Received: 14 June 2021 | Revised: 21 August 2021 | Accepted: 30 August 2021
DOI: 10.1111/pai.13662

LETTER TO THE EDITOR

WILEY

# Childhood asthma and land-use characteristics in school and residential neighborhoods: A decision tree learning approach

To the Editor,

Growing urbanization imposes a huge pressure on environment and physical landscape. With hasty urbanization, there is an increasing interest in understanding how urban settings and environment affect children's health.[1] The number of studies on the effect of specific characteristics of child's school and residential neighborhood on asthma and allergy is increasing.[2–4] Yet, given that children spend large periods of time at home and at school, whether and how different land-use characteristics of both neighborhoods are related to respiratory health, including asthma, in school-aged children, is unclear. Thus, we aimed to identify the land-use characteristics in school and residential surrounding associated with asthma prevalence in school-aged children using a decision tree learning approach.

Data on demographic and health outcomes from 858 children (aged 7–12 years) were obtained from a cross-sectional study performed in Porto, Portugal (supporting information and Table S1). We did a cross-sectional study with an ecological research design. Land-use characteristics within a 500 m buffer centered in each child's school ($n$ = 20) and residential ($n$ = 174) address were assessed based on the 2015 Portuguese official maps of land cover using a Geographical Information System (ArcMap 10.7.2) (Table S2). For residences without information on land-use characteristics, data from each child's school were imputed. The mean imputation method was used for residences without information on land-use characteristics, and the missing values are replaced with the mean of the same land-use type for each child's school. The analysis of land-use characteristics was aggregated at school level. Regression tree models were used to identity the land-use characteristics in school and residential neighborhoods associated with the prevalence of asthma among children. Linear regression models were fitted to estimate the association between each branch and asthma prevalence using the split values. The Ethical Committee of the University of Porto approved the study, and informed consent was obtained from children's legal guardians.

The results demonstrated that three land-use characteristics around schools and residences had a significant effect in the prevalence of asthma in schoolchildren (Figure 1). Schools with children living in residential neighborhoods surrounded by an area of agriculture, natural, and semi-natural spaces smaller than 0.77 ha, and those surrounded by an area of continuous urban fabric smaller than 65.3 ha had a higher prevalence of asthma (11%) compared to schools with children living in residential neighborhoods surrounded by an area of agriculture, natural, and semi-natural spaces larger than or equal to 0.77 ha and those surrounded by an area of cultural facilities and larger than or equal to 0.67 ha (4.7%).

Compared to the branch with larger areas of agriculture, natural, and semi-natural spaces in residential's neighborhood and cultural facilities and historic areas around schools, the branch with smaller areas of agriculture, natural, and semi-natural spaces in residential's neighborhood and continuous urban fabric around schools was associated with a higher prevalence of asthma [$\beta$ (95% CI) = 0.068 (0.024; 0.112)] (Table 1).

Our findings suggested that some characteristics of land use around schools and residences may be determinants of a higher prevalence of asthma in school-aged children. Smaller areas of agriculture, natural, and semi-natural spaces around residences, and the smaller presence of continuous urban fabric in school's neighborhood were significantly associated with a higher prevalence of asthma compared with larger areas of agriculture, natural, and semi-natural spaces in residential's neighborhood and cultural facilities and historic areas around schools. Additionally, children living in neighborhoods surrounded by larger areas of agriculture, (semi)-natural, and cultural and historic facilities had a lower prevalence of asthma. The presence of natural land-use areas may enhance environmental biodiversity in children's neighborhood, affecting the composition of their microbiota, which in turn may decrease the risk of asthma.[5] In addition, the presence of natural spaces may also provide a protection from anthropogenic air pollution through absorption, providing physical barriers against emission sources, or by limiting the overall area available to sources of pollution such as traffic and industry.[6] Previous studies have also evaluated the health impacts of specific land cover types, as agricultural areas and forests,[5,7] corroborating our results and highlighting the influence of land cover use on respiratory health in children. Additionally, the presence of cultural and historic facilities and continuous urban fabric areas in our study was related to a lower presence of transit roads in school's neighborhood (data not shown), and possibly to a lower exposure to traffic-related air pollution, decreasing the risk of asthma development among children.[8,9] Growing evidence supports the link between air pollution and the traffic-related pollution exposure from roadways near schools and the risk of asthma in children.[10–12] In 2020, Zeng et al.[13] also suggested a protective

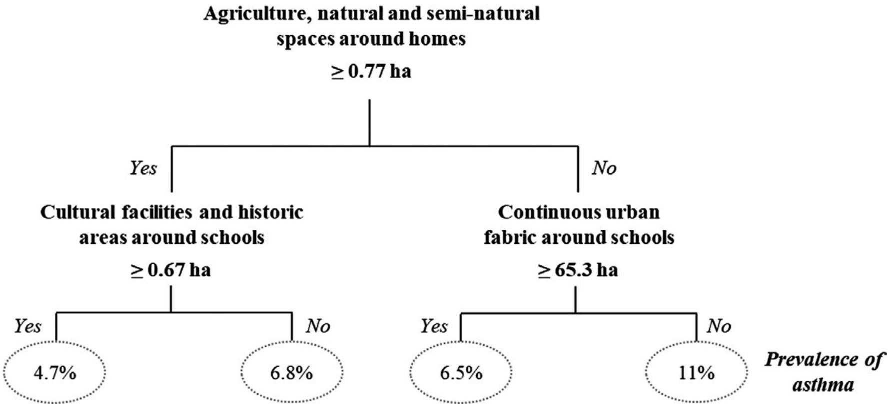


**FIGURE 1** Decision tree reporting the prevalence of asthma considering the land-use characteristics in school and residential surrounding

**TABLE 1** Association (β, 95% Confidence Intervals, 95% CI) between tree branches and asthma

| Tree branches | β (95% CI) |
|---|---|
| Branch 1 | Reference |
| Branch 2 | 0.021 (−0.032; 0.074) |
| Branch 3 | 0.018 (−0.040; 0.076) |
| Branch 4 | 0.068 (0.024; 0.112) |

*Note:* Branch 1: Agriculture, natural, and semi-natural spaces around residences ≥0.77 ha and cultural facilities and historic areas around schools ≥0.67 ha; Branch 2: Agriculture, natural, and semi-natural spaces around residences ≥0.77 ha and cultural facilities and historic areas around schools <0.67 ha; Branch 3: Agriculture, natural, and semi-natural spaces around residences <0.77 ha and continuous urban fabric around school ≥65.3 ha; Branch 4: Agriculture, natural, and semi-natural spaces around residences <0.77 ha and continuous urban fabric around school <65.3 ha.

effect of greenness in school's neighborhoods on asthma, particularly in higher air pollution areas. Our study has some limitations. As a cross-sectional ecological study, we cannot assume that associations observed at the aggregate level are necessarily reflections of associations at individual level. We did not access the levels of outdoor air pollution. However, different studies on urban environmental effects reported that land use could be used as an indicator of urban-related air pollution, such as traffic, without outdoor air monitoring.[14,15] Rosenlund et al.[16] also found a reasonable agreement between land-use and traffic emissions. In addition, we did not measure the effect of other important factors, such as socioeconomic conditions, family history of asthma and allergic disease, day-care attendance, or exposure to pets, which may impact the results. Although the mean imputation method may result in an underestimation of variability and affect the strength of relationships, this method is usually used for missing data [17–19]and the schools were usually located in the residence's neighborhoods, minimizing the bias effect of mean imputation. Moreover, Maheswari et al.[20] showed that there is a good improvement in the accuracy of classification algorithm after imputing mean value in the dataset. The results may not be representative of the children in the Porto metropolitan area and may not be generalizable in a different geographical context. The small sample size may also be a limitation, since at each time the decision tree splits the data using a land-use characteristic, the remaining sample size reduces, and may limit the identification of further predictors that are still relevant. However, relatively small samples will often produce models whose accuracy approaches that of an unlimited sample size.[21] To our knowledge, this is the first study to identity the land-use characteristics in school and residential surrounding associated with asthma among children. In addition, the use of land cover maps, such as the Portuguese official maps of land cover, which is suitable for analyzing land-use dynamics at different levels (local, regional, national), allowed to evaluate the role of several land-use characteristics in schools and residences' neighborhood on asthma development in children. Our results indicate that larger of agriculture, (semi)-natural spaces around residences, and cultural and historic facilities in school's neighborhoods may be associated with a lower prevalence of asthma among school-age children. Our findings also provide further evidence for the need to assess the influence of land use around schools and residences on asthma and to develop sustainable and healthy urban planning for children.

## KEYWORDS

asthma, children, environment, regression tree models, respiratory disease predictors

## FUNDING INFORMATION

European Regional Development Fund; Foundation for Science and Technology, Grant/Award Number: POCI-01-0145-FEDER-03019302/SAICT/2017; FEDER; FCT, Grant/Award Number: UIDB/04750/2020; Programa Operacional de Saúde—Saúde XXI; Quadro Comunitário de Apoio III and Administração Regional de Saúde Norte; National Funds, Grant/Award Number: CEECIND/02386/2018

## ACKNOWLEDGEMENTS

Authors gratefully acknowledge the Project EXALAR 21 funded by the European Regional Development Fund (FEDER), through the Competitiveness and Internationalization Operational Programme, and by national funding from the Foundation for Science and Technology (FCT) under the scope of the project PTDC/GES-AMB/30193/2017 (POCI-01-0145-FEDER-030193, 02/SAICT/2017—Project nº 30193). This study was also funded by FEDER through the Operational Programme Competitiveness and Internationalization and national funding from the Foundation for Science and Technology—FCT (Portuguese Ministry of Science, Technology and Higher Education) under the Unidade de Investigação em Epidemiologia—Instituto de Saúde Pública da Universidade do Porto (EPIUnit) (UIDB/04750/2020). G21 was funded by Programa Operacional de Saúde—Saúde XXI, Quadro Comunitário de Apoio III and Administração Regional de Saúde Norte (Regional Department of Ministry of Health). Ana Isabel Ribeiro was supported by National Funds through FCT, under the program of "Stimulus of Scientific Employment—Individual Support" within the contract CEECIND/02386/2018. João Cavaleiro Rufo was supported by National Funds through FCT, under the program of "Stimulus of Scientific Employment—Individual Support" within the contract 2020.01350.CEECIND.

## CONFLICTS OF INTEREST

Authors declare no conflicts of interest.

## AUTHOR CONTRIBUTION

**Inês Paciência:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Writing-original draft (lead); Writing-review & editing (equal). **João Cavaleiro Rufo:** Data curation (equal); Formal analysis (equal); Writing-review & editing (equal). **Ana Isabel Ribeiro:** Data curation (equal); Formal analysis (equal); Writing-review & editing (equal). **Milton Severo:** Conceptualization (equal); Formal analysis (equal); Methodology (equal); Writing-review & editing (equal). **André Moreira:** Conceptualization (equal); Funding acquisition (lead); Writing-review & editing (equal).

## PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/pai.13662.

Inês Paciência[1,2,3]
João Cavaleiro Rufo[1,2,3]
Ana Isabel Ribeiro[1,2,4]
Milton Severo[1,2,5]
André Moreira[1,2,3,6]

[1]*EPIUnit, Instituto de Saúde Pública da Universidade do Porto, Porto, Portugal*
[2]*Laboratório para a Investigação Integrativa e Translacional em Saúde Populacional (ITR), Porto, Portugal*
[3]*Departamento de Patologia, Serviço de Imunologia Básica e Clínica, Faculdade de Medicina da Universidade do Porto, Porto, Portugal*
[4]*Faculdade de Medicina da Universidade do Porto, Porto, Portugal*
[5]*Instituto de Ciências Biomédicas Abel Salazar, Porto, Portugal*
[6]*Faculdade de Ciências da Nutrição e Alimentação da Universidade do Porto, Porto, Portugal*

**Correspondence**
Inês Paciência, EPIUnit, Instituto de Saúde Pública da Universidade do Porto, Porto, Portugal.
Email: inespaciencia@gmail.com

**Editor:** Jon Genuneit

## ORCID

*Inês Paciência* https://orcid.org/0000-0001-5762-7814
*João Cavaleiro Rufo* https://orcid.org/0000-0003-1175-242X
*André Moreira* https://orcid.org/0000-0002-7294-9296

## REFERENCES

1. Elmqvist T, Setälä H, Handel SN, et al. Benefits of restoring ecosystem services in urban areas. *Curr Opin Env Sust*. 2015;14:101-108.
2. Paciência I, Moreira A, Moreira C, et al. Neighbourhood green and blue spaces and allergic sensitization in children: a longitudinal study based on repeated measures from the Generation XXI cohort. *Sci Total Environ*. 2021;772:145394.
3. Cavaleiro Rufo J, Ribeiro AI, Paciência IR, Delgado JLD, Moreira AMA. The influence of species richness in primary school surroundings on children lung function and allergic disease development. *Pediatr Allergy Immunol*. 2020;31:358-363.
4. Hartley K, Ryan P, Brokamp C, Gillespie GL. Effect of greenness on asthma in children: a systematic review. *Public Health Nurs*. 2020;37(3):453-460.
5. Ruokolainen L, von Hertzen L, Fyhrquist N, et al. Green areas around homes reduce atopic sensitization in children. *Allergy*. 2015;70(2):195-202.
6. Kumar P, Druckman A, Gallagher J, et al. The nexus between air pollution, green infrastructure and human health. *Environ Inter*. 2019;133:105181.
7. Parmes E, Pesce G, Sabel CE, et al. Influence of residential land cover on childhood allergic and respiratory symptoms and diseases: evidence from 9 European cohorts. *Environ Res*. 2020;183:108953.
8. Khreis H, Ramani T, de Hoogh K, et al. Traffic-related air pollution and the local burden of childhood asthma in Bradford, UK. *Int J Transp Sci Technol*. 2019;8(2):116-128.
9. Son J-Y, Kim H, Bell ML. Does urban land-use increase risk of asthma symptoms? *Environ Res*. 2015;142:309-318.
10. McConnell R, Islam T, Shankardass K, et al. Childhood incident asthma and traffic-related air pollution at home and school. *Environ Health Perspect*. 2010;118(7):1021-1026.
11. Hwang BF, Lee YL, Lin YC, Jaakkola JJK, Guo YL. Traffic related air pollution as a determinant of asthma among Taiwanese school children. *Thorax*. 2005;60(6):467-473.
12. Perez L, Declercq C, Iniguez C, et al. Chronic burden of near-roadway traffic pollution in 10 European cities (APHEKOM network). *Eur Respir J*. 2013;42(3):594-605.
13. Zeng X-W, Lowe AJ, Lodge CJ, et al. Greenness surrounding schools is associated with lower risk of asthma in schoolchildren. *Environ Int*. 2020;143:105967.
14. Ebisu K, Holford TR, Belanger KD, Leaderer BP, Bell ML. Urban land-use and respiratory symptoms in infants. *Environ Res*. 2011;111(5):677-684.
15. Raaschou-Nielsen O, Andersen ZJ, Beelen R, et al. Air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). *Lancet Oncol*. 2013;14(9):813-822.
16. Rosenlund M, Forastiere F, Stafoggia M, et al. Comparison of regression models with land-use and emissions data to predict the

spatial distribution of traffic-related air pollution in Rome. *J Expo Sci Environ Epidemiol*. 2008;18(2):192-199.
17. Baneshi MR, Talei AR. Does the missing data imputation method affect the composition and performance of prognostic models? *Iran Red Crescent Med J*. 2012;14(1):31-36.
18. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. *BMC Med Res Methodol*. 2017;17(1):162.
19. Dziura JD, Post LA, Zhao Q, Fu Z, Peduzzi P. Strategies for dealing with missing data in clinical trials: from design to analysis. *Yale J Biol Med*. 2013;86(3):343-358.
20. Maheswari K, Packia Amutha Priya P, Ramkumar S, Arun M. Missing data handling by mean imputation method and statistical analysis of classification algorithm. In: Haldorai A, Ramu A, Onn CC, Mohanram S, eds. *EAI International conference on big data innovation for sustainable cognitive computing*. Springer; 2020:137-149
21. Morgan J, Dougherty R, Hilchie A, Carey B. Sample size and modeling accuracy with decision tree based data mining tools. *Acad Inf Manag Sci J*. 2003;6:71-99.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.