Zealand's Integrated Data Infrastructure (IDI), which is a collection of national individual-level datasets joined by a common ID. Currently, it holds 166 billion pieces of information in datasets describing health, education and benefits, criminal justice, population (births, deaths and immigration), income and work, and housing (social-housing use and tenancy bond data)[25]. The IDI has a central spine to which individual datasets (hospital discharges, for example) are linked. This spine is based on data from the Inland Revenue Department, births and visa applications and covers almost 95% of New Zealand residents[41]. Individual datasets have been added to IDI over time, so the coverage of the data varies. The datasets we used had the following temporal range: births and deaths (1840 to present), census (2013), pharmaceutical (2005 to present), hospital discharges (1988 to present) and immigration (1997 to present).

Our IDI-based sample consisted of all children born in New Zealand in 1998 ($n = 57{,}451$). After removing children without an address history, multiple births and children who died or emigrated before age 18, the sample was reduced to 49,956.

This study was assessed as low risk by Massey University Ethics Committee (#4000017243) and was approved by Statistics New Zealand (MAA2017-11).

Outcomes. We identified children with asthma using pharmacy and hospital-discharge records. We defined asthma on the basis of two criteria: (1) a diagnosis of asthma in hospital (ICD-10 codes J45 or J46); or (2) having received seven or more prescriptions for inhaled corticosteroids or an inhaled beta-adrenoceptor agonist between 2005 and 2016. We chose seven or more prescriptions, as this corresponded to an asthma prevalence rate of 14.4%, which closely matches the 15.1% prevalence of childhood asthma in New Zealand[42]. To assess whether results were sensitive to our definition of asthma, we used three other definitions in our analysis: four or more asthma prescription or hospital diagnosis (prevalence 19.0%), ten or more asthma prescriptions or hospital diagnosis (prevalence 9.4%) and only hospital diagnosis (prevalence 5.1%).

Residential history. IDI provides residential history for each individual based on residential meshblock, which is the smallest geographic unit for which Statistics New Zealand reports data. On average, 95 people live in a meshblock (46,637 meshblocks nationally in 2013). Residential history is derived from multiple sources and algorithms are used to resolve conflicts based on the reliability of data sources. The accuracy of a person's residential history, therefore, depends upon these algorithms and the frequency with which addresses are updated. Children don't interact as frequently with bureaucratic systems as adults, so we used the residential history of a child's mother to supplement a child's residential history, if both shared the same address. Specifically, we assumed that mother and child lived together, if, for years when residential history was available for both, their residential meshblock matched more than half the time. We also used the mother's residential history to determine the child's prenatal meshblock. Finally, we classified each meshblock as either rural or urban based on definitions from the 2013 Census. The mean area for an urban meshblock is 20.4 hectares and the mean area for a rural meshblock is 2,498 hectares. The mean area of an urban meshblock is equal to the area of a circle with a 255 m radius, which is equivalent, or somewhat smaller, than the buffers used in previous studies of asthma and the natural environment[15,18,43].

Environmental exposures. We used two external data sources to describe exposure to the natural environment at the meshblock level, which we linked to

IDI data. The NDVI is a greenness index (bounded by −1 and 1) derived from satellite imagery. Specifically, we used maximum annual NDVI (1998–2016) derived from Landsat imagery archives (30 m resolution)[44]. The NDVI was calculated at the top of the atmosphere, so all atmospheric effects were normalized (the atmosphere-induced noise embedded in the satellite signal has been largely removed). Using these data, we calculated the mean NDVI for each meshblock for each year from 1998 to 2016. We also used the 2012 Land-Cover Database, produced by Landcare Research New Zealand, which categorizes all land cover in New Zealand into one of 35 classes (Supplementary Table 1). We calculated the proportion of each meshblock covered by each of the land-cover classes. As measures of vegetation diversity, we calculated the total number of natural land-cover types present in each meshblock, the number of native land-cover types present in each meshblock and the number of non-native land-cover types present in each meshblock (Supplementary Table 1). In contrast to NDVI, we only had access to land-cover data from 2001, 2008 and 2012. However, the classification schemes were not consistent across the three years. In addition, when we compared 2008 and 2012 data, we found that the net area of New Zealand that changed from one land class to another was only 0.903%. Therefore, we used 2012 data to represent land cover throughout the study period. Finally, to ensure that exposure measures were not confounded by meshblock size, we included meshblock area in regression models.

We used two measures of air pollution. First, to estimate exposure to traffic-related air pollution, we calculated the length of major roads in each meshblock using data from Land Information New Zealand[45]. We first calculated the length of roads in each meshblock and then converted from a linear to a planar and finally to a raster representation relying on the number of lanes present in each road segment and an estimate of the mean lane width (3.5 m). The latter was obtained by randomly selecting 500 locations on the road network, stratified across lane number classes, where width was measured manually using overlays with high-resolution imagery from Google Maps.

In addition to road density, we used data on mean annual nitrogen dioxide concentration at the meshblock level developed by the National Institute of Water and Atmospheric Research. These nitrogen dioxide estimates are based on the sum of two semi-empirical models: an urban background estimate and a local increment from nearby roads. The urban background model is based on observational data from 41 locations in Greater Wellington, 15 in Hamilton, 3 in Auckland, 2 in Christchurch and 1 site each from a further 10 towns. The roadside increment model is based on data from 19 sites alongside major roads in central Auckland and predicts additional nitrogen dioxide as a function of traffic volume weighted by the proportion of that meshblock that lies within the road's influence.

We merged these exposure estimates with children's residential history to create exposure estimates for each year of a child's life, early life (prenatal to age 2), late life (age 2+) and lifetime mean (prenatal to age 18).

Covariates. We used IDI data to control for a range of variables that previous research has shown to be associated with childhood asthma including being born prematurely or low birth weight[46], antibiotic use[47], parental smoking[48], ethnicity[42], birth order and number of siblings[49], and parental occupation[50]. Using birth records, we identified children whose birth weight was below the 15th percentile[51] based on their ethnicity and gender (the reference population was all New Zealand births from 1993 to 2003); children born before 37 weeks' gestational age; ethnicity of children and parents; parents' occupations; parents' ages; birth order; and number of siblings. Using prescription data, we identified the number of asthma and antibiotic prescriptions dispensed to each child from 2005 to 2016. Parental smoking and education level were specified using 2013 Census data. Note that parental education may vary across the study period (2005–2016) but was only measured in 2013.

To account for the social conditions within a child's residential meshblock, we used NZDep, which is a well-validated social deprivation index based on nine census variables[52]. The index ranks meshblocks into deciles from 1 (least deprived) to 10 (most deprived).

Statistical analysis. We used a three-stage modelling approach. First, we estimated a logit model of asthma occurrence at the individual level accounting for a child's overall greenness exposure as measured by the NDVI (model 1). Asthma occurrence was defined using pharmacy and hospital records (see Outcomes). We controlled for roads, air pollution, ethnicity, gender, birth outcomes, parents' occupation, parents' education, parents' smoking status, antibiotic use, number of siblings, meshblock size and birth order. Second, we added variables describing the diversity of a child's greenness exposure (number of natural land-cover types in a child's residential meshblock) (model 2). Third, we added variables describing a child's exposure to specific land-cover types (model 3). To allow for a consistent comparison of regression coefficients, all continuous variables describing the natural environment were standardized by subtracting the mean and dividing by the standard deviation. Although our data were longitudinal, we estimated a cross-sectional logit model. We created average exposures (NDVI, for example) over three different periods of a child's life: prenatal to age 18, prenatal to age 2 and age 2+ to age 18.

Within each modelling stage, we used a backwards stepwise model selection process. Variables were dropped from the model using progressively lower *P*-value thresholds. For groups of correlated variables (those describing parents' education, for example), we used the variable with the lowest *P*-value when individually regressed against asthma. To avoid including highly correlated variables in the same model, we consulted a correlation coefficient matrix of all candidate variables. To avoid more complex patterns of multicollinearity, we calculated the variance inflation factor of all independent variables in simple linear models in which the dependent variable was the number of asthma prescriptions a child received rather than a binary definition of asthma. We systematically reintroduced insignificant variables and retained them, if the coefficients on variables of interest changed by more than 10%[53].

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Code availability. All code used in the analysis is available from the authors

Data availability. The data that support the findings of this study are available from Statistics New Zealand but restrictions apply to the availability of these data, which were used under license for the current study and accessed via a secure data laboratory, and so are not publicly available. Access to these data is, however, available from Statistics New Zealand upon request.

Received: 20 December 2017; Accepted: 10 April 2018;
Published online: 7 May 2018

References

1. Ulrich, R. S. View through a window may influence recovery from surgery. *Science* **224**, 420–421 (1984).
2. James, P., Hart, J. E., Banay, R. F. & Laden, F. Exposure to greenness and mortality in a nationwide prospective cohort study of women. *Environ. Health Perspect.* **124**, 1344–1352 (2016).
3. Mitchell, R. & Popham, F. Effects of exposure to natural environment on health inequalities: an observational population study. *Lancet* **372**, 1655–1660 (2008).
4. Dadvand, P. et al. Green space, health inequality and pregnancy. *Environ. Int.* **44**, 3–30 (2012).
5. Donovan, G. H., Michael, Y. L., Butry, D. T., Sullivan, A. D. & Chase, J. M. Urban trees and the risk of poor birth outcomes. *Health Place* **17**, 390–393 (2011).
6. Hystad, P. Residential greenness and birth outcomes: evaluating the influence of spatially correlated built-environment factors. *Environ. Health Perspect.* **122**, 1095–102, (2014).
7. Donovan, G. H. et al. The relationship between trees and human health: evidence from the spread of the emerald ash borer. *Am. J. Prev. Med.* **44**, 139–145 (2013).
8. Alcock, I., White, M. P., Wheeler, B. W., Fleming, L. E. & Depledge, M. H. Longitudinal effects on mental health of moving to greener and less green urban areas. *Environ. Sci. Technol.* **48**, 1247–1255 (2014).
9. Nutsford, D., Pearson, A. L. & Kingham, S. An ecological study investigating the association between access to urban green space and mental health. *Public Health* **127,** 1005–1011 2013).
10. Pope, C. A. III & Dockery, D. W. Acute health effects of PM10 pollution on symptomatic and asymptomatic children. *Am. Rev. Respir. Dis.* **145**, 1123–1128 (1992).
11. Grandjean, P. & Landrigan, P. J. Developmental neurotoxicity of industrial chemicals. *Lancet* **368**, 2167–2178 (2006).
12. Lambert, K. A.et al. Residential greenness and allergic respiratory diseases in children and adolescents – a systematic review and meta-analysis . *Environ. Res.* **159,** 212–221 2017).
13. Asher, I. & Pearce, N. Global burden of asthma among children. *Int. J. Tuberc. lung Dis.* **18**, 1269–1278 (2014).
14. Lovasi, G. S., Quinn, J. W., Neckerman, K. M., Perzanowski, M. S. & Rundle, A. Children living in areas with more street trees have lower prevalence of asthma. *J. Epidemiol. Commun. H.* **62**, 647–649 (2008).
15. Lovasi, G. S. et al. Urban tree canopy and asthma, wheeze, rhinitis, and allergic sensitization to tree pollen in a New York City birth cohort. *Environ. Health Perspect.* **121**, 494–500 (2013).
16. Sbihi, H., Tamburic, L., Koehoorn, M. & Brauer, M. Greenness and incident childhood asthma: a 10-year follow-up in a population-based birth cohort. *Am. J. Respir. Crit. Care Med.* **192**, 1131–1133 (2015).
17. Pilat, M. A. et al. The effect of tree cover and vegetation on incidence of childhood asthma in metropolitan statistical areas of Texas. *HortTechnology* **22**, 631–637 (2012).
18. Dadvand, P. et al. Risks and benefits of green spaces for children: a cross-sectional study of associations with sedentary behavior, obesity, asthma, and allergy. *Environ. Health Perspect.* **122**, 1329–1335 (2014).

19. Liddicoat, C. et al. Landscape biodiversity correlates with respiratory health in Australia. *J. Environ. Manag.* **206**, 113–122 (2018).
20. Brokamp, C., LeMasters, G. K. & Ryan, P. H. Residential mobility impacts exposure assessment and community socioeconomic characteristics in longitudinal epidemiology studies. *J. Expo. Sci. Env. Epid.* **26**, 428 (2016).
21. Nowak, D. J., Crane, D. E. & Stevens, J. C. Air pollution removal by urban trees and shrubs in the United States. *Urban For. Urban Gree.* **4**, 155–123 (2006).
22. Ward Thompson, C. et al. More green space is linked to less stress in deprived communities: evidence from salivary cortisol patterns. *Landscape Urban Plan.* **105**, 221–229 (2012).
23. Jones, B. A. Work more and play less? Time use impacts of changing ecosystem services: the case of the invasive emerald ash borer. *Ecol. Econ.* **124,** 49–58 2016).
24. Rook, G. A. Regulation of the immune system by biodiversity from the natural environment: an ecosystem service essential to health. *Proc. Natl Acad. Sci. USA* **110**, 18360–18367 (2013).
25. *Integrated Data Infrastructure* (Statistics New Zealand, 2017); http://www.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure.aspx
26. Hanski, I. et al. Environmental biodiversity, human microbiota, and allergy are interrelated. *Proc. Natl Acad. Sci. USA* **109**, 8334–8339 (2012).
27. Yazdanbakhsh, M., Kremsner, P. G. & Van Ree, R. Allergy, parasites, and the hygiene hypothesis. *Science* **296**, 490–494 (2002).
28. Ege, M. J. et al. Exposure to environmental microorganisms and childhood asthma. *N. Engl. J. Med.* **364**, 701–709 (2011).
29. Douwes, J. et al. Farm exposure in utero may protect against asthma, hay fever and eczema. *Eur. Respir. J.* **32**, 603–611 (2008).
30. Douwes, J. et al. Lifelong farm exposure may strongly reduce the risk of asthma in adults. *Allergy* **62**, 1158–1165 (2007).
31. Takayama, N. et al. Emotional, restorative and vitalizing effects of forest and urban environments at four sites in Japan. *Int. J. Environ. Res. Public Health* **11**, 7207–7230 (2014).
32. Fan, Y., Das, K. V. & Chen, Q. Neighborhood green, social support, physical activity, and stress: assessing the cumulative impact. *Health Places* **17,** 1202–1211 2011).
33. Aniansson, G. et al. Nasopharyngeal colonization during the first year of life. *J. Infect. Dis.* **165**, S38–S41 (1992).
34. Jakobsson, H. E. et al. Short-term antibiotic treatment has differing long-term impacts on the human throat and gut microbiome. *PloS ONE* **5**, e9836 (2010).
35. Johnston, S. L. et al. The relationship between upper respiratory infections and hospital admissions for asthma: a time-trend analysis. *Am. J. Respir. Crit. Care Med.* **154**, 654–660 (1996).
36. Norton, D. A. Species invasions and the limits to restoration: learning from the New Zealand experience. *Science* **325**, 569–571 (2009).
37. Vila, M. et al. Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. *Ecol. Lett.* **14**, 702–708 (2011).
38. Fountain, D. W. & Cornford, C. A. Aerobiology and allergenicity of *Pinus radiata* pollen in New Zealand. *Grana.* **30**, 71–75 (1991).
39. Nielsen, A. B., van den Bosch, M., Maruthaveeran, S. & van den Bosch, C. K. Species richness in urban parks and its drivers: a review of empirical evidence. *Urban Ecosyst.* **17**, 305–327 (2013).
40. Markevych, I. et al. Surrounding greenness and birth weight: results from the GINIplus and LISAplus birth cohorts in Munich. *Health Place* **26**, 39–46 (2014).
41. Black, A. *The IDI Prototype Spine's Creation and Coverage*. Statistics New Zealand Working Paper No 16–03 (Statistics New Zealand, Wellington, 2016).
42. New Zealand Ministry of Health. *Annual Update of Key Results 2014/15: New Zealand Health Survey* (New Zealand Ministry of Health, Wellington, 2015).
43. Fuertes, E. et al. Residential greenness is differentially associated with childhood allergic rhinitis and aeroallergen sensitization in seven birth cohorts. *Allergy* **71**, 1461–1471 (2016).
44. Chander, G., Markham, B. L. & Helder, D. L. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sens. Environ.* **113**, 893–903 (2009).
45. *NZ Roads (Addressing) Set* (Land Information New Zealand, 2016); https://data.linz.govt.nz/set/88
46. Sonnenschein-van der Voort, A. M. et al. Preterm birth, infant weight gain, and childhood asthma risk: a meta-analysis of 147,000 European children. *J. Allergy Clin. Immunol.* **133**, 1317–1329 (2014).
47. Aboal, J. R., Fernandez, J. A., Boquete, T. & Carballeira, A. Is it possible to estimate atmospheric deposition of heavy metals by analysis of terrestrial mosses?. *Sci. Total. Environ.* **408,** 6291–6297 2010).
48. Strachan, D. P. & Cook, D. G. Parental smoking and childhood asthma: longitudinal and case-control studies. *Thorax* **53**, 204–212 (1998).
49. Goldberg, S. et al. Asthma prevalence, family size, and birth order. *Chest* **131**, 1747–1752 (2007).
50. Tagiyeva, N. et al. Parental occupation is a risk factor for childhood wheeze and asthma. *Eur. Respir. J.* **35**, 987–993 (2010).
51. Xu, H., Simonet, F. & Luo, Z. C. Optimal birth weight percentile cut-offs in defining small- or large-for-gestational-age. *Acta. Paediatr.* **99**, 550–555 (2010).
52. Salmond, C. E. & Crampton, P. Development of New Zealand's deprivation index (NZDep) and its Uptake as a national policy tool. *Can. J. Public Health* **103**, S7–S11 (2012).
53. Rothman, K. J., Greenland, S. & Lash, T. L. *Modern Epidemiology*. *Modern Epidemiology*. 3rd edn. (Wolters Kluwer, Philadelphia, 2008).

Acknowledgements

Thanks to C. Brooks for his valuable insights on mechanisms linking the natural environment and asthma. The results in this paper are not official statistics. They have been created for research purposes from the IDI, managed by Statistics NZ. The opinions, findings, recommendations and conclusions expressed in this paper are those of the authors, not Statistics NZ. Access to the anonymized data used in this study was provided by Statistics NZ under the security and confidentiality provisions of the Statistics Act 1975. Only people authorized by the Statistics Act 1975 are allowed to see data about a particular person, household, business or organization, and the results in this paper have been confidentialized to protect these groups from identification and to keep their data safe. Careful consideration has been given to the privacy, security and confidentiality issues associated with using administrative and survey data in the IDI. Further detail can be found in the privacy impact assessment for the IDI available from www.stats.govt.nz.

Author contributions

G.H.D. developed the research idea with J.D., conducted the analysis and took the lead on writing the paper. D.G. conducted all geo-spatial analysis and wrote the geo-spatial section of the methods. I.L. provided air quality data, wrote the air-quality section of the methods and reviewed drafts of the paper. J.D. developed the research idea with G.H.D., wrote significant parts of the paper and reviewed multiple drafts of the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41477-018-0151-8.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to G.H.D.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Geoffrey H. Donovan

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study.
For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

Experimental design

1. Sample size

Describe how sample size was determined.

We selected all children born in New Zealand in 1998. We picked 1998, because it was the youngest birth cohort that we could track until adulthood using our data.

2. Data exclusions

Describe any data exclusions.

We excluded multiple births, because we used birth outcomes as a covariate. In addition, we excluded anyone without an address history, because without an address history, we weren't able to determine exposure metrics.

3. Replication

Describe the measures taken to verify the reproducibility of the experimental findings.

This is an observational study, so we didn't reproduce findings.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

This is an observational study, so there was no assignment

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

This is an observational study, so there was no blinding

Note: all in vivo studies must report how sample size was determined and whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☒	☐	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	A statement indicating how many times each experiment was replicated
☐	☒	The statistical test(s) used and whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☐	Test values indicating whether an effect is present *Provide confidence intervals or give results of significance tests (e.g. P values) as exact values whenever appropriate and with effect sizes noted.*
☐	☒	A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒	Clearly defined error bars in all relevant figure captions (with explicit mention of central tendency and variation)

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	Microsoft SQL studio and Stata 14

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a third party.	n/a

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	n/a

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	n/a
b. Describe the method of cell line authentication used.	n/a
c. Report whether the cell lines were tested for mycoplasma contamination.	n/a
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	n/a

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide all relevant details on animals and/or animal-derived materials used in the study.	n/a

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	This study used routinely collected data. See table 1 for descriptive statistics of the sample.