NOVA
NOVA SCHOOL OF SCIENCE & TECHNOLOGY

DEPARTAMENT OF MATHEMATICS

LARA PINHEIRO GUEDES

Master in Medicine

LIMITATIONS OF LOGISTIC REGRESSIONS TO ESTIMATE MEASURES OF ASSOCIATION FOR BINARY HEALTH OUTCOMES

A HANDS-ON STUDY WITH TWO EPIDEMIOLOGICAL EXAMPLES

MASTER IN STATISTICS FOR HEALTH

NOVA University Lisbon

September 2023

NOVA
NOVA SCHOOL OF
SCIENCE & TECHNOLOGY

DEPARTAMENT OF
MATHEMATICS

LIMITATIONS OF LOGISTIC REGRESSIONS TO ESTIMATE MEASURES OF ASSOCIATION FOR BINARY HEALTH OUTCOMES

A HANDS-ON STUDY WITH TWO EPIDEMIOLOGICAL EXAMPLES

LARA PINHEIRO GUEDES
Master in Medicine

Adviser: Maria do Rosário Oliveira Martins
Full Professor, NOVA University Lisbon

Examination Committee:

Chair: Isabel Natário
Full Professor, NOVA University Lisbon

Rapporteurs:

Adviser: Maria do Rosário Oliveira Martins
Full Professor, NOVA University Lisbon

Members: Sara Simões Dias
Professor, NOVA University Lisbon

MASTER IN STATISTICS FOR HEALTH

NOVA University Lisbon
September 2023

Limitations of logistic regressions to estimate measures of association for binary health outcomes

Copyright © Lara Pinheiro Guedes, NOVA School of Science and Technology, NOVA University Lisbon.

The NOVA School of Science and Technology and the NOVA University Lisbon have the right, perpetual and without geographical boundaries, to file and publish this dissertation through printed copies reproduced on paper or on digital form, or by any other means known or that may be invented, and to disseminate through scientific repositories and admit its copying and distribution for non-commercial, educational or research purposes, as long as credit is given to the author and editor.

This document was created with Microsoft Word text processor and the NOVA thesis Word template.

To my friends, to Lisbon.

ACKNOWLEDGMENTS

To my adviser, Professor Maria do Rosário Martins, who accepted me as her advisee since the first discussion of the thesis topic, and who helped me organize my ideas with pragmatism and sympathy in a record time.

To all the friends I met and re-met in Lisbon (especially, Carlos and Margarida).

To Luísa, Isa, Rita, Marisa, Jamy, and Artur, who always supported me, even from afar.

To my family, who is invincible in being present.

To all the people who made me laugh and have a good time.

“There are three kinds of lies: lies, damned lies, and statistics.” (attributed to the British Prime Minister Benjamin Disraeli).

ABSTRACT

Introduction: In medical literature, the logistic regression is frequently used to estimate measures of association between an exposure, a health determinant or an intervention, and a binary outcome. However, when the outcome is frequent (>10%), model estimates for relative risks and prevalence ratios might be biased, with potential impact on medical decision and policymaking. Despite the availability of several alternatives, many studies still rely on logistic regression models, and a consensus on this matter is yet to be reached. We aimed to compare the estimation and goodness-of-fit of logistic, log-binomial and robust Poisson regression models, in cross-sectional studies involving frequent binary outcomes.

Methodology: Two cross-sectional studies with distinct characteristics and on different topics were conducted to estimate measures of association between an exposure and a frequent binary outcome. Study 1 was a nationally-representative study on the impact of air pollution on mental health. Study 2 was a local study on immigrants' access to urgent health care services. Odds ratios (OR) were obtained through logistic regression models, while prevalence ratios (PR) were obtained through log-binomial and robust Poisson regression models. Confidence intervals (CI), their ranges, and standard-errors (SE) were also computed, along with models' relative goodness-of-fit through Akaike Information Criterion (AIC), when applicable.

Results: In Study 1, the OR (95%IC) was 1.015 (0.970-1.063), while the PR (95%CI) obtained through the robust Poisson mode was 1.012 (0.979-1.045). The log-binomial regression model did not converge in this study. In Study 2, the OR (95%CI) was 1.584 (1.026-2.446), the PR (95%CI) for the log-binomial model was 1.217 (0.978-1.515), and 1.130 (1.013-1.261) for the robust Poisson model. The 95%CI, their ranges, and the SE of the OR were higher than those of the PR, in both studies. However, in Study 2, the AIC value was lower for the logistic regression model, followed by the log-binomial regression and the robust Poisson regression (1.345, 1.350, and 1.656, respectively).

Discussion and conclusions: In the two presented examples, OR overestimated PR, with wider 95%CI and higher EPs. The extent of overestimation was greater as the outcome under study became more prevalent, in line with previous studies. Employing logistic regression models by default might lead to misinterpretations, especially by less experienced researchers. Robust Poisson models are viable alternatives to logistic regression models, in cross-sectional studies with frequent binary outcomes,

avoiding the non-convergence issues of log-binomial models. Nonetheless, in Study 2, the logistic regression was the model with the best fit, which illustrates the need to consider multiple criteria, rather than just one, when selecting the most appropriate statistical model for each study. This dissertation highlights the need for statistical guidelines to support the selection of the most appropriate models and to facilitate the correct reporting, interpretation, and communication of scientific results.

Keywords: logistic models, odds ratio, risk ratio, log-binomial models, robust Poisson models, prevalence ratio, measures of association, health statistics, research methodologies, epidemiology.

RESUMO

Introdução: *Na literatura médica, a regressão logística é frequentemente utilizada para estimar medidas de associação entre uma exposição, um determinante de saúde ou uma intervenção e um desfecho binário. No entanto, quando o desfecho é frequente (>10%), as estimativas destes modelos para o risco relativo ou razão de prevalências (RP) podem ser enviesadas. Apesar de existirem modelos estatísticos alternativos, muitos estudos continuam a aplicar modelos de regressão logística indiscriminadamente. O objetivo desta dissertação foi comparar as estimativas e o ajuste de modelos de regressão logística, log-binomial e* Poisson *robusta, em estudos transversais com desfechos binários frequentes.*

Metodologia: *Elaboraram-se dois estudos transversais com características distintas e sobre diferentes tópicos, de modo a estimar medidas de associação entre uma exposição e um desfecho binário frequente. O Estudo 1 tratou-se de um estudo representativo a nível nacional sobre o impacto da poluição atmosférica na saúde mental. O Estudo 2 tratou-se de um estudo local sobre o acesso de imigrantes a serviços de urgência. Obtiveram-se* odds ratio *(OR) através de modelos de regressão logística e RP através de modelos log-binomiais e* Poisson *robustos. Foram ainda obtidos os intervalos de confiança a 95% (IC95%), suas amplitudes, os erros-padrão (EP) das estimativas e comparados os valores* Akaike Information Criteria *(AIC) entre os modelos elaborados.*

Resultados: *No Estudo 1, a OR (IC95%) foi de 1,015 (0,970-1,063) e a RP (IC 95%) obtida através do modelo de* Poisson *robusto foi de 1,012 (0,979-1,045). O modelo de regressão log-binomial não convergiu. No Estudo 2, a OR (IC95%) foi de 1,584 (1,026-2,446), a RP (IC95%) para o modelo de regressão log-binomial foi de 1,217 (0,978-1,515) e para o modelo de* Poisson *robusto foi de 1,130 (1,013-1,261). Os IC95%, as suas amplitudes e os EP das OR foram superiores ao das RP, em ambos os estudos. No entanto, no Estudo 2, o valor do AIC foi inferior no modelo de regressão logística, seguido pela regressão log-binomial e pela* Poisson *robusta (1,345; 1,350 e 1,656, respetivamente).*

Discussão e conclusões: *Nos dois exemplos apresentados, as OR sobrestimaram as RP, apresentando também IC95% mais amplos e EP superiores. A magnitude da sobrestimação foi tanto maior quanto mais prevalente o desfecho em estudo, em linha com estudos prévios. No entanto, no Estudo 2, a regressão logística foi a que melhor se ajustou aos dados. Este exemplo ilustra a necessidade de avaliar vários critérios, ao invés de apenas um, para a seleção do modelo estatístico mais apropriado*

a cada estudo. Os modelos de Poisson *robustos são uma alternativa viável à regressão logística, em estudos transversais com desfechos binários frequentes e evitam o problema de não convergência dos modelos log-binomiais.*

Palavras-chave: *regressão logística,* odds ratio*, risco relativo, regressão log-binomial, regressão de* Poisson *robusta, razão de prevalências, medidas de associação, estatística aplicada à saúde, metodologias de investigação, epidemiologia.*

CONTENTS

1 INTRODUCTION 1

1.1 Objective 3

2 METHODS 5

2.1 Study 1 - Association Between Long-term Exposure to Ambient Air Pollution and Probable Diagnosis of Depressive and Anxiety Disorders 6

2.2 Study 2 - Association Between Immigration Status and Urgent Care Use in the Paediatric Population 8

3 RESULTS 11

3.1 Study 1 - Association Between Long-term Exposure to Ambient Air Pollution and Probable Diagnosis of Depressive and Anxiety Disorders 11

3.1.1 Regression Models 14

3.2 Study 2 - Association Between Immigration Status and Urgent Care Use in the Paediatric Population 14

3.2.1 Regression Models 16

4 DISCUSSION 18

4.1 Strengths and Limitations 20

4.2 Conclusions 21

List of Tables

Table 1. Assumptions, estimation, and other characteristics of logistic, log-binomial and Poisson regression models ... 2

Table 2. Study 1 models' specification ... 7

Table 3. Study 2 models' specification ... 9

Table 4. Estimates for the characteristics in study of the included individuals ... 11

Table 5. Characteristics of the included individuals (n=2 398) and comparison between groups of probable diagnosis of common mental health disorders ... 12

Table 6. Characteristics of logistic, log-binomial and robust Poisson regression models fitted for the estimation of measures of association between the exposure to particulate matter with an aerodynamic diameter equal than or lower to 10 micrometres and the frequency of common mental disorders ... 14

Table 7. Baseline characteristics of the children and their caregivers, by immigration status ... 15

Table 8. Characteristics of children and their caregivers, by having or not used urgent care in 2019 . 16

Table 9. Characteristics of logistic, log-binomial, and robust Poisson regression models fitted for the estimation of measures of association between immigration status and urgent care use ... 17

ACRONYMS

AIC	Akaike Information Criteria
CI	Confidence interval
CMD	Common mental disorders
CRIAS	Cohort of immigrant children in Amadora
INSEF	Portuguese National Health Examination Survey (from the Portuguese: *Inquérito Nacional de Saúde com Exame Físico*)
IQR	Interquartile range
ISCED	International Standard Classification of Education
ISCO-08	International Standard Classification of Occupations
MHI-5	5-item Mental Health Inventory
OR	Odds ratio
PM_{10}	Particulate matter with an aerodynamic diameter equal than or lower to 10 micrometres
PR	Prevalence ratio
RR	Relative risk

1

INTRODUCTION

In medical research, logistic regression models, a generalized linear models, are frequently used to estimate measures of association between an exposure, treatment or health determinant and a binary outcome.[1, 2] Modelling the logarithm odds of the outcome, these models are employed mostly in cross-sectional and case-control studies to estimate odds ratios (OR) through the maximum-likelihood estimation method.[1] Logistic regression models are very attractive for their simplicity. Their logit link function ensures that model-predicted probabilities fall within the range of 0 to 1, allowing them to accurately estimate binomial variances.[1] Another appealing property of these models, particularly to unexperienced researchers, is reciprocity. It means that changing the reference category of a binary independent variable will result in a reciprocal estimate.[3]

Although, under certain conditions, OR might allow to infer prevalence ratios (PR) and relative risks (RR), in many other, these inferences cannot be done as they introduce bias (Annex 1). This means that the expected values or estimates are different from the population parameter.[1] If the outcome in study is rare (≤10%), logistic regression models may appropriately estimate PR and RR. However, if the outcome is frequent (>10%), OR might not be suitable estimators of PR and RR.[2, 4] Despite that, in a recent veterinary medicine review, 96.5% of the examined articles employed logistic regression models regardless of the outcome's prevalence.[5]

In the case of a frequent outcome, OR tend to overestimate PR and RR when their values are greater than one, or, conversely, to underestimate PR and RR when values are lower than one.[2, 3] A study reported that 40% of the papers that employed logistic regression models estimated OR which deviated by over 20% from the corresponding RR.[6] OR can thus lead to misinterpretations, being frequently mistakenly portrayed as PR and RR, which is reflected in the results section of research papers, where "risk" and "probability" are wrongly used instead of "odds" and "possibility".[2, 4]

While alternative methods to estimate valid adjusted PR and RR, in the presence of frequent binary outcomes, have been suggested, a consensus on the matter is yet to be reached. Among the available options, log-binomial[7-9], and modified Poisson regression models[7, 8, 10] are the most widely accepted (Table 1).

Table 1. Assumptions, estimation, and other characteristics of logistic, log-binomial and Poisson regression models

	Logistic	Log-binomial	Poisson
Model assumptions	Categorical or ordinal outcome, independent observations, linear relationship between explanatory variables and logit	Binary outcome, independent observations, linear relationship between explanatory variables and the log probability	The outcome is count data (non-negative integers), independent observations, linear relationship between explanatory variables and the logarithm of the mean, variance equal to the mean (no overdispersion)
Model estimation method	Maximum likelihood estimation	Maximum likelihood estimation	Maximum likelihood estimation
Model link function	Logit of the proportion	Logarithm of the proportion	Logarithm of the mean
Distribution of errors	Binomial	Binomial	Poisson
Model equation	$\ln(\frac{\pi}{1-\pi}) = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + \varepsilon$, where π is the probability of success, X_i the independent variables, *k* the number of independent variables, ε represents the error term.	$\log(\pi) = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + \varepsilon$, where π is the probability of success, X_i the independent variables, and *k* the number of independent variables, ε represents the error term.	$\log\frac{n}{t} = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + \varepsilon$, where *n* is the count of events for a given individual, *t* the follow-up time, *Xi* the independent variables, *k* the number of independent variables, ε represents the error term.
Model estimates	OR, the exponential of the model coefficients ($e^{\beta k}$)	RR/PR, the exponential of the model coefficients ($e^{\beta k}$)	RR/PR, the exponential of the model coefficients ($e^{\beta k}$)
Other considerations	Probabilities between 0-1, reciprocity	Probabilities can be > 1, may have convergence issues, non-reciprocity	Probabilities can be >1, non-reciprocity

The log-binomial regression model, another generalized linear model, is characterized by a binomial distribution of errors and uses the maximum likelihood estimation method, similarly to the logistic regression model. However, the log-binomial model uses a logarithmic link function, instead of the logit. This function models the logarithm of the proportion of a binary outcome, allowing for the direct estimation of PR and RR. But, unlike the logit link function, which confines the model-predicted probabilities between 0 and 1, the logarithmic only constrains this probability to be ≥ 0. Besides that, log-binomial models may encounter other challenges like non-reciprocity and non-convergence, which might limit their use.[8, 11]

The Poisson regression model, another generalized linear model, and the nominal model for count data, is usually regarded as an appropriate approach for analysing rare events when subjects are followed for a length of time.[8, 10] This model uses a logarithmic link function, similarly to the log-binomial model, but with a Poisson distribution of errors instead of a binomial distribution. By assuming that the logarithm of the Poisson parameter (the mean) is linearly related to a set of independent

variables, the exponentiation of the model coefficients will yield estimates of a ratio of Poisson parameters. Poisson regression models are commonly used in cohort and other incidence studies to estimate RR. In cross-sectional studies with binary data, the computed ratios can still be interpreted, but this time as an approximation to the PR.[8, 10] This approach yields correct estimates of the PR directly like log-binomial models, but typically with larger variances and standard errors.[8, 10] Modified Poisson regression models, like the robust Poisson model, are available to address this issue. Robust standard errors, determined through the Huber sandwich estimation (Annex 2), correct the Poisson model for data overdispersion (when variance is higher than the mean) and variance inflation, being this correction the essence of the robust Poisson model. This robust estimator performs well even when data does not perfectly meet the model's requirement, like in this specific case where Poisson regression models are applied to cross-sectional binary data.[10]

As most researchers still rely on logistic regression models because of their simplicity, the complex process of statistical model selection might pose a real challenge, and lead to inaccurate decisions.[3] Since methodological consistency is important to epidemiological and clinical research and given the potential impact of biased estimates on clinical practice, policymaking, public health communications and community behaviour, the selection of the appropriate statistical model holds paramount importance and should be clarified.

The hypothesis in study was that log-binomial and robust Poisson models are more appropriate than logistic regression models to estimate measures of association in cross-sectional studies involving frequent binary outcomes, provided the assumptions for the models' application are met.

1.1 Objective

This dissertation aimed to compare the estimation and goodness-of-fit of logistic, log-binomial, and robust Poisson regression models, in cross-sectional studies with frequent (>10%) binary outcomes.

2

METHODS

We carried out two cross-sectional studies (Study 1 and Study 2) to estimate the measures of association between two different exposures and two different frequent (>10%) binary outcomes, accounting for confounding from several variables identified as potential confounders through directed acyclic graphs, elaborated based on a literature review of the associations in study. We conducted a comparative analysis of the estimates of the measures of association (OR in the case of logistic regression models, and PR in the case of log-binomial and robust Poisson regression models), in terms of their magnitude and significance. The estimates' confidence intervals, their ranges and standard errors were also compared.

The maximum likelihood estimation method was employed in all the computed models, even for those using complex survey data (Study 1). In this case, we fitted models to complex data incorporating the sampling weights in a designed-based analysis (a statistical analysis of survey data which took the survey design, its stratification and clustering, and the sampling weights into consideration).

As the Akaike Information Criterion (AIC), ideally suited to generalized linear modelling applications, is a widely used tool in model selection and allows to compare the relative goodness-of-fit of different models, AIC values were obtained for every model in study fitting non-complex data (Study 2) and compared. The optimal model should be the one with the minimum AIC value. Any model yielding an AIC value within two units of the minimum AIC value might also be an appropriate candidate.[12]

All statistical analyses were conducted using Stata®, version 15. The significance level was established at 5%.

2.1 Study 1 - Association Between Long-term Exposure to Ambient Air Pollution and Probable Diagnosis of Depressive and Anxiety Disorders

Particulate matter smaller than 10 μm (PM_{10}) accounts for much of the impact of air pollution on health.[13, 14] The physiological path between PM_{10} and mental disorders has been proposed.[15, 16] Some studies have assessed the association between long-term exposure to PM_{10} and common mental disorders (CMD), namely depressive and anxiety disorders, but evidence is inconsistent.[17-22] To estimate the association between long-term exposure to PM_{10} and the frequency of probable diagnosis of CMD, a population-based, nationally-representative cross-sectional study was conducted, in mainland Portugal.

Long-term exposure to PM_{10} was estimated through 1-year average concentrations of PM_{10}, calculated with data from the Portuguese Environment Agency's air quality monitoring stations, and attributed to each individual considering their 7-digit postal codes of residence.

The probable diagnosis of CMD was ascertained through the scores obtained in the 5-item Mental Health Inventory (MHI-5). MHI-5 is a self-report scale from 0 to 100 points, which measures negative and positive dimensions of mental health, in the four weeks prior to the day of assessment (Supplementary Figure 1).[23-25] It was nationally[26, 27] and internationally[25, 28, 29] validated for the screening of major depressive disorder, affective disorders (unipolar and bipolar depressive disorders), and anxiety disorders. A score ≤ 52 in the MHI-5 represents a situation that implies proper clinical evaluation by a doctor (a "Probable diagnosis of CMD"), while a score > 52 represents a situation that does not imply proper clinical evaluation.

Based on our literature review, demographic variables (sex, age), individual socioeconomic variables (education level, employment status, professional occupation), aggregate socioeconomic variables (area-level socioeconomic deprivation), and environmental variables (individually allocated 1-year average temperature, degree of urbanization, area-level walkability) were included in the minimal set of adjustment required to correct for confounding in the study of the present association.

MHI-5 scores and the independent variables' data were obtained from a restricted sample of the participants of the first Portuguese National Health Examination Survey (INSEF), implemented in Portugal, from February 2 to December 21, 2015. INSEF targeted non-institutionalized individuals from 25 to 74 years old, living in Portugal for at least 12 months, who were able to follow an interview in Portuguese, and collected self-reported data on demographic, socioeconomic, lifestyle and health variables.[30] The individuals in study were the INSEF participants from mainland Portugal who consented on data linkage, had their 7-digits postal codes of residence available in the database, were living within a 30-km radius from a background air quality monitoring station (with available data on PM_{10} atmospheric concentration), like previously detailed.[31-34], and who had answered all the five items of the

MHI-5. Data on area-level socioeconomic deprivation, the Portuguese version of the European Deprivation Index, was available online for all the Portuguese parishes at https://figshare.com/s/3a4226d520df3b18cb71 and was studied in terciles of increasing deprivation.[35] Individually allocated 1-year average temperatures were obtained through a similar methodology to the one applied to estimate exposure to PM_{10}, making use of 1-year average temperatures collected from the National Oceanic and Atmospheric Administration database, available online at https://www.ncdc.noaa.gov/.[31, 32, 34] Data on area-level walkability was obtained through the weighted sum of residential density, street connectivity, and a land use mix index, for all the parishes of mainland Portugal, and is available at request, in terciles of increasing walkability.

Age was categorized into four categories: 25-34 years, 35-49 years, 50-64 years, and 65-74 years, according to the cut-offs used in the World Mental Health Survey.[36] Sex and degree of urbanization were studied as available in INSEF database: female/male and rural/urban, respectively.[30] Education level was categorized into three categories, considering the highest level of education completed and according to the 2011 International Standard Classification of Education (ISCED, 2011)[37]: low education (levels 0–2 of ISCED, 2011), medium education (levels 3–4 of ISCED, 2011), and high education (levels 5–8 of ISCED, 2011). Employment status was categorized into three categories, as in previous studies[38]: employed, unemployed, and other without a professional activity (student, retired, disabled, housewife, or other). Professional occupation was categorized into two categories, according to the International Standard Classification of Occupations (ISCO-08)[39]: white-collar (managers; professionals; technicians, and associate professionals; clerical support workers; and services, and sales workers) and blue-collar (skilled agricultural workers; craft, and related trades workers; plant, machine operators, and assemblers; and elementary occupations).

As the outcome's frequency was higher than 10%, logistic, log-binomial, and robust Poisson regression models were computed (Table 2). We performed single-level multivariate analyses since, even if we used individual and aggregated variables (at parish level) in our models, the assumptions for carrying out a multilevel analysis were not met (some parishes had just one individual).

Table 2. Study 1 models' specification

	Logistic	Log-binomial	Robust Poisson
Hypothesis in study	Long-term exposure to PM_{10} is associated to an increased probable diagnosis of common mental disorders		
Exposure	Long-term exposure to PM_{10} (individually allocated 1-year average measured concentrations of PM_{10}, in μm/m^3)		
Other independent variables	Sex (female/male, reference), age (25-34, reference/35-49/50-64/65-74 years), education level (low, reference/medium/high), employment status (employed, reference/unemployed/other), professional occupation (white, reference/blue-collar), area-level socioeconomic deprivation (terciles), individually allocated 1-year average temperature (continuous, in ºC), degree of urbanization (urban/rural, reference), area-level walkability (terciles)		

Outcome	Probable diagnosis of common mental disorders (yes, if ≤ 52 in the MHI-5 score/no, if >52)		
Fitted model equation	$\ln\left(\frac{p}{1-p}\right) = b_0 + b_1 \times \text{long-term exposure to } PM_{10} + b_2 \times \text{sex (female)} + b_3 \times \text{age } (35-49) + b_4 \times \text{age } (50-64) + b_5 \times \text{age } (\geq 65) + b_6 \times \text{education level (medium)} + b_7 \times \text{education level (high)} + b_8 \times \text{employment status (unemployed)} + b_9 \times \text{employment status (other)} + b_{10} \times \text{professional occupation (blue-collar)} + b_{11} \times \text{area-level socioeconomic deprivation (mediu} \; b_{12} \times \text{area-level socioeconomic deprivation(high)} \; b_{13} \times \text{temperature} + b_{14} \times \text{degree of urbanization (urban)} + b_{15} \times \text{area-level walkability (medium)} + b_{16} \times \text{area-level walkability (high)}$, where *p* is the probability of having a probable diagnosis	Did not converge	$\log(p) = b_0 + b_1 \times \text{long-term exposure to } PM_{10} + b_2 \times \text{sex (female)} + b_3 \times \text{age } (35-49) + b_4 \times \text{age } (50-64) + b_5 \times \text{age } (\geq 65) + b_6 \times \text{education level (medium)} + b_7 \times \text{education level (high)} + b_8 \times \text{employment status (unemployed)} + b_9 \times \text{employment status (other)} + b_{10} \times \text{professional occupation (blue-collar)} + b_{11} \times \text{area-level socioeconomic deprivation (medium)} + b_{12} \times \text{area-level socioeconomic deprivation(high)} + b_{13} \times \text{temperature} + b_{14} \times \text{degree of urbanization (urban)} + b_{15} \times \text{area-level walkability (medium)} + b_{16} \times \text{area-level walkability (high)}$, where *p* is the probability of having a probable diagnosis

PM_{10} – particulate matter lower than 10 micrometres, MHI-5 – Mental Health Inventory 5.

All estimates were weighted to account for different selection probabilities resulting from the complex sampling design, and to match the population distribution in terms of geographic region, age group and sex, unless specifically stated.

In this study, no ethical or legal issues of confidentiality were raised, since all the data came from anonymized databases, and the due approvals were gathered from the different competent authorities.

2.2 Study 2 - Association Between Immigration Status and Urgent Care Use in the Paediatric Population

Foreign residents have been increasing in Portugal, in the past years.[40] In 2021, 542 165 migrants were estimated to be living in Portugal.[40] One of the biggest challenges immigrants face in the host countries is to obtain access to health care services.[41]

The Portuguese National Health Service offers universal and free healthcare services for children up to 18 years old. As a result, healthcare provisions are consistent for all children, irrespective of their immigration status.[42] Several studies have already reported the increased urgent care use by migrants, compared to non-migrants[43-45], but evidence is inconsistent[43, 46-50], and differences exist according to migrants' characteristics and to countries' different health systems.[43, 51]

To estimate the association between being an immigrant and the urgent care use, in the paediatric population living in Amadora, a population-based, non-representative cross-sectional study was

conducted there, in the most densely populated Portuguese municipality, and the second with the highest density of foreign residents.[42]

Data was obtained from the participants of the first wave of the CRIAS (Health trajectories of Immigrant Children in Amadora) cohort, implemented in Portugal from June 2019 to March 2020. CRIAS targeted children born in 2015, aged 4 or 5 years old, and who had records of attending at least one of the primary health care centres of Amadora in the 2 years before the assessment time, also targeting their parents. Self-reported data on the families' demographic and socioeconomic characteristics, migration history, and on the children's lifestyle and health was collected and integrated with data on the children's health care use, obtained through electronic health records.[42] The objective of the cohort managers was to include 50% of immigrants in the cohort sample. From the CRIAS participants, those with missing data on the variables immigration status or urgent care use were excluded from the present study.

Immigration status was categorized in "immigrant" (born to non-native parents, even if in Portugal) and "non-immigrant" (born in Portugal, to native parents). Urgent care use was categorized in "yes" or "no", according to the children having or not at least one visit to the urgent care service of the *Hospital Professor Doutor Fernando Fonseca* (the referral hospital of Amadora's primary health care centres) in 2019.

Based on our literature review, a demographic variable (caregiver's age), and two individual socioeconomic variables (caregiver's education level and caregiver's professional occupation) were included in the minimal set of adjustment required to correct for confounding in the study of this association. Caregiver's age was categorized in two categories: <35 and ≥35 years old. Caregiver's education level and professional occupation were categorized into three and two categories, according to the ISCED, 2011 and to the ISCO-08, respectively, as in Study 1.

As the outcome's frequency was higher than 10%, logistic, log-binomial, and robust Poisson regression models were computed (Table 3). AIC values were obtained for all the fitted models.

Table 3. Study 2 models' specification

	Logistic	**Log-binomial**	**Robust Poisson**
Hypothesis in study	The immigration status of a child is associated to the urgent care service use, in the paediatric population living in Amadora		
Exposure	Immigration status (immigrant/non-immigrant, reference)		
Other independent	Caregiver's age (<35, reference/≥35 years), caregiver's education level (less than secondary, reference/secondary/ professional, or higher) and caregiver's professional occupation (white, reference/blue-collar)		

variables			
Outcome	Urgent care service use (yes, if the child had at least one visit to the urgent care service of the *Hospital Professor Doutor Fernando Fonseca*, in the year 2019/no, if there were no visits)		
Fitted model equation	$\ln\left(\frac{p}{1-p}\right) = b_0 + b_1 \times immigrant\ status\ (immigrant) + b_2 \times caregiver's\ age\ (\geq 35) + b_3 \times education\ level\ (secondary) + b_4 \times education\ level\ (professional) + b_5 \times professional\ occupation\ (blue-collar)$, where *p* is the probability of having used the urgent care service	log (*p*) $= b_0 + b_1 \times immigrant\ status\ (immigrant) + b_2 \times caregiver's\ age\ (\geq 35) + b_3 \times education\ level\ (secondary) + b_4 \times education\ level\ (professional) + b_5 \times professional\ occupation\ (blue-collar)$, where *p* is the probability of having used the urgent care service	log (*p*) $= b_0 + b_1 \times immigrant\ status\ (immigrant) + b_2 \times caregiver's\ age\ (\geq 35) + b_3 \times education\ level\ (secondary) + b_4 \times education\ level\ (professional) + b_5 \times professional\ occupation\ (blue-collar)$, where *p* is the probability of having used the urgent care service

In this study, no ethical or legal issues of confidentiality were raised, since all the data came from anonymized databases, and the due approvals were gathered from the different competent authorities.

3

Results

3.1 Study 1 - Association Between Long-term Exposure to Ambient Air Pollution and Probable Diagnosis of Depressive and Anxiety Disorders

A total of 2 398 individuals were included in the study, following the application of the inclusion and exclusion criteria (Supplementary Figure 2). Included and excluded individuals were similar regarding most of the analysed characteristics. Differences between these two groups were only found regarding the individual allocated 1-year average temperature, and the area-level walkability terciles (Supplementary Table 1).

The study population had higher percentages of females (52.6%, 95%CI: 50.2-55.1) and individuals belonging to the age groups 35-49 years (34.1%, 95%CI: 32.0-36.4) and 50-64 years (31.5%, 95%CI: 29.4-33.7) (Table 4). The median [interquartile range (IQR)] individual allocated 1-year average PM_{10} concentration was 18.6 (15.3-19.3) μg/m³ (Table 5). A probable diagnosis of DMC was found in 22.7% (95%CI: 20.0-25.6) of the study population (Table 4).

Table 4. Estimates for the characteristics in study of the included individuals

	n	Sample estimates (%/median)$	Weighted estimates (95%CI) (%/median)$	
Exposure				
Individual allocated 1-year average PM_{10} (μg/m³)	2398	19.0	18.6	(18.4-18.7)
Independent variables				
Sex				
Female	1294	54.0	52.6	(50.2-55.1)
Male	1104	46.0	47.4	(44.9-49.9)
Age group				
25-34 years	353	14.7	18.9	(17.2-20.7)

35-49 years	826	34.4	34.1	(32.0-36.4)
50-64 years	827	34.5	31.5	(29.4-33.7)
65-74 years	392	16.4	15.5	(13.9-17.2)
Education level*				
Low[a]	679	28.3	26.5	(24.4-28.7)
Medium[b]	1279	53.4	53.5	(50.2-56.8)
High[c]	439	18.3	20.0	(16.7-23.7)
Employment status*				
Employed	1448	60.4	62.3	(59.4-65.0)
Unemployed	272	11.4	11.4	(9.7-13.4)
Other[d]	677	28.2	26.3	(24.1-28.7)
Professional occupation*				
White-collar[e]	1376	62.3	62.9	(58.5-67.1)
Blue-collar[f]	834	37.7	37.1	(32.9-41.5)
Area-level socioeconomic deprivation terciles				
Low deprivation (T1)	475	19.8	16.0	(10.5-24.5)
Moderate deprivation (T2)	609	25.4	31.6	(19.3-47.2)
High deprivation (T3)	1314	54.8	52.4	(37.8-66.6)
Individual allocated 1-year average temperature (ºC)	2398	15.2	15.0	(14.8-15.3)
Area-level walkability terciles				
Low walkability (T1)	84	3.5	1.3	(0.6-2.9)
Moderate walkability (T2)	574	23.9	18.4	(13.5-24.6)
High walkability (T3)	1740	72.6	80.3	(73.2-85.9)
Degree of urbanization				
Rural	702	29.3	28.2	(22.7-34.5)
Urban	1696	70.7	71.8	(65.5-77.3)
Outcomes				
Probable diagnosis of common mental health disorders				
Yes	555	23.1	22.7	(20.0-25.6)
No	1843	76.9	77.3	(74.4-80.0)

95%CI - 95% confidence intervals, PM_{10} - particulate matter with an aerodynamic diameter less than or equal to 10 micrometres, T1 - first tercile, T2 - second tercile, T3 - third tercile.
*1 missing in education level, 1 missing in employment status, 188 missings in professional occupation.
$ percentages or medians were presented according to the type of variable being described (categorical or continuous, respectively).
a: Less than secondary: levels 0–2 of the International Standard Classification of Education 2011 .
b: Secondary: levels 3–4 of the International Standard Classification of Education 2011 .
c: Professional or higher: levels 5–8 of the International Standard Classification of Education 2011.
d: Other: students, retired, stay-at-home parents, other.
e: White-collar: managers, professionals, technicians and associate professional, clerical support workers, and services and sales workers (according to the International Standard Classification of Occupations).
f: Blue-collar: skilled agricultural workers, craft and related trades workers, plant and machine operators, and elementary occupations (according to the International Standard Classification of Occupations).

Table 5. Characteristics of the included individuals (n=2 398) and comparison between groups of probable diagnosis of common mental health disorders

	Total (n=2 398)	**Probable diagnosis**	**Without probable**	**p-value/ 95%CI[$]**

		of CMD (n=555)	diagnosis of CMD (n=1 843)	
Individual allocated 1-year average PM_{10} (μg/m³) – median (IQR)	18.6 (15.3-19.3)	18.3 (15.2-19.2)	18.6 (15.3-19.3)	**(0.012, 0.423)**
Sex - %				**<0.001**
Female	52.6	70.4	47.4	
Male	47.4	29.6	52.6	
Age group - %				0.072
25-34 years	18.9	14.3	20.2	
35-49 years	34.1	31.1	35.0	
50-64 years	31.5	37.3	29.9	
65-74 years	15.5	17.3	14.9	
Education level* - %				**<0.001**
Low[a]	26.5	33.2	24.5	
Medium[b]	53.5	53.2	53.6	
High[c]	20.0	13.6	21.9	
Employment status* - %				**<0.001**
Employed	62.3	51.5	65.4	
Unemployed	11.4	14.3	10.6	
Other[d]	26.3	34.2	24.0	
Professional occupation* - %				0.071
White-collar[e]	62.9	58.1	64.2	
Blue-collar[f]	37.1	41.9	35.8	
Area-level socioeconomic deprivation terciles - %				0.087
Low deprivation (T1)	16.0	16.4	15.9	
Moderate deprivation (T2)	31.6	36.0	30.3	
High deprivation (T3)	52.4	47.6	53.8	
Individual allocated 1-year average temperature (ºC) – median (IQR)	15.0 (14.8-16.7)	14.9 (14.6-16.7)	15.2 (14.8-16.7)	**(0.015, 0.141)**
Area-level walkability terciles - %				0.489
Low walkability (T1)	1.3	1.5	1.3	
Moderate walkability (T2)	18.4	20.6	17.7	
High walkability (T3)	80.3	77.9	81.0	
Degree of urbanization - %				0.051
Rural	28.2	33.0	26.9	
Urban	71.8	67.0	73.1	

CMD - common mental health disorders, 95%CI - 95% confidence interval, PM_{10} - particulate matter with an aerodynamic diameter less than or equal to 10 micrometres, T1 - first tercile, T2 - second tercile, T3 - third tercile.
*1 missing in education level, 1 missing in employment status, 188 missings in professional occupation.
$ p-values or 95%CI were presented according to the test used to compare proportions (Chi-squared test) or medians (robust confidence intervals for generalized Hodges-Lehmann median differences), respectively. Results in bold are those with statistically significant differences between individuals with and without probable diagnosis of CMD.
a: Less than secondary: levels 0–2 of the International Standard Classification of Education 2011 .

b: Secondary: levels 3–4 of the International Standard Classification of Education 2011 .
c: Professional or higher: levels 5–8 of the International Standard Classification of Education 2011.
d: Other: students, retired, stay-at-home parents, other.
e: White-collar: managers, professionals, technicians and associate professional, clerical support workers, and services and sales workers (according to the International Standard Classification of Occupations).
f: Blue-collar: skilled agricultural workers, craft and related trades workers, plant and machine operators, and elementary occupations (according to the International Standard Classification of Occupations).
All the estimates were weighted to account for different selection probabilities and population distribution.

3.1.1 Regression Models

The adjusted logistic regression model obtained an OR (95%CI) = 1.015 (0.970-1.063), and the robust Poisson model a PR (95%CI) = 1.012 (0.979-1.045). The adjusted log-binomial regression model did not converge. Adjusted OR overestimated the adjusted PR obtained through the robust Poisson model by 0.003 and presented wider 95%CI (95%CI range: 0.093 vs 0.066, respectively) and a higher standard-error (0.022 vs 0.016, respectively) (Table 6).

Table 6. Characteristics of logistic, log-binomial and robust Poisson regression models fitted for the estimation of measures of association between the exposure to particulate matter with an aerodynamic diameter equal than or lower to 10 micrometres and the frequency of common mental disorders

	Logistic	Log-binomial	Robust Poisson
Number of individuals included in the adjusted models	2210	2210	2210
Adjusted OR/PR* (95%CI)	1.015 (0.970-1.063)	Did not converge	1.012 (0.979-1.045)
95%CI range	0.093	Did not converge	0.066
Standard-error of the adjusted estimates	0.022	Did not converge	0.016

OR - odds ratio, PR - prevalence ratio, 95%CI - 95% confidence interval, PM10 - particulate matter with an aerodynamic diameter less than or equal to 10 micrometres.
* Adjusted for sex, age groups, education level, employment status, professional occupation, area-level socioeconomic deprivation terciles, individual allocated 1-year average temperature, area-level walkability terciles, degree of urbanization [OR for logistic and log-binomial regression models, PR for robust Poisson regression model].

No statistically significant association between long-term exposure to PM_{10} and the frequency of probable CMD diagnosis was observed in any of the models, after adjustment.

3.2 Study 2 - Association Between Immigration Status and Urgent Care Use in the Paediatric Population

From the CRIAS participants, 410 children were included in the study, since 10 were excluded due to missing data in the variable urgent care use.

Most of the children were males (50.7%), being most of their caregivers (or questionnaire respondents) females (87.6%) with a median (IQR) age of 34 (18-75) years old (Table 7). Among the children in study, 50.5% were immigrants (Table 7). Approximately 48.0% of all children used urgent care services (58.4% of them being immigrants) (Tables 7 and 8).

Table 7. Baseline characteristics of the children and their caregivers, by immigration status

	Total (n = 410)	**Immi-grants** (n = 207)	**Non-immi-grants** (n=203)	**p-value**
Children's sex - %				0.697**
Male	50.7	49.8	51.7	
Female	49.3	50.2	48.3	
Caregivers' sex - %				0.331**
Male	12.4	14.0	10.8	
Female	87.6	86.0	89.2	
Caregivers' age* - median(min-max)	34 (18-75)	34 (20-75)	35 (18-68)	0.196$
Caregivers' relationship with the child - %				0.348**
Parent	97.3	96.1	98.5	
Other	2.7	3.9	1.5	
Caregivers' education level* - %				0.093**
Less than secondary[a]	37.4	38.8	36.0	
Secondary[b]	37.7	40.8	34.5	
Professional or higher[c]	24.9	20.4	29.5	
Caregivers' employment status - %				0.181**
Employed	82.2	79.7	84.7	
Unemployed or other[d]	17.8	20.3	15.3	
Caregivers' professional occupation* - %				**<0.001****
White-collar[e]	35.6	20.5	51.3	
Blue-collar[f]	64.4	79.5	48.7	
Household monthly income* - %				**<0.001****
< 750 euros	41.1	52.0	30.0	
≥ 750 euros	58.9	48.0	70.0	
Children urgent care use* - %				**0.002****
No	52.0	44.4	59.6	
Yes	48.0	55.6	40.4	
Children number of urgent care visits* - median (min-max)	2 (1-10)	2 (1-10)	2 (1-10)	0.438$

Results in bold are those with statistically significant differences between immigrants and non-immigrants.
*3 missings in caregivers' age, 1 missing in caregivers' education level, 18 missings in caregivers' professional occupation, 22 missings in household monthly income, 10 missings in children urgent care use, 23 missings in children number of urgent care visits.

**Chi-squared test. $Mann-Whitney U test.
a: Less than secondary: levels 0–2 of the International Standard Classification of Education 2011 .
b: Secondary: levels 3–4 of the International Standard Classification of Education 2011 .
c: Professional or higher: levels 5–8 of the International Standard Classification of Education 2011.
d: Other: students, retired, stay-at-home parents, other.
e: White-collar: managers, professionals, technicians and associate professional, clerical support workers, and services and sales workers (according to the International Standard Classification of Occupations).
f: Blue-collar: skilled agricultural workers, craft and related trades workers, plant and machine operators, and elementary occupations (according to the International Standard Classification of Occupations).

Table 8. Characteristics of children and their caregivers, by having or not used urgent care in 2019

	Total (n=410)	**Urgent care use** (n=197)	**No urgent care use** (n=213)	**p-value**
Children's immigration status - %				**0.002****
Non-immigrant	49.5	41.6	56.8	
Immigrant	50.5	58.4	43.2	
Caregivers' age* (years) - median(min-max)	34 (18-75)	33 (20-69)	36 (18-75)	**<0.001$**
Caregivers' education level* - %				**0.004****
Less than secondary[a]	37.4	43.3	31.9	
Secondary[b]	37.7	38.8	36.6	
Professional or higher[c]	24.9	17.9	31.5	
Caregivers' professional occupation* - %				**<0.001****
White-collar[d]	35.6	24.7	45.8	
Blue-collar[e]	64.4	75.3	54.2	

Results in bold are those with statistically significant differences between immigrants and non-immigrants.
*3 missings in caregivers' age, 1 missing in caregivers' education level, 17 missings in caregivers' professional occupation.
**Chi-squared test. $Mann-Whitney U test.
a: Less than secondary: levels 0–2 of the International Standard Classification of Education 2011 .
b: Secondary: levels 3–4 of the International Standard Classification of Education 2011 .
c: Professional or higher: levels 5–8 of the International Standard Classification of Education 2011.
d: White-collar: managers, professionals, technicians and associate professional, clerical support workers, and services and sales workers (according to the International Standard Classification of Occupations).
e: Blue-collar: skilled agricultural workers, craft and related trades workers, plant and machine operators, and elementary occupations (according to the International Standard Classification of Occupations).

3.2.1 Regression Models

The adjusted logistic regression model obtained an OR (95%CI) = 1.584 (1.026-2.446), the adjusted log-binomial regression model a PR (95%CI) = 1.217 (0.978-1.515), and the robust Poisson regression model a PR (95%CI) = 1.130 (1.013-1.261). The adjusted OR overestimated the adjusted PR obtained through the log-binomial and robust Poisson regression models by 0.367 and 0.454, respectively, presented wider 95%CI (95%CI range: 1.420 vs 0.537 and 0.248, respectively) and higher standard-errors (0.351 vs 0.136 and 0.063, respectively). However, the AIC value was lower for the logistic regression model than for the log-binomial and robust Poisson models (1.345 vs 1.350 and 1.656, respectively) (Table 9).

Table 9. Characteristics of logistic, log-binomial, and robust Poisson regression models fitted for the estimation of measures of association between immigration status and urgent care use

	Logistic	Log-binomial	Robust Poisson
Number of individuals included in the adjusted models	389	389	389
Adjusted OR/PR* (95%CI)	**1.584 (1.026-2.446)**	1.217 (0.978-1.515)	**1.130 (1.013-1.261)**
95%CI range	**1.420**	0.537	**0.248**
AIC for the adjusted model	1.345	1.350	1.656
Standard-error of the adjusted estimate	0.351	0.136	0.063

OR - odds ratio, PR - prevalence ratio, 95%CI - 95% confidence interval, AIC - Akaike information criteria.
* Adjusted for caregiver's age, education level and professional occupation [OR for logistic and log-binomial regression models, PR for robust Poisson regression model]. Results in bold are those statistically significant.

Only logistic and robust Poisson regression models showed statistically significant associations between immigration status and urgent care service use, in the adjusted models. The log-binomial model, however, did not yield statistically significant associations.

4

DISCUSSION

This dissertation, based on two distinct real-life epidemiological examples involving outcomes with different prevalence (22.7% in Study 1, and 48.0% in Study 2), contributed to the knowledge about the estimation of measures of association, in cross-sectional studies with frequent outcomes, in different settings. Study 1 assessed a nationally-representative sample of 2 398 individuals, whereas Study 2 focused on a local sample of 410 individuals. Study 1 estimated the association between a continuous exposure and a binary outcome, adjusting for 9 independent variables (including one continuous variable); and Study 2 estimated the association between a binary exposure and a binary outcome, adjusting for 3 covariates.

In Study 1, the log-binomial regression model did not converge. This likely happened due to the inclusion of two continuous independent variables in the model (one variable which was a potential confounder and the exposure in study).[9, 52, 53] When in the presence of quantitative variables, the maximum likelihood estimate can lie on the boundary of the parameter space which can lead to model non-convergence as the instantaneous slope of the likelihood may not reach zero at this boundary.[7] This issue supports authors who advocate for the robust Poisson regression model as the preferred choice to logistic models, when addressing frequent outcomes in cross-sectional studies, instead of log-binomial models.[52, 53]

In this study, as the obtained measures of association were higher than 1, the OR overestimated the PR obtained through the robust Poisson regression model. The magnitude of the overestimation was 0.003 units (OR = 1.015 vs PR = 1.012). The precision of the OR was lower than the precision of the PR, which presented narrower 95%CI, with the 95%CI range for the OR being 0.027 units higher than the obtained for the PR through the robust Poisson regression model (0.093 vs 0.066, respectively). Consistently, the standard-error was 0.006 units lower for the robust Poisson (Robust Poisson: 0.016 vs Logistic: 0.022, respectively).

In Study 2, the obtained measures of association were also higher than 1, with the consequent overestimation of PR by OR already reported in Study 1. The magnitude of the overestimation was higher when the OR was compared with the PR obtained through the robust Poisson model, than with the PR obtained through the log-binomial model (overestimations of 0.454 and 0.367 units, respectively). This happened because the robust Poisson model obtained a lower value for the PR than the log-binomial model (PR: 1.130 vs 1.217, respectively). The robust Poisson model also obtained the most precise estimate (the one with the lowest 95%CI range), followed by the log-binomial and, lastly, by the logistic regression model (95%CI ranges: 0.248, 0.537, and 1.420, respectively), as well as the

lowest standard-error for the estimate. However, the AIC values were lower for the logistic regression model, followed by the log-binomial model, and, lastly, by the robust Poisson model, indicating that the logistic was the one better fitting to the data, which was counterintuitive. This might have been because the binomial distribution is ultimately more appropriate to module binary outcomes than the Poisson distribution, and because the mean and the variance of the outcome were different in this study (variance = 0.250, mean = 0.480), affecting the fit of the Poisson model, even if a robust estimator was being used.[8] This example illustrates why relying solely on the AIC value for the selection of the most appropriate statistical model is not adequate.

In this study, we also observed that the log-binomial model could not obtain a statistically significant association, while the logistic and robust Poisson models did. This might have been due to the conjunction of the estimation of a not so inflated estimate through the log-binomial model, as through the logistic model, with wider 95%CI for the log-binomial than for the robust Poisson model.

Moreover, in Study 2, the magnitude of the overestimation exceeded the overestimation observed in Study 1 (Study 2: 0.454 and 0.367 units, Study 1: 0.003 units), which was also verified in terms of the 95%CI ranges. This discrepancy was likely due to the higher prevalence of the outcome in Study 2 compared to Study 1 and illustrates that higher outcome prevalence represent larger OR inflations.[1, 4]

In both Study 1 and Study 2, the robust Poisson model exhibited greater precision (narrower 95%CI and lower 95%CI ranges) than log-binomial and logistic regression models. This characteristic is particularly relevant when the aim of a study is to estimate unbiased measures of association (where the expected value is identical to the population parameter being estimated), but not so when the aim is solely to comprehend the general trend or direction of it.[2]

In both examples, we observed that OR obtained through logistic regression models provide biased estimates of PR, in the presence of frequent binary outcomes, providing also wider confidence intervals, which magnified the probability of a type II error. Additionally, we verified that the extension of the bias became more pronounced as the prevalence of the outcome increased, consistently with findings from prior studies.[3, 5, 7, 9, 52-55]

The challenges associated with estimating measures of association through logistic regression models when there are frequent outcomes in study are not limited to cross-sectional designs. Longitudinal studies and randomized clinical trials might encounter similar estimation issues, and also need to be investigated.[6]

This dissertation reinforces the viability of alternative statistical models and reopens a discussion that has yet to reach a definitive solution. These alternative models are readily accessible within commonly used statistical software packages like SPSS®, Stata® and R Studio®. Nonetheless, they have potential disadvantages that could lead researchers to discard them and stick to the logistic regression models. A disadvantage of the log–binomial regression is that the model might not converge, like in Study 1.[11] While robust Poisson regression does not have convergence issues, this model, like the

log-binomial, may yield individual predicted probabilities above 1. The robust Poisson regression is indeed a suitable method to obtain valid, unbiased estimates, but not to predict individual probabilities (as in diagnostic or prognostic studies).[4, 8, 10] Moreover, both models lack reciprocity, underscoring the importance of being mindful of the reference category selected.[3]

Given the curriculum requirements that often compel medical doctors and other health professionals to publish, recognizing the time constraints clinicians face to master research methodologies and statistics, and considering the need for consistency in research, it is imperative to establish a simple and, as much as possible, standardized approach for the selection of the most suitable statistical model to employ in a study. Several processes have been proposed. One of them involves constructing a contingency table between the exposure and the outcome and comparing the OR and RR/PR obtained through epidemiological formulas of calculation. Relevant disparities between the values would suggest the employment of a model that estimates RR/PR instead of OR (such as the log-binomial and robust Poisson models).[55] However, this methodology is too simplistic, can only be applied in the presence of a categorical exposure, and does not help to choose between log-binomial and robust Poisson models. Another proposal was made by a group of authors who created a flowchart to support researchers, but its interpretation could be complex as it requires some methodological knowledge, and it was not agreed upon consensus.[56]

While other options exist, the log-binomial and the robust Poisson regression models are relatively straightforward to apply, easy to interpret, and offer the ability to control for confounding.[8, 10]

4.1 Strengths and Limitations

This dissertation has several strengths. First, it contributed to reignite the discussion around the validity of the OR to infer RR/PR, in the context of frequent binary outcomes, in cross-sectional studies. Second, it used real-life datasets applied to different settings - in terms of topics investigated (the health impact of air pollution, and the access to healthcare services), types of exposure assessed (continuous, and categorical), types of independent variables included (categorical and continuous, or just categorical), and outcomes' prevalence (22.7%, and 48.0%) -, allowing for a comprehensive discussion, pertinent to several different scientific fields and aims. Third, the example studies were conducted according to robust and previously published methodological approaches (for instance, in terms of exposure assessment), allowing us to diminish biases and confounding, and to add findings relevant for the scientific literature with each of the example studies developed.

Some limitations should also be noted. First, the assessed associations did not have substantial effect magnitudes, making the estimates' values closer to one and, across models, to each other (due to a reduced variability). Second, it uses just two examples, and does not make use of simulation. Third, logistic regression models were just compared against log-binomial and robust Poisson regression models. This choice was made not only based on the appropriateness of these models, but also on their popularity and user-friendliness.

4.2 Conclusions

In this dissertation, we presented two different examples of the overestimation of the PR made by OR, in cross-sectional studies with frequent (>10%) outcomes.

Even if both the OR and the RR or PR might be appropriate to understand the direction of a certain measure of association, in a study whose aim is to obtain unbiased estimates, using OR instead of RR/PR might lead to misinterpretations and potential consequences in terms of health communication, policymaking, medical decisions, and, lastly, human health. OR should not be interpreted as RR or PR in cross-sectional studies whose outcomes are not rare. When the outcome’s frequency is >10%, robust Poisson models could be a viable alternative to logistic regression models, avoiding the non-convergence of log-binomial models.

However, there is neither consensus nor a standardized approach of the process of selection of the most suitable model. The type of outcome variable (for instance, binary or not), the aim of the study (the need or not of unbiased estimates), its expected epidemiological/clinical interpretation, the fulfilment of the assumptions for causal inference, and the fulfilment of the assumptions of the available statistical models should guide this complex decision. Several statistical measures and processes could be used to support the models’ selection, including comparing models' estimates, their standard-errors, 95%CI, and AIC values, but no single value can serve as the sole criterion to determine the most appropriate model to choose.

More studies should be conducted to compare different alternatives to logistic regression models (like modified Cox, quasi-Poisson, negative binomial, and probit models), not only within cross-sectional studies, but also across longitudinal studies and clinical trials. A standard approach to the estimation of measures of association involving frequent binary outcomes should be established.

Bibliographic references

[1] Jr. Hosmer, S. Lemeshow, R.X. Sturdivant, *Applied Logistic Regression*, 3rd edition. Hoboken, NJ: John Wiley & Sons, 2013.

[2] T.L. Lash, T.J. VanderWeele, S. Haneuse, K.J. Rothman, *Modern Epidemiology*. Lippincott Williams & Wilkins, 2020.

[3] A. R. Tamhane, A. O. Westfall, G. A. Burkholder, and G. R. Cutter, "Prevalence odds ratio versus prevalence ratio: choice comes with consequences," (in eng), *Stat Med,* vol. 35, no. 30, pp. 5730-5735, Dec 30 2016, doi: 10.1002/sim.7059.

[4] D. Celentano and M. Szklo, *Gordis Epidemiology*. Elsevier, 2020.

[5] B. A. F. Martinez, V. B. Leotti, G. d. S. e. Silva, L. N. Nunes, G. Machado, and L. G. Corbellini, "Odds Ratio or Prevalence Ratio? An Overview of Reported Statistical Methods and Appropriateness of Interpretations in Cross-sectional Studies with Dichotomous Outcomes in Veterinary Medicine," (in English), *Frontiers in Veterinary Science,* Original Research vol. 4, 2017-November-10 2017, doi: 10.3389/fvets.2017.00193.

[6] M. J. Knol, S. Le Cessie, A. Algra, J. P. Vandenbroucke, and R. H. Groenwold, "Overestimation of risk ratios by odds ratios in trials and cohort studies: alternatives to logistic regression," (in eng), *Cmaj,* vol. 184, no. 8, pp. 895-9, May 15 2012, doi: 10.1503/cmaj.101715.

[7] M. R. Petersen and J. A. Deddens, "A comparison of two methods for estimating prevalence ratios," (in eng), *BMC Med Res Methodol,* vol. 8, p. 9, Feb 28 2008, doi: 10.1186/1471-2288-8-9.

[8] P. McCullagh, *Generalized Linear Models*. Routledge, 2019.

[9] A. J. Barros and V. N. Hirakata, "Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio," (in eng), *BMC Med Res Methodol,* vol. 3, p. 21, Oct 20 2003, doi: 10.1186/1471-2288-3-21.

[10] J.L. Fleiss, B. Levin, M.C. Paik, *Statistical Methods for Rates and Proportions*. John Wiley & Sons, 2003.

[11] T. Williamson, M. Eliasziw, and G. H. Fick, "Log-binomial models: exploring failed convergence," *Emerging Themes in Epidemiology,* vol. 10, no. 1, p. 14, 2013/12/13 2013, doi: 10.1186/1742-7622-10-14.

[12] J. E. Cavanaugh and A. A. Neath, "The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements," *WIREs Computational Statistics,* vol. 11, no. 3, p. e1460, 2019/05/01 2019, doi: https://doi.org/10.1002/wics.1460.

[13] I. Manisalidis, E. Stavropoulou, A. Stavropoulos, and E. Bezirtzoglou, "Environmental and Health Impacts of Air Pollution: A Review," (in English), *Frontiers in Public Health,* Review vol. 8, 2020-February-20 2020, doi: 10.3389/fpubh.2020.00014.

[14] World Health Organization. WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide: WHO Regional Office for Europe; 2021.

[15] H. Kim, W.-H. Kim, Y.-Y. Kim, and H.-Y. Park, "Air Pollution and Central Nervous System Disease: A Review of the Impact of Fine Particulate Matter on Neurological Disorders," (in English), *Frontiers in Public Health,* Review vol. 8, 2020-December-16 2020, doi: 10.3389/fpubh.2020.575330.

[16] H. Huang *et al.*, "Long-term ambient air pollution exposure and DNA methylation of peripheral brain-derived neurotrophic factor promoter," *Ecotoxicology and Environmental Safety,* vol. 244, p. 114061, 2022/10/01/ 2022, doi: https://doi.org/10.1016/j.ecoenv.2022.114061.

[17] I. Braithwaite, S. Zhang, J. B. Kirkbride, D. P. J. Osborn, and J. F. Hayes, "Air Pollution (Particulate Matter) Exposure and Associations with Depression, Anxiety, Bipolar, Psychosis and Suicide Risk: A Systematic Review and Meta-Analysis," (in eng), *Environ Health Perspect,* vol. 127, no. 12, p. 126002, Dec 2019, doi: 10.1289/ehp4595.

[18] E. Borroni, A. C. Pesatori, V. Bollati, M. Buoli, and M. Carugno, "Air pollution exposure and depression: A comprehensive updated systematic review and meta-analysis," *Environmental Pollution,* vol. 292, p. 118245, 2022/01/01/ 2022, doi: https://doi.org/10.1016/j.envpol.2021.118245.

[19] Y. Zeng, R. Lin, L. Liu, Y. Liu, and Y. Li, "Ambient air pollution exposure and risk of depression: A systematic review and meta-analysis of observational studies," (in eng), *Psychiatry Res,* vol. 276, pp. 69-78, Jun 2019, doi: 10.1016/j.psychres.2019.04.019.

[20] Q. Liu *et al.*, "Association between particulate matter air pollution and risk of depression and suicide: a systematic review and meta-analysis," *Environmental Science and Pollution Research,* vol. 28, no. 8, pp. 9029-9049, 2021/02/01 2021, doi: 10.1007/s11356-021-12357-3.

[21] T. Trushna, V. Dhiman, D. Raj, and R. R. Tiwari, "Effects of ambient air pollution on psychological stress and anxiety disorder: a systematic review and meta-analysis of epidemiological evidence," (in eng), *Rev Environ Health,* vol. 36, no. 4, pp. 501-521, Dec 20 2021, doi: 10.1515/reveh-2020-0125.

[22] P. Cuijpers *et al.*, "Impact of climate events, pollution, and green spaces on mental health: an umbrella review of meta-analyses," (in eng), *Psychol Med,* vol. 53, no. 3, pp. 638-653, Feb 2023, doi: 10.1017/s0033291722003890.

[23] A. L. Stewart *et al.*, "Functional status and well-being of patients with chronic conditions. Results from the Medical Outcomes Study," (in eng), *Jama,* vol. 262, no. 7, pp. 907-13, Aug 18 1989.

[24] A. L. Stewart, R. D. Hays, and J. E. Ware, Jr., "The MOS Short-form General Health Survey: Reliability and Validity in a Patient Population," *Medical Care,* vol. 26, no. 7, 1988. [Online]. Available:https://journals.lww.com/lww-medicalcare/Fulltext/1988/07000/The_MOS_Short_form_General_Health_Survey_.7.aspx.

[25] D. M. Berwick, J. M. Murphy, P. A. Goldman, J. E. Ware, Jr., A. J. Barsky, and M. C. Weinstein, "Performance of a five-item mental health screening test," (in eng), *Med Care,* vol. 29, no. 2, pp. 169-76, Feb 1991, doi: 10.1097/00005650-199102000-00008.

[26] J. Pais-Ribeiro, "Mental Health Inventory: A study of adaptation to Portuguese population," *Psicologia, Saúde & Doenças,* vol. 2, pp. 77-99, 01/01 2001.

[27] A. M. S. Santos and R. F. Novo, "Mental Health Inventory: Sensitivity and Specificity of the Portuguese Version of the MHI-38 and MHI-5," *Psychological Reports,* vol. 123, no. 4, pp. 1452-1469, 2020/08/01 2019, doi: 10.1177/0033294119850490.

[28] H. J. Rumpf, C. Meyer, U. Hapke, and U. John, "Screening for mental health: validity of the MHI-5 using DSM-IV Axis I psychiatric disorders as gold standard," (in eng), *Psychiatry Res,* vol. 105, no. 3, pp. 243-53, Dec 31 2001, doi: 10.1016/s0165-1781(01)00329-8.

[29] P. Cuijpers, N. Smits, T. Donker, M. ten Have, and R. de Graaf, "Screening for mood and anxiety disorders with the five-item, the three-item, and the two-item Mental Health Inventory," (in eng), *Psychiatry Res,* vol. 168, no. 3, pp. 250-5, Aug 15 2009, doi: 10.1016/j.psychres.2008.05.012.

[30] B. Nunes *et al.*, "The first Portuguese National Health Examination Survey (2015): design, planning and implementation," (in eng), *J Public Health (Oxf),* vol. 41, no. 3, pp. 511-517, Sep 30 2019, doi: 10.1093/pubmed/fdy150.

[31] V. Gaio and C. M. Dias. E. N. d. S. Pública. (2021). Association between ambient air pollution exposure and biomarkers of cardiovascular risk: link between the first Portuguese Health Examination Survey and the air quality data.

[32] V. Gaio *et al.*, "Exposure to ambient particulate matter increases blood count parameters with potential to mediate a cardiovascular event: results from a population-based study in Portugal,"

Air Quality, Atmosphere & Health, vol. 14, no. 8, pp. 1189-1202, 2021/08/01 2021, doi: 10.1007/s11869-021-01007-9.

[33] V. Gaio, R. Roquette, A. Monteiro, J. Ferreira, C. Matias Dias, and B. Nunes, "Investigating the association between ambient particulate matter (PM(10)) exposure and blood pressure values: Results from the link between the Portuguese Health Examination Survey and air quality data,", *Rev Port Cardiol,* vol. 42, no. 3, pp. 251-258, Mar 2023, doi: 10.1016/j.repc.2022.02.011.

[34] V. Gaio *et al.*, "PM10 exposure interacts with abdominal obesity to increase blood triglycerides: a cross-sectional linkage study," *European Journal of Public Health,* vol. 32, no. 2, pp. 281-288, 2022, doi: 10.1093/eurpub/ckab190.

[35] A. I. Ribeiro, L. Launay, E. Guillaume, G. Launoy, and H. Barros, "The Portuguese version of the European Deprivation Index: Development and association with all-cause mortality," *PLOS ONE,* vol. 13, no. 12, p. e0208320, 2018, doi: 10.1371/journal.pone.0208320.

[36] J.M. Caldas de Almeida, M. Xavier, *Estudo Epidemiológico Nacional de Saúde Mental: 1º relatório.* Universidade Nova de Lisboa, 2013.

[37] United Nations Educational, Scientific and Cultural Organization (UNESCO). (2012). International Standard Classification of Education: ISCED 2011. UNESCO.

[38] A. J. Santos *et al.*, "Psychological distress – prevalence and associated factors in the portuguese population in 2015: results from the National Health Examination Survey," *Boletim Epidemiológico Observações,* 2017.

[39] International Labour Office. (2012). International Standard Classification of Occupations: ISCO-08. International Labour Organization.

[40] Statistics Portugal. (2023, Sept. 10) *National Statistical Institute - Censos 2021 dissemination platform (definitive results)* [Online]. Available: https://censos.ine.pt/xportal/xmain?xpgid=censos21_produtos&xpid=CENSOS21&xlang=ptdb_censos_2021.html.

[41] A. Buja *et al.*, "Characteristics, processes, management and outcome of accesses to accident and emergency departments by citizenship," (in eng), *Int J Public Health,* vol. 59, no. 1, pp. 167-74, Feb 2014, doi: 10.1007/s00038-013-0483-0.

[42] M. Zélia *et al.*, "Cohort profile: Health trajectories of Immigrant Children (CRIAS)–a prospective cohort study in the metropolitan area of Lisbon, Portugal," *BMJ Open,* vol. 12, no. 10, p. e061919, 2022, doi: 10.1136/bmjopen-2022-061919.

[43] V. Graetz, B. Rechel, W. Groot, M. Norredam, and M. Pavlova, "Utilization of health care services by migrants in Europe-a systematic literature review," (in eng), *Br Med Bull,* vol. 121, no. 1, pp. 5-18, Jan 1 2017, doi: 10.1093/bmb/ldw057.

[44] M. Norredam, A. Krasnik, T. Moller Sorensen, N. Keiding, J. Joost Michaelsen, and A. Sonne Nielsen, "Emergency room utilization in Copenhagen: a comparison of immigrant groups and Danish-born residents," (in eng), *Scand J Public Health,* vol. 32, no. 1, pp. 53-9, 2004, doi: 10.1080/14034940310001659.

[45] S. E. Ruud, P. Hjortdahl, and B. Natvig, "Reasons for attending a general emergency outpatient clinic versus a regular general practitioner - a survey among immigrant and native walk-in patients in Oslo, Norway," (in eng), *Scand J Prim Health Care,* vol. 35, no. 1, pp. 35-45, Mar 2017, doi: 10.1080/02813432.2017.1288817.

[46] L. A. Gimeno-Feliu, M. Pastor-Sanz, B. Poblador-Plou, A. Calderón-Larrañaga, E. Díaz, and A. Prados-Torres, "Overuse or underuse? Use of healthcare services among irregular migrants in a north-eastern Spanish region," (in eng), *Int J Equity Health,* vol. 20, no. 1, p. 41, Jan 20 2021, doi: 10.1186/s12939-020-01373-3.

[47] M. A. Schoevers, M. J. Loeffen, M. E. van den Muijsenbergh, and A. L. Lagro-Janssen, "Health care utilisation and problems in accessing health care of female undocumented immigrants in the Netherlands," (in eng), *Int J Public Health,* vol. 55, no. 5, pp. 421-8, Oct 2010, doi: 10.1007/s00038-010-0151-6.

[48] L. A. Gimeno-Feliu, A. Calderón-Larrañaga, E. Diaz, B. Poblador-Plou, R. Macipe-Costa, and A. Prados-Torres, "Global healthcare use by immigrants in Spain according to morbidity burden, area of origin, and length of stay," (in eng), *BMC Public Health,* vol. 16, p. 450, May 27 2016, doi: 10.1186/s12889-016-3127-5.

[49] S. Dias, A. Gama, M. Cortes, and B. de Sousa, "Healthcare-seeking patterns among immigrants in Portugal," (in eng), *Health Soc Care Community,* vol. 19, no. 5, pp. 514-21, Sep 2011, doi: 10.1111/j.1365-2524.2011.00996.x.

[50] O. Sauzet, M. David, B. Naghavi, T. Borde, J. Sehouli, and O. Razum, "Adequate Utilization of Emergency Services in Germany: Is There a Differential by Migration Background?," (in eng), *Front Public Health,* vol. 8, p. 613250, 2020, doi: 10.3389/fpubh.2020.613250.

[51] C. Tzogiou, S. Boes, and B. Brunner, "What explains the inequalities in health care utilization between immigrants and non-migrants in Switzerland?," *BMC Public Health,* vol. 21, no. 1, p. 530, 2021/03/18 2021, doi: 10.1186/s12889-021-10393-9.

[52] L. M. Coutinho, M. Scazufca, and P. R. Menezes, "Methods for estimating prevalence ratios in cross-sectional studies,", *Rev Saude Publica,* vol. 42, no. 6, pp. 992-8, Dec 2008. [Online]. Available: https://www.scielo.br/j/rsp/a/XdMZzxdGrSTS77Y8ZgpGdWj/?lang=en&format=pdf.

[53] J. Lee, C. S. Tan, and K. S. Chia, "A practical guide for multivariate analysis of dichotomous outcomes," (in eng), *Ann Acad Med Singap,* vol. 38, no. 8, pp. 714-9, Aug 2009.

[54] T. Behrens, D. Taeger, J. Wellmann, and U. Keil, "Different methods to calculate effect estimates in cross-sectional studies. A comparison between prevalence odds ratio and prevalence ratio," (in eng), *Methods Inf Med,* vol. 43, no. 5, pp. 505-9, 2004.

[55] C. Gnardellis, V. Notara, M. Papadakaki, V. Gialamas, and J. Chliaoutakis, "Overestimation of Relative Risk and Prevalence Ratio: Misuse of Logistic Modeling," (in eng), *Diagnostics (Basel),* vol. 12, no. 11, Nov 17 2022, doi: 10.3390/diagnostics12112851.

[56] M. E. Reichenheim and E. S. Coutinho, "Measures and models for causal inference in cross-sectional studies: arguments for the appropriateness of the prevalence odds ratio and related logistic regression," (in eng), *BMC Med Res Methodol,* vol. 10, p. 66, Jul 15 2010, doi: 10.1186/1471-2288-10-66.

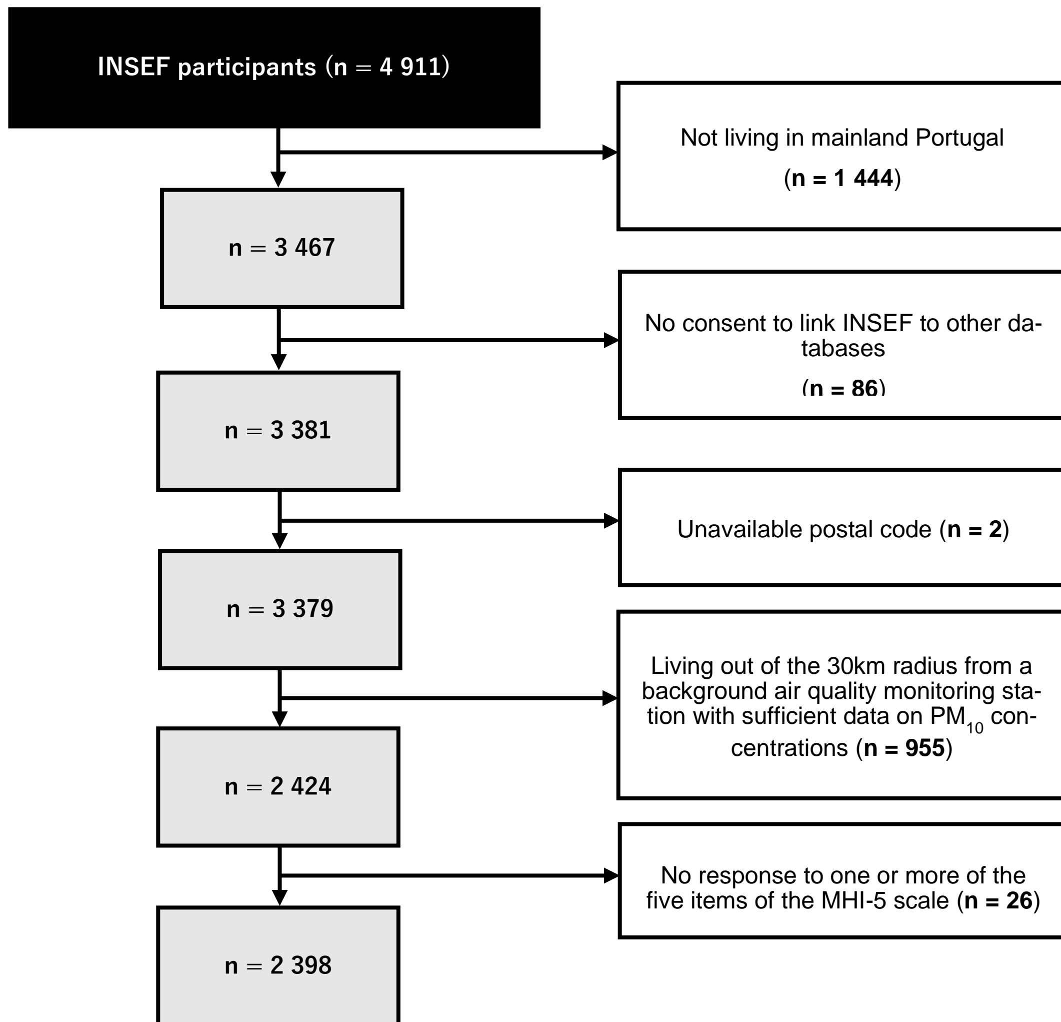
A

SUPPLEMENTARY MATERIAL

A.1 Supplementary Figures

The 5-item Mental Health Inventory
How much of the time in the previous 4 weeks:
1. Have you been a very nervous person?
2. Have you felt so down in the dumps that nothing could cheer you up?
3. Have you felt calm and peaceful?
4. Have you felt downhearted and blue?
5. Have you been a happy person?

Supplementary Figure 1. Items of the 5-item Mental Health Inventory



Supplementary Figure 2. Selection flow diagram for the individuals in study

PM_{10} - particulate matter with an aerodynamic diameter less than or equal to 10 micrometres, INSEF – First Portuguese National Health Survey with Physical Examination, MHI-5 - 5-item Mental Health Inventory.

A.2 Supplementary Tables

Supplementary Table 1. Comparison of the characteristics between the included and the excluded INSEF participants

	All (n=4 911)	**Included** (n=2 398)	**Excluded** (n=2 513)	**p-value/ 95%CI** [$]
Exposure				
Individual allocated 1-year average PM10* (µg/m³) – median (IQR)	18.5 (15.3-19.3)	18.6 (15.3-19.3)	18.5 (16.7-19.1)	(-0.345, 0.178)
Independent variables				
Sex - %				0.805
Female	52.5	52.6	52.0	
Male	47.5	47.4	48.0	
Age group - %				0.245
25-34 years	18.3	18.9	16.3	
35-49 years	34.4	34.1	35.6	
50-64 years	31.4	31.5	30.7	
65-74 years	15.9	15.5	17.4	
Education level* - %				0.024
Low[a]	27.7	26.5	32.2	
Medium[b]	52.9	53.5	50.5	
High[c]	19.4	20.0	17.3	
Employment status* - %				0.441
Employed	61.9	62.3	60.7	
Unemployed	11.3	11.4	10.8	
Other[d]	26.8	26.3	28.5	
Professional occupation* - %				0.236
White-collar[e]	62.2	62.9	59.7	
Blue-collar[f]	37.8	37.1	40.3	
Area-level socioeconomic deprivation terciles* - %				0.512
Low deprivation (T1)	16.9	16.0	20.4	
Moderate deprivation (T2)	33.4	31.6	40.2	
High deprivation (T3)	49.7	52.4	39.4	
Individual allocated 1-year average temperature (ºC) – median (IQR)	15.0 (14.5-16.7)	**15.0 (14.8-16.7)**	**14.9 (14.5-16.2)**	**(-0.440, -0.326)**
Area-level walkability terciles* - %				**<0.001**
Low walkability (T1)	5.3	**1.3**	**20.8**	
Moderate walkability (T2)	18.2	**18.4**	**17.6**	
High walkability (T3)	76.4	**80.3**	**61.6**	
Degree of urbanization- %				0.282
Rural	25.4	28.2	19.2	
Urban	73.6	71.8	80.8	
Outcome variables				

Probable diagnosis of common mental health disorders* - %				0.797
Yes	22.6	22.7	22.1	
No	77.4	77.3	77.9	

INSEF – First Portuguese National Health Survey with Physical Examination, 95%CI - 95% confidence interval. IQR - interquartile range, PM_{10} - particulate matter with an aerodynamic diameter less than or equal to 10 micrometres, T1 - first tercile, T2 - second tercile, T3 - third tercile, MHI-5 - 5-item Mental Health Inventory.
* 2 339 missings in individual allocated 1-year average PM_{10} and PM_{10} exposure terciles, 2 missings in area-level socioeconomic deprivation, 2 missings in area-level walkability, 4 missings in education level, 3 missings in employment status, 476 missings in professional occupation, 53 missings in at least one of the 5 items composing the MHI-5 (instrument whose score was categorized in two to obtain the categorical variable "probable diagnosis of common mental health disorders", being a score ≤52 indicative of a probable diagnosis).
$ p-values or 95%CI were presented according to the test used to compare proportions (Chi-squared test) or medians (robust confidence intervals for generalized Hodges-Lehmann median differences), respectively. Results in bold are those with statistically significant differences between individuals included and excluded of the study.
a: Less than secondary: levels 0–2 of the International Standard Classification of Education 2011 .
b: Secondary: levels 3–4 of the International Standard Classification of Education 2011 .
c: Professional or higher: levels 5–8 of the International Standard Classification of Education 2011.
d: Other: students, retired, stay-at-home parents, other.
e: White-collar: managers, professionals, technicians and associate professional, clerical support workers, and services and sales workers (according to the International Standard Classification of Occupations).
f: Blue-collar: skilled agricultural workers, craft and related trades workers, plant and machine operators, and elementary occupations (according to the International Standard Classification of Occupations).
All the estimates were weighted to account for different selection probabilities and population distribution.

A.3 Annex 1 - Odds ratio, prevalence ratio and relative risk

In cross-sectional studies, the association between an exposure, treatment or health determinant and a binary outcome/event is assessed through the odds ratio (OR), the odds of the disease occurrence between exposed and unexposed. The prevalence ratio (PR) is the equivalent to the relative risk or risk ratio (RR), used in prospective studies, but applied to cross-sectional studies and, thus, prevalence (instead of incidence like in prospective studies).

RR is defined as the ratio of the probability of the outcome in those exposed to the probability of the outcome in the unexposed. Therefore, as a measure, RR estimates how much higher (or lower) is the probability of developing the outcome for those exposed to the risk factor compared to the unexposed. The data of such an investigation can be summarized in a 2-by-2 table (Table A.3.1).

Table A.3.1. Notation for entries in a 2-by-2 table

	Event	**No event**	**Total**
Exposed	a	b	a + b
Unexposed	c	d	c + d
Total	a + c	b + d	n = a + b + c + d

The RR (and PR) can be estimated by the equation:

$$RR = \frac{\frac{a}{a+b}}{\frac{c}{c+d}} = \frac{a \times (c + d)}{c \times (a+b)}$$

The OR can be estimated by the equation:

$$OR = \frac{\frac{\frac{a}{a+b}}{\frac{b}{a+b}}}{\frac{\frac{c}{c+d}}{\frac{d}{c+d}}} = \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a \times d}{b \times c}$$

When the probability of the outcome is low, the difference between the odds ratio and relative risk is negligible, as expressed by the formula:

$$RR = \frac{a \times (c + d)}{c \times (a+b)} \sim \frac{a \times d}{b \times c}$$

A.4 Annex 2 - Huber sandwich estimator

In regression and time-series modelling, models make use of the assumption that errors have the same variance across all observation points (homoscedasticity). When this is not the case, errors are said to be heteroskedastic, and this behaviour will be reflected in the residuals estimated from a fitted model.

Heteroskedasticity-consistent standard errors are used to allow the fitting of a model that does contain heteroskedastic residuals. The first such approach was proposed by Huber (1967).

The Poisson regression with robust standard errors specify that the variance-covariance matrix neither assumes $E(y_i) = Var(y_i)$, nor requires $Var(y_i)$ to be constant across all i. This alternative is known to underestimate the true variability with moderately sized samples.

The sandwich idea is to estimate the second partial derivative of the log likelihood function [L(θ) = $\sum_{i=1}^{n} \log fi \ (Yi|\hat{O})$], with respect to θ, L"(θ), directly from the sample data, as L"(Ô).The Huber sandwich estimator is represented by $(-A)^{-1} B(-A)^{-1}$, where A= L"(Ô) and B = $\sum_{i=1}^{n} gi \ (Yi \ |\hat{O})^{T} \ gi \ (Yi|\hat{O})$.

2023 **LARA PINHEIRO GUEDES**

LIMITATIONS OF LOGISTIC REGRESSIONS TO ESTIMATE MEASURES OF ASSOCIATION FOR BINARY HEALTH OUTCOMES

NOVA SCHOOL OF SCIENCE & TECHNOLOGY