RESIDENTIAL EXPOSURE TO AGRICULTURAL POLLUTANTS

THE USE OF ADMINISTRATIVE DATA TO EXPLORE THE HEALTH OF THE DUTCH RURAL POPULATION

MARIANA SIMÕES

RESIDENTIAL EXPOSURE TO AGRICULTURAL POLLUTANTS

The use of administrative data to explore the health of the Dutch rural population

Mariana Simões

Residential exposure to agricultural pollutants:
The use of administrative data to explore the health of the Dutch rural population

PhD Thesis. Institute for Risk Assessment Sciences,
Utrecht University, Utrecht, The Netherlands

Author: Mariana Simões

The studies of this thesis were funded by the Ministry of Health, Welfare and Sport (VWS), in the context of the Policy Advisory on Plant Protection Products and by internal funding of the Institute of Risk Assessment Sciences, Utrecht University, which financial support is gratefully acknowledged.

ISBN: 978–94–93353–15–2
Cover and Design: IsaSilva.com
Printed by: Proefschrift-AIO

RESIDENTIAL EXPOSURE TO AGRICULTURAL POLLUTANTS

The use of administrative data to explore the health of the Dutch rural population

De agrarische sector en omwonenden: het gebruik van administratieve data om potentiële blootstellingen en gezondheidseffecten in kaart te brengen

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

dinsdag 31 oktober 2023 des middags te 4.15 uur

door

Mariana de Oliveira Lobo Fernandes da Silva Simões
geboren op 22 februari 1985
te Lisboa, Portugall

Promotor:
Prof. dr. R.C.H. Vermeulen

Copromotoren:
Dr. N.A.H. Janssen
Dr. A. Huss

Beoordelingscommissie:
Prof. dr. V. Geissen
Prof. dr. ir. J. Kromhout
Prof. dr. A.M. May
Dr. Stafoggia
Prof. dr. J.A. Stegeman

CONTENTS

CHAPTER 1

GENERAL INTRODUCTION

There has been increasing interest in the health effects of residential exposure to agricultural pollutants in the last years. This is evidenced by increased public health concern, political involvements and scientific research. Rural populations are the most exposed to these agricultural pollutants given their residential proximity to their sources, namely livestock farms and crop fields where pesticides are applied. Epidemiological studies have been applying different approaches with regard to study design and assessment of exposures, outcomes and potential confounders. Because environmental exposures usually result in small but often relevant public health effects, it is important to conduct studies with sufficient power to detect these subtle effects. This thesis compiles large epidemiological studies based on administrative databases to explore associations between exposure to agricultural pollutants, namely pesticides used in crops and livestock farming emissions, and health of the Dutch rural population.

HEALTH OF THE DUTCH RURAL POPULATION

According to the World Health Organization, health is "is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity".[1] Health is not only a fundamental human right, it is also a valuable backbone of a country's socio-economic development, making it one of the most important features of a population to monitor.

Increased urbanization during the last century has shifted demographic characteristics of populations between urban and rural settings, ultimately affecting population health. Rapid increase in urbanization has resulted in higher population density, higher levels of pollution and noise in cities, and socioeconomic inequities in urban areas.[2] These have in turn contributed to higher incidence and prevalence of infectious diseases, chronic diseases and psychological distress, which have a heterogenous distribution across population subgroups.[2,3] Despite this, urban populations in high income countries remain, generally speaking, at an advantage, in terms of their overall health and wellbeing, compared to their counterparts in more rural areas.[4] Relatively poorer health among the rural population has been associated with overall lower socioeconomic position (lower levels of income and education), reduced access to relevant information sources, barriers to, or underservice of, healthcare and poorer effectiveness of public health policies. Furthermore, rural populations are often exposed to a range of emissions from agricultural activities, such as livestock farming or application of pesticides in crops. These agricultural emissions in the home environment can be characterized by short periods of exposure to high concentration levels (for example, a pesticide spraying event) or by prolonged exposure to low concentration levels. Ulti-

mately, both can contribute to health problems and, since a large number of people is exposed, there can be impactful consequences for public health and society.
Interestingly, and in contradition to the discussion above, a 2020 report from Statistic Netherlands (CBS) showed that while Dutch rural residents experienced lower material wealth and higher distances to important facilities (such as primary schools), they seemed to enjoy better health, as evidenced by better perceived health, better overall wellbeing and less chronic diseases than found in urban populations.[5] This observation may be due to established advantages of living in rural areas, such as stronger social cohesion, less stressful lifestyle and proximity to nature.[6] Notwithstanding, as the Netherlands is a very densely populated country, an important proportion of people live in close proximity to livestock farms and crop fields. For example, about 30% of all residences in the Netherlands are located within 250m of a crop field. Therefore, even if the effects of agricultural exposures are small, their consequences and impacts can be important because of the large number of exposed individuals. Take birth weight as an example: a small downward shift of the distribution of birth weight in a population of pregnant women exposed to a certain environmental risk factor translates into a substantial increase of the proportion of babies born with low birth weight. This, in turn, has been shown to be associated with newborn and infant mortality and morbidity, poor cognitive development and increased risk of chronic diseases in adulthood.[7]

ADMINISTRATIVE DATA AND THEIR USE IN EPIDEMIOLOGY

Administrative data refers to structured data that is routinely collected, usually by governments or other (private) organizations, and encompasses registration, record keeping and transactions keeping. It is not collected for statistical nor research purposes, but rather to manage services and monitor their performance, such as administering taxes and benefits or registering deaths and births and respond to their legal requirements. One type of administrative data is *administrative registries*, which record changes at unit level. Such data can become an extremely useful tool in epidemiological studies by offering important advantages such as virtually complete population coverage, detailed information on demographic characteristic (age, sex, marital status, income, residential history) and information on *administrative units* (individuals or groups such as companies). There are different levels of accessibility to administrative data, shaping both its utility and the quality of research that can be done using these data. Much administrative data is *public data*, that is data that were produced by public bodies, but that may not be easily accessible depending on whether they contain sensitive information. *Microdata*, for instance, is data at the level of individuals,

subject to anonymization rules that permit their use by researchers under strict conditions. *Open data*, on the other hand, are generated by both public and non-public organizations and are published on open portals. These data are thus easily accessible and can be re-used freely but may not have information at the detailed level that is required for research.
Given that administrative data are often not collected with a health or research purpose in mind, they generally provide less detail on clinical information and little or no information on lifestyle factors. In certain settings that guarantee privacy and compliance with the General Data Protection Regulation (GDPR), such data can be linked to other data sets allowing the study of a broad range of outcomes and exposures. These linkages reduce time and costs of research projects by providing insights to possible associations and identifying small but relevant effects before setting up huge target studies. This is especially important for policy-relevant research, as it provides evidence-based information.

THE WORK OF THIS THESIS

This work used nationwide administrative databases to explore associations between residential exposure to agricultural pollutants, namely livestock emissions and pesticides, and a number of health outcomes in the Netherlands. All studies considered made use of microdata to determine the demographic characteristics of individuals and health outcomes. This strategy enabled studies to include large numbers of individuals. Large studies, such as these, preclude taking personal measurements or measurements in the home environment to assess levels of exposure. Accordingly, exposure assessment relied upon modelling approaches for which input information on agricultural exposures was available, either from open data or from public data, specifically purposed for research.
The general objective of this thesis was, using extant data, to identify potential associations between proxies for exposure to agricultural pollutants and various health outcomes. The aim was to generate leads for more targeted research on relevant (groups of) pollutants.
In the light of previous reports of adverse respiratory problems in people living near livestock farms, the goal of **chapter 2** was to assess whether proximity to farms raising certain types of livestock animals (which likely produce specific mixtures of agricultural pollutants) was associated with respiratory mortality, a health outcome that had yet to be explored in this context. The Geographic Information System for Agricultural Holdings (Geografisch Informatiesysteem Agrarische Bedrijven, GIAB) database, allowed

the computation of exposure proxies for the emissions from livestock farms, taking into account the number and type of animals being raised. A nationwide adult administrative cohort living in rural areas was built and included information on demographics and residential history at individual level. The cohort was linked to the death registry to explore how living near livestock farms could be related to overall respiratory mortality and to chronic lower respiratory disease and pneumonia mortality in particular.

Subsequent chapters aim to shed light upon inconsistent results in literature regarding the associations between residential exposure to pesticides and health. A cohort, similar to that of chapter 2, was used in **chapter 3**, which explores potential associations between residential proximity to specific types of crops (employing crop-specific mixtures of pesticides) and several cause-specific mortality endpoints. Endpoints included several types of cancer and respiratory, cardiovascular and neurologic diseases. In this chapter, the Landelijk Grondgebruik Nederland (LGN) land use raster maps from 1995, 1997, 2000 and 2004 were used to compute the average area of specific crop groups around residences. This was used as a proxy for residential exposure to the mixture of pesticides used in those crops.

In **chapter 4**, these areas were computed using the annual Basisregistratie Gewaspercelen (BRP) vector crop maps from 2009 to 2013, which provided better spatial and temporal resolution for the computation of the exposure proxies. This chapter explored the association between residential proximity to specific crops and self-reported depression and perceived health of a subset of participants of the Health Monitor 2012 living in rural areas.

Finally, the association between residential exposure to specific active ingredients and birth outcomes was studied in **chapter 5**. Exposure proxies computed previously and information on the dosage of pesticides used in specific crops (as reported in farmer surveys), were combined to compute the exposure proxies used for this chapter. Using the birth registry compiled by Perined, the chapter starts by investigating active ingredients for which there is some evidence of reproductive and developmental adverse effects. Chapter 5 then further explores other potentially relevant associations using a variable selection method.

Given that the exposure proxies in **chapters 2** to **4** reflected mixtures of pollutants, the aim in these chapters was to give insights into which sorts of livestock farms or which types of crops could be associated with adverse health effects. Conversely, **chapter 5**, specific active ingredients were investigated in a hypothesis testing approach, with the aim of understanding whether residential exposures might be associated with relevant (adverse) effects. Other active ingredients were investigated in an exploratory approach, the goal being to identify potential risks from unsuspected active ingredients.

REFERENCES

1. Constitution of the World Health Organization. Constitution of the World Health Organization [Internet]. [cited 2022 Oct 6]. Available from: https://www.who.int/about/governance/constitution
2. Commission E, Centre JR. The future of cities : opportunities, challanges and the way forward. Aurambout J, Vandecasteele I, Baranzelli C, Siragusa A, editors. Publications Office; 2019.
3. Nieuwenhuijsen M, Khreis H. Integrating Human Health into the Urban Development and Transport Planning Agenda: A Summary and Final Conclusions BT – Integrating Human Health into Urban and Transport Planning: A Framework. In: Nieuwenhuijsen M, Khreis H, editors. Cham: Springer International Publishing; 2019. p. 707–18. Available from: https://doi.org/10.1007/978–3–319–74983–9_34
4. Rydin Y, Bleahu A, Davies M, Dávila JD, Friel S, De Grandis G, et al. Shaping cities for health: complexity and the planning of urban environments in the 21st century. Lancet (London, England) [Internet]. 2012/05/30. 2012 Jun 2;379(9831):2079–108. Available from: https://pubmed.ncbi.nlm.nih.gov/22651973
5. Statistics Netherlands (CBS). Regionale Monitor Brede Welvaart toont grote verschillen stad en platteland [Internet]. [cited 2022 Oct 4]. Available from: https://www.cbs.nl/nl-nl/nieuws/2020/49/regionale-monitor-brede-welvaart-toont-grote-verschillen-stad-en-platteland
6. Eurostat. Statistics on rural areas in the EU [Internet]. 2018 [cited 2022 Oct 6]. Available from: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Archive:Statistics_on_rural_areas_in_the_EU
7. Organization WH. Global nutrition targets 2025: low birth weight policy brief [Internet]. Geneva PP – Geneva: World Health Organization; Available from: https://apps.who.int/iris/handle/10665/149020

CHAPTER 2

RESIDENTIAL PROXIMITY TO LIVESTOCK ANIMALS AND MORTALITY FROM RESPIRATORY DISEASES IN THE NETHERLANDS: A PROSPECTIVE CENSUS-BASED COHORT STUDY

Mariana Simões, Nicole Janssen, Dick J.J. Heederik, Lidwien A.M. Smit, Roel Vermeulen, Anke Huss

Published: Environ Int. 2022 Mar;161:107140.
doi: 10.1016/j.envint.2022.107140

ABSTRACT

Background: There is increasing evidence of associations between residential proximity to livestock farms and respiratory morbidity, but less is known about potential effects on respiratory mortality among residents.
Objectives: We aimed to assess potential associations between respiratory mortality and residential proximity to (intensive) livestock farming.
Methods: In DUELS, a national census-based cohort, we selected all inhabitants from rural and semi-urban areas of the Netherlands, aged ≥30 years and living at the same address for five years up to baseline (2004). We followed these ~4 million individuals for respiratory mortality (respiratory system diseases, chronic lower respiratory diseases, pneumonia) from 2005 to 2012. We computed the average number of cattle, pigs, chicken, and mink present in 500m, 1000m, 1500m and 2000m of each individual's residence in the period 1999–2003. Analyses were conducted using Cox proportional hazards regression, adjusting for potential confounders at individual and neighborhood level.
Results: We found evidence that living up to 2000m of pig farms was associated with respiratory mortality, namely from chronic lower respiratory diseases, with Hazard Ratios ranging from 1.06 (1.02, 1.10) in people living close to low numbers (<median number of animals) of pigs in 1000m and 1.18 (1.13, 1.24) in those living near high numbers (≥median) of pigs in 2000m. We also found indications of higher pneumonia mortality in people living near mink farms.
Conclusion: Our results are in line with previous findings of adverse respiratory effects in people living near livestock farms. Little is known about the physical, chemical, and biological exposures leading to respiratory morbidity and mortality warranting further explorations of air contaminants in the vicinity of livestock farms.

INTRODUCTION

Livestock farms have been shown to be major sources of zoonotic pathogens and air pollutants including particulate matter, endotoxins, ammonia, volatile organic compounds and greenhouse gases[1–7]. Despite concentrations of these compounds being considerably lower in ambient air compared to inside farms, some studies have shown that residents living near farms are at increased risk of respiratory health effects, such as exacerbation of chronic obstructive pulmonary disease (COPD) and symptoms indicative of asthma (wheezing), decreased lung function, increased respiratory symptoms and pneumonia[8–12]. Heterogeneity of outcome definitions and limited evidence of exposure-response relationships do not allow firm conclusions about causality of exposure to livestock farms' emissions and adverse respiratory outcomes in residents living near farms[13].

The Netherlands is a densely populated country with a large livestock industry and currently witnesses a debate about the future of intensive animal farming, namely regarding loss of biodiversity due to nitrogen deposition, sustainability of farming practices, animal welfare, and possible adverse public health effects[14]. Particularly effects on the health of residents living near (intensive) livestock farms have received considerable attention in recent years in the Netherlands after the emergence of antimicrobial resistant bacteria (MRSA and ESBLs) starting in 2005, the Q fever epidemic of 2007–2010, and SARS-CoV-2 outbreaks in mink farms in 2020[15–19]. Several studies conducted after the Q-fever epidemic consistently showed increased risks for pneumonia in people living near goat farms and, to a lesser extent, poultry farms, although the exact causal mechanisms are still under investigation[1,20–22]. These studies also showed that people living near livestock farms had poorer lung function and higher risk of COPD exacerbations, while, in contrast, a lower prevalence of asthma, allergies, and COPD was observed[8,23]. Respiratory problems were weakly associated with living in the vicinity of cattle, pigs and mink[8,9,24]. Most of these studies focused on incidence of respiratory diseases and symptoms using predominantly data from general practitioners in two rural regions of the Netherlands where density of intensive farming is high. To the best of our knowledge, no studies have investigated associations between respiratory diseases mortality and proximity to livestock farms at a nationwide level.

In this paper, we aimed to investigate the association between living near cattle, pigs, chicken, and mink and mortality due to respiratory system diseases in general, and chronic lower respiratory diseases and pneumonia specifically. Using historical data on the location of farms and registry data, we followed the entire rural Dutch population for respiratory mortality from 2005 to 2012.

METHODS

Study population

The Dutch Environmental Longitudinal Study (DUELS) is an administrative cohort that includes all inhabitants aged 30 years or older on 01–01–2004 and registered in the Dutch population registry (GBA – Gemeentelijke Basisadministratie Persoonsgegevens); registration in GBA is mandatory in the Netherlands. The cohort was built integrating data from several databases from Statistics Netherlands (Centraal Bureau voor de Statistiek, CBS) including mortality, individual characteristics, residential history, and neighborhood characteristics. In this study, we excluded persons who lived within 2000m of the border with Germany or Belgium (for whom we were unable to compute the livestock specific exposure reliably), persons who changed address in the five years prior to enrolment, and persons living in the more urbanized areas of the Netherlands (≥1500 addresses per km^2, at neighborhood level) (Figure 1).

Residential proximity to livestock farms

We determined the presence of livestock farms located in the vicinity of residences using the Geographic Information System for Agricultural Holdings (Geografisch Informatiesysteem Agrarische Bedrijven, GIAB) database, which provides spatial information on agricultural land use, namely data on Dutch agricultural holdings, obtained through the annual agricultural census by CBS and the Netherlands Enterprise Agency (RVO). Data on farm type, farm size and average annual numbers per animal group, among others, are linked to the main farm location of each agricultural holding. These data were available for the years 1999 to 2003. For each year we computed the number of cattle, pigs, chicken, and mink present within buffers of 500, 1000, 1500 and 2000m around each residence in the Netherlands as proxies for farm exposure. We averaged the number of (specific) animals over the exposure period (1999–2003) and categorized the obtained exposure variables into "no animals" (0 animals within buffer), "low" (< median number of animals within buffer) and "high" (≥ median number of animals within buffer) – see Table S2.1 for cut-off points (medians). The types of farms and buffer sizes were chosen based on results from the "Livestock Farming and Neighbouring Residents' Health" (VGO) project for which some evidence for associations to respiratory health was determined[8].

Cohort follow-up and mortality endpoints

Each individual in the cohort was assigned five years of exposure period, from 01–01-1999 to 31–12–2003. We included a one-year lag period (01–01–2004 to 31–12–2004), to allow for a latency period. Follow-up started on 01–01–2005 and terminated at the end of the follow-up period (31–12–2012), at the time of death or when individuals were lost to follow-up, whichever came first. Data on mortality due to respiratory system diseases (RSD), chronic lower respiratory diseases (CLR) and pneumonia (PNE) were retrieved from the mortality database from CBS, where primary causes of death are classified according to the International Classification of Diseases, 10th revision (ICD-10; Figure 2)[25].

Statistical analysis

We studied the association between respiratory mortality and number of livestock present within our *a priori* defined buffer sizes using age-stratified (one-year age strata) Cox proportional hazards regression, including all considered livestock animal species in the models (Figure 2). We applied a combination of increasingly adjusted models by adding potential confounders at individual and neighborhood level and compared residents living within 500, 1000, 1500 and 2000m from livestock farms ('exposed') to residents that did not have livestock farms within those distances from their residences ('unexposed'):

- *basic model*, adjusted for sex
- *intermediate model*, basic model further adjusted for origin (based on the mother's country of birth or, if unavailable, father's country of birth), marital status and standardized household income (an individual socioeconomic indicator adjusted for differences in household size and composition)
- *full model*, intermediate model further adjusted for socioeconomic position (SEP) as defined by the SCP (Sociaal en Cultureel Planbureau; a social status score taking into account average income, percentage of people with a low income, percentage of people with a low education and percentage of people not working in a postal code area)26 at four-digit postcode level, urbanization degree at neighborhood level as defined in the "Wijk- en buurtkaart" (neighborhood maps) from 1999 and 2003, the proportion of low educated residents in the neighborhood in 2007 and ambient Particulate Matter < 2.5 µm in diameter ($PM_{2.5}$) and nitrogen dioxide (NO_2) levels, as estimated by land use regression models using data for the year 2010[27].

All potential confounders were used as categorical variables (see Table 1 for classes).

Sensitivity analyses

We conducted six sensitivity analyses. First, we excluded people who worked in agriculture for at least one year in the period 1999 to 2003, to assess the influence of possible occupational exposure on the estimates. Second, we restricted analyses to people living in neighborhoods with less than 1000 addresses per km^2 to assess potential bias from a semi-urban environment. Third, we combined the two previous sensitivity analyses, since most farmers will live in the more rural areas and in or near farms. Fourth, we reran analyses using redefined exposure variables' categories, where we assigned a zero if the farm had less than a minimum number of animals, as done previously in the VGO study (Supplementary material, S1), so not to assign people living near stables with only a few hobby animals or a farm with an obsolete license with a few animals contributing to the "exposed" categories. Fifth, we conducted an analysis on the VGO study region in the east of Noord-Brabant and the North of Limburg only. Sixth, for completeness, we ran analyses including equines (horses and donkeys), sheep, and other poultry (mainly turkey and ducks) since data was available, although we had no prior reason for investigation. Furthermore, because (intensive) livestock farming is a regional activity in the Netherlands, we conducted stratified analyses by the four major socio-economic regions (according to the Nomenclature des Unités Territoriales Statistiques, NUTS1), followed by a random effects meta-analysis to assess heterogeneity (I^2 statistic) of regional estimates. Finally, to identify potential residual bias we conducted negative control analyses using colon cancer, bladder cancer, liver cirrhosis and alcoholic liver disease mortality as the endpoints. These mortality endpoints are strongly associated with smoking and/or other unhealthy lifestyle behaviors, namely alcohol consumption, but unlikely related to environmental pollution from livestock farming[28–31]. For completeness, we also explored associations to 'all cause' and 'non-accidental' mortality.

Software

The geospatial assignment of exposure variables was conducted in R version 3.6.1 (2019–07–05), using the "sf" and "rgdal" packages. Statistical analyses were performed in R version 3.6.2 (2019–12–12), within a secured remote access environment of CBS.

RESULTS

We included 4,040,845 persons in our analyses, of which a total of 412,532 (10.2%) participants died, including 40,131 (1.0%) from RSD, 19,054 (0.5%) from CLR and 15,189 (0.4%)

from PNE during follow-up (2005–2012). In this study population there were 26,309 (0.7%) persons lost to follow-up. There were 2,203,650 (54.5%), 3,525,961 (87.3%), 3,884,771 (96.1%) and 3,993,150 (98.8%) people exposed to at least on type of animal in the 500, 1000, 1500 and 2000m buffer, respectively. We observed few unexposed persons in the larger buffers for the most ubiquitous types of farms in the Netherlands (namely cattle, S2). Table 1 describes the demographic characteristics of the study population and the exposed population within each buffer. A table reporting the number of (un)exposed people for each mortality endpoint status can be found in S2.

People living near pigs presented consistently higher risk for all mortality endpoints (RSP, CLR and PNE) across the four buffers (Table 2). We saw no clear pattern indicating that living near cattle, chicken, or mink was associated with these mortality endpoints. Nevertheless, we observed some elevated risk estimates of PNE in people living near a high number of cattle in the 500m and 1500m and 2000m buffers, accompanied by a lower risk of CLR in the two larger buffers in both exposure categories. People living within 1500 and 2000m of a high number of mink showed higher risk for PNE. In general, estimates obtained in the main analyses were robust to sensitivity analyses (S4), except for the analysis in the VGO areas where we observe weak, absent or even inverse associations compared to the main analysis. Stratified analysis showed that, generally, heterogeneity of the regional estimates was not high (I^2<75%) (S5). The negative control analyses showed no associations between colon cancer, bladder cancer, liver cirrhosis, and alcoholic liver disease mortality and residential proximity to livestock animals (S6).

DISCUSSION

We investigated the association between living near livestock animals and mortality from respiratory system diseases, chronic lower respiratory diseases and pneumonia using a national administrative cohort. We found higher risk of mortality due to all three respiratory endpoints in people living near farms raising pigs, observing Hazard Ratios above unity consistently across all buffer sizes and a tendency for increasing risks in people living in proximity to higher as compared to lower animal counts. In addition, generally homogeneous results across the Netherlands were observed. There was no clear evidence of associations for the other animals, although several increased risk estimates also emerged for associations between living near cattle and mink farms and risk of PNE mortality.

We conducted a nationwide prospective census-based cohort study using a large non-urban study population of over 4 million individuals for which we objectively assessed individual proxies for livestock farm exposure and included all major groups of animals raised in the Netherlands. By including the entire rural and semi-urban Dutch

population aged ≥30 years, not only did we preclude recall and selection bias for the exposure, outcome and considered confounders, but we also conducted, to the best of our knowledge, the largest study on the topic to date. Additionally, most studies on this topic have focused on short-term exposures or have a cross-sectional design. Our long exposure and follow-up periods allowed the study of long-term exposure and potential respiratory health effects.

While access to registry data allowed for the advantages described above, use of these data was accompanied by disadvantages regarding obtaining detailed information about outcomes, exposure and potential confounders. First, we are unable to identify the specific causes of respiratory diseases underlying death. Second, we could not adjust for behavioral or lifestyle factors and relevant risk factors for respiratory mortality endpoints such as active or passive smoking, body mass index (BMI), nutrition, indoor air pollution and underlying comorbidities. Nevertheless, there is no reason to assume that any residual confounding would be present only for the association found between living close to pig farms and respiratory mortality, and not with the other types of animals. Furthermore, our negative control analyses show no indication for strong confounding by smoking and poor lifestyle behaviors, although the potential for some residual confounding cannot be completely dismissed. Of note, other studies on residential proximity to livestock farms and respiratory outcomes conducted in the Netherlands where some of the abovementioned confounders were taken into account showed no appreciable changes in the estimates when compared to more parsimonious models controlling only for age and sex[22,24]. Third, our geographical data on farms pertained to the address of the farm's company which may not correspond to the location where animals were held. This is unlikely an issue for animals such as pigs, chicken, and mink that usually stay in barns/coops, often close to the farmers' home address. However, different housing systems are used for dairy and beef (veal calves) cattle. According to CBS, most dairy cattle had access to outdoor pastures during grazing season in 1997 in the Netherlands (CBS 1997), while beef cattle is mainly raised indoors. In the Netherlands, sheep are also typically managed in grazing systems and equines (horses) usually have high mobility due to their use in sports; uncertainty about the location of these animals also hampers the interpretation of the results of the sensitivity analyses where we included them (S4). Overall, however, high uncertainty of location information of a specific animal type would mean that the absence of statistically increased risks does not preclude that such risks may exist. We did not define a cut-off for the minimum number of animals a farm should have, thus exploring the effect living near a relatively (very) low number of livestock animals. This may have introduced some exposure misclassification if people lived near a few hobby animals or obsolete licenses with a few animals in the "low exposed" group. Still, for compari-

son, we provide the VGO cut-offs and conducted a sensitivity analysis using these cut-offs, observing that our main findings and conclusions remained unchanged (S1 and S3). Moreover, livestock farming is a ubiquitous activity in the Netherlands resulting in lack of exposure contrasts, especially in the larger buffers. Spatial analysis and adjustment for other types of livestock farming than the one of interest was further complicated by the fact that some types of animals are more predominant in some regions. Nevertheless, our stratified analyses by region showed no indication that the results were influenced by big heterogeneity between these regions, although meta-analytical risk estimates were slightly attenuated (S5). Sensitivity analyses focused on the VGO project study area revealed clear attenuation of HRs towards unity. In this region, most people lived in close proximity to pig farms resulting in reduced exposure contrasts. We observed that there were both somewhat more cases among unexposed and somewhat fewer cases among the exposed in the VGO area as compared to the whole country. In combination, this hampered the interpretation of the results of a sensitivity analysis limited to a smaller regional unit. Stratified analysis by region showed that heterogeneity was overall low.

Our results suggest increased risks of CLR (which is dominated by COPD, S2.4) in residents living close to pigs. Several previous studies have shown negative associations between residential proximity to farms and COPD, which conflicts with what was shown in studies conducted among farmers[24,33]. A possible explanation could be that these studies were cross sectional, a design not best suited to study the relationship between long term exposures and chronic diseases. Studies among farmers have indeed shown a higher risk of developing COPD, especially in cattle, poultry and pig farmers, probably due to long-term exposure to indoor air contaminants[34,35].

We observed indications of reduced risks of CLR mortality and increased risks of PNE mortality in people living near cattle. Although a recent study, also conducted in the Netherlands, showed a decreased risks for asthma and COPD prescriptions, especially among people living near cattle[36], because we were unable to distinguish dairy and beef cattle in our proxy, estimates obtained for cattle are difficult to interpret. While previous studies reported several adverse respiratory effects in people living near poultry farms, we did not find a clear pattern indicating higher risk of respiratory mortality in people living near chicken, which could be related to the difference in outcomes explored (mortality vs symptoms/diagnosis)[1,15,20–22]. We also observed an indication for higher risk of PNE in people living within 1500m and 2000m of mink farms. Note that mink farming is banned from the Netherlands as of January.

Although research shows that living near rabbit and goat farms can be associated with adverse respiratory effects and we did include their presence in the models (Tables S4), we refrained from interpreting the results, due to the following reasons: first, there are 40–50 rabbit farms in the Netherlands, and most farms keep just few rabbits, there-

fore resulting in very few people exposed to rabbit farms. Second, the goat farm industry has seen an important increase between 2000 and 2009, with a doubling of the number of animals and an increasing number of farms. Our exposure (1999–2003) and follow-up (2005–2012) periods encompass, each, part of this rise, resulting in underestimation of "goat exposure", especially during the follow-up period. This is particularly relevant in the context of the large Q fever outbreak of 2007–2010. That our results show no evidence of an association between living near goat farms and increased pneumonia mortality may be related to this underestimation and is in contrast to the clear increase in the risk of having pneumonia in people living near goat farms, even after the epidemic, as reported by several studies[1,3,15,22,37,38].

Despite our data showing lower animal counts (possible related to that smaller farms) were not more frequently located closer to residences than very high animal counts (possibly related to larger farms), we observed that some HRs were, counterintuitively, higher in the low category compared to the high category. These results could be possibly explained by differences in type of housing system, type ventilation systems and hygiene practices affecting emissions rates, reinforcing the importance of using quantitative exposure information. In fact, most studies so far, including this one, have used exposure proxies for farm emissions, such as distance and number of farms/animals near residences, which can be prone to ecological fallacies. Because we did not perform direct measurements of exposure nor did we have access to information on animal housing systems and other farming practices that can influence emissions, it remains unclear which compounds emitted by animal farming are responsible for the effects seen. Possible underlying exposures include endotoxins and pathogenic infectious agents as well as particulate matter, reactive nitrogen gases and volatile organic compounds, all shown to have deleterious effects on health [15]. A model to quantify national agricultural emissions such as ammonia, methane, particulate matter and carbon dioxide has been used for the Netherlands since 2011 (the National Emission Model for Agriculture, NEMA)[39]. Furthermore, De Rooij, et al. have done extensive work on improving modelling of farm related exposures. They quantified residential exposure to livestock farms' emissions in the Dutch agricultural setting by developing land use regression and dispersion models, analogue to traffic related air pollution models, to predict endotoxin exposure at residential addresses[33,40]. This work has demonstrated that predicted PM_{10} and endotoxin concentrations are well, if not better, suited for individual exposure assessment. Because these models rely on data collected in a later time period than that assessed by us, they are unsuited to be applied to the period of this study (1999–2003). De Rooij's models were furthermore developed in a specific region of the Netherlands and may be unfit for application to the whole country. Still, they constitute a valuable tool to evaluate associations in more detail in the future.

CONCLUSIONS

In conclusion, this is the first exploratory study conducted in the Netherlands assessing possible associations between respiratory mortality and residential proximity to (intensive) animal farming. We observed an association between residential proximity to pig farms and increased mortality from respiratory diseases, namely COPD and pneumonia and some indications of higher risk of pneumonia in people living near mink. Deeper insights and better guidance towards interventions warrant both additional analyses using improved exposure assessment methodology, using either quantitative molecular techniques or modelled particulate matter and endotoxin residential exposure on a national scale, and identification of the pollutants driving respiratory health effects observed in this and other Dutch studies.

ACKNOWLEDGEMENTS

This work was funded by internal funding of the Institute of Risk Assessment Sciences, Utrecht University. We thank the Netherlands Enterprise Agency (RVO) for providing access to the GIAB datasets and Statistics Netherlands (CBS) for providing access to the administrative cohort data and facilitating analyses.

TABLES

Table 1: Demographic characteristics of the study population and exposed population (at least one livestock animal within 500, 1000, 1500 and 2000m from the residence), at baseline (2004).

Characteristic	Study Population (N=4,040,845)	Exposed Population			
		500m (N=2,203,650)	1000m (N=3,525,961)	1500m (N=3,884 771)	2000m (N=3,993,150)
AGE [MEAN ± STANDARD DEVIATION]	54.3 ± 13.8	54.1 ±13.7	54.2 ± 13.8	54.3 ± 13.8	54.3 ± 13.8
SEX [N (%)]					
Female	2,082,912 (51.5%)	1,124,139 (51.0%)	1,812,901 (51.4%)	2,000,863 (51.5%)	2,058,022 (51.5%)
Male	1,957,933 (48.5%)	1,079,511 (49.0%)	1,713,060 (48.6%)	1,883,908 (48.5%)	1,935,128 (48.5%)
ORIGIN [N (%)]					
Dutch	3,688,549 (91.3%)	2,046,999 (92.9%)	3,235,409 (91.8%)	3,553,857 (91.5%)	3,648,072 (91.4%)
Western	278,290 (6.9%)	130,105 (5.9%)	232,409 (6.6%)	262,824 (6.8%)	273,233 (6.8%)
Non-Western	74,006 (1.8%)	26,546 (1.2%)	58,143 (1.6%)	68,090 (1.8%)	71,845 (1.8%)
CIVIL STATUS [N (%)]					
Married/partner	3,077,930 (76.2%)	1,703,409 (77.3%)	2,699,958 (76.6%)	2,964,955 (76.3%)	3,043,824 (76.2%)
Widowed	328,450 (8.1%)	174,242 (7.9%)	282,552 (8.0%)	314,315 (8.1%)	324,521 (8.1%)
Divorced	222,307 (5.5%)	102,427 (4.6%)	185,796 (5.3%)	209,866 (5.4%)	217,866 (5.5%)
Single	412,060 (10.2%)	223,522 (10.1%)	357,568 (10.1%)	395,541 (10.2%)	406,843 (10.2%)
Unknown	98 (0.0%)	50 (0.0%)	87 (0.0%)	94 (0.0%)	96 (0.0%)
HOUSEHOLD INCOME [N (%)]					
<1 percentile	29,409 (0.7%)	19,353 (0.9%)	26,224 (0.7%)	28,375 (0.7%)	29,063 (0.7%)
1-<5 percentile	45,281 (1.1%)	27,663 (1.3%)	39,859 (1.1%)	43,516 (1.1%)	44,680 (1.1%)
5-<10 percentile	89,520 (2.2%)	52,256 (2.4%)	78,591 (2.2%)	86,126 (2.2%)	88,437 (2.2%)
10-<25 percentile	453,406 (11.2%)	249,034 (11.3%)	396,257 (11.2%)	436,811 (11.2%)	448,558 (11.2%)
25-<50 percentile	1,009,797 (25.0%)	543,970 (24.7%)	880,901 (25.0%)	971,012 (25.0%)	998,198 (25.0%)
50-<75 percentile	1,138,095 (28.2%)	617,294 (28.0%)	995,349 (28.2%)	1,094,677 (28.2%)	1,124,407 (28.2%)
75-<90 percentile	748,297 (18.5%)	408,160 (18.5%)	654,468 (18.6%)	720,239 (18.5%)	739,377 (18.5%)
90-<95 percentile	261,202 (6.5%)	142,363 (6.5%)	227,208 (6.4%)	250,803 (6.5%)	258,034 (6.5%)
95-<99 percentile	207,369 (5.1%)	112,008 (5.1%)	177,933 (5.0%)	198,043 (5.1%)	204,653 (5.1%)
99-100 percentile	54,102 (1.3%)	29,430 (1.3%)	45,621 (1.3%)	51,107 (1.3%)	53,467 (1.3%)
Unknown	4,367 (0.1%)	2,119 (0.1%)	3,550 (0.1%)	4,062 (0.1%)	4,276 (0.1%)

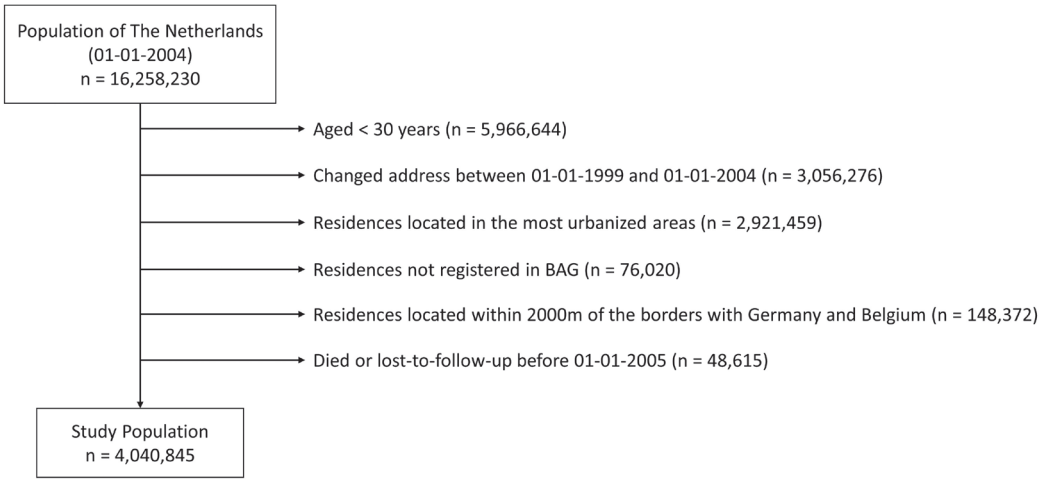
SOCIOECONOMIC POSITION [N (%)]					
1st quintile	298,062 (7.4%)	115,141 (5.2%)	232,868 (6.6%)	277,870 (7.2%)	292,719 (7.3%)
2nd quintile	769,701 (19.0%)	387,401 (17.6%)	653,861 (18.5%)	732,784 (18.9%)	760,863 (19.1%)
3rd quintile	1,070,146 (26.5%)	636,141 (28.9%)	966,907 (27.4%)	1,044,219 (26.9%)	1,063,152 (26.6%)
4th quintile	1,082,734 (26.8%)	663,151 (30.1%)	975,373 (27.7%)	1,042,292 (26.8%)	1,066,287 (26.7%)
5th quintile	791,645 (19.6%)	377,745 (17.1%)	669,472 (19.0%)	759,409 (19.5%)	781,771 (19.6%)
Unknown	28,557 (0.7%)	24,071 (1.1%)	27,480 (0.8%)	28,197 (0.7%)	28,358 (0.7%)
URBANIZATION DEGREE [N (%)]					
<500 addresses per km2	1,366,905 (33.8%)	1,048,318 (47.6%)	1,307,841 (37.1%)	1,351,076 (34.8%)	1,362,100 (34.1%)
500–1000 addresses per km2	1,361,141 (33.7%)	455,184 (20.7%)	1,040,593 (29.5%)	1,253,290 (32.3%)	1,324,294 (33.2%)
1000–1500 addresses per km2	1,312,799 (32.5%)	700,148 (31.8%)	1,177,527 (33.4%)	1,280,405 (33.0%)	1,306,756 (32.7%)
PROPORTION OF PEOPLE WITH LOW EDUCATION [N (%)]					
1st quintile	544,598 (13.5%)	257,759 (11.7%)	451,699 (12.8%)	517,145 (13.3%)	536,404 (13.4%)
2nd quintile	888,532 (22.0%)	493,668 (22.4%)	786,101 (22.3%)	853,556 (22.0%)	874,482 (21.9%)
3rd quintile	969,472 (24.0%)	546,933 (24.8%)	857,073 (24.3%)	940,832 (24.2%)	963,065 (24.1%)
4th quintile	921,454 (22.8%)	515,918 (23.4%)	814,439 (23.1%)	891,406 (22.9%)	912,901 (22.9%)
5th quintile	715,681 (17.7%)	388,857 (17.6%)	615,723 (17.5%)	680,797 (17.5%)	705,203 (17.7%)
Unknown	1,108 (0.0%)	515 (0.0%)	926 (0.0%)	1,035 (0.0%)	1,095 (0.0%)
MEAN CONCENTRATION OF NITROGEN DIOXIDE (NO_2), µg/m³ [N (%)]					
1st quartile [7.27,26.9]	1,704,937 (42.2%)	1,106,561 (50.2%)	1,540,337 (43.7%)	1,656,606 (42.6%)	1,693,048 (42.4%)
2nd quartile (26.9,31.7]	1,390,917 (34.4%)	732,236 (33.2%)	1,215,605 (34.5%)	1,331,406 (34.3%)	1,367,766 (34.3%)
3rd quartile (31.7,36.6]	793,081 (19.6%)	309,628 (14.1%)	657,190 (18.6%)	757,547 (19.5%)	783,600 (19.6%)
4th quartile(36.6,93]	126,416 (3.1%)	39,881 (1.8%)	90,747 (2.6%)	114,723 (3.0%)	123,438 (3.1%)
Unknown	25,494 (0.6%)	15,344 (0.7%)	22,082 (0.6%)	24,489 (0.6%)	25,298 (0.6%)
MEAN CONCENTRATION OF FINE PARTICULATE MATTER ($PM_{2.5}$), µg/m³ [N (%)]					
1st quartile [8.57,15.4]	1,319,422 (32.7%)	713,768 (32.4%)	1,097,988 (31.1%)	1,244,281 (32.0%)	1,295,899 (32.5%)
2nd quartile (15.4,16.5]	892,046 (22.1%)	517,181 (23.5%)	783,213 (22.2%)	855,853 (22.0%)	880,305 (22.0%)
3rd quartile (16.5,17.4]	836,048 (20.7%)	484,008 (22.0%)	753,564 (21.4%)	814,429 (21.0%)	827,862 (20.7%)
4th quartile (17.4,20.8]	967,835 (24.0%)	473,349 (21.5%)	869,114 (24.6%)	945,719 (24.3%)	963,786 (24.1%)
Unknown	25,494 (0.6%)	15,344 (0.7%)	22,082 (0.6%)	24,489 (0.6%)	25,298 (0.6%)

Table 2: Associations between living within 500, 1000, 1500 and 2000m from livestock animals and mortality due to Respiratory system diseases, Chronic lower respiratory diseases and Pneumonia. Results are presented as Hazard Ratios (HR) and its corresponding 95% Confidence Interval (95% CI) and P value (full model).

Type animal	Buffer size	Category	Respiratory System Diseases		Chronic Lower Respiratory Diseases		Pneumonia	
			HR (95% CI)	P value	HR (95% CI)	P value	HR (95% CI)	P value
CATTLE	500m	low	1.01 (0.98, 1.03)	0.521	0.99 (0.96, 1.03)	0.618	1.04 (1.00, 1.08)	0.081
		high	1.02 (0.99, 1.05)	0.277	0.99 (0.94, 1.03)	0.518	1.05 (1.00, 1.10)	0.045
	1000m	low	0.99 (0.96, 1.02)	0.589	0.97 (0.93, 1.02)	0.235	0.99 (0.95, 1.04)	0.741
		high	1.01 (0.97, 1.04)	0.630	1.01 (0.96, 1.06)	0.645	0.99 (0.94, 1.04)	0.692
	1500m	low	0.97 (0.93, 1.02)	0.217	0.89 (0.84, 0.95)	2.2e-04	1.06 (0.99, 1.14)	0.114
		high	1.00 (0.96, 1.05)	0.866	0.91 (0.85, 0.97)	0.007	1.09 (1.00, 1.18)	0.038
	2000m	low	0.92 (0.86, 0.99)	0.020	0.84 (0.76, 0.93)	6.8e-04	0.99 (0.88, 1.11)	0.858
		high	0.97 (0.90, 1.04)	0.363	0.86 (0.77, 0.95)	0.004	1.06 (0.94, 1.20)	0.312
PIGS	500m	low	1.05 (1.01, 1.09)	0.018	1.02 (0.97, 1.08)	0.407	1.11 (1.05, 1.18)	5.5e-04
		high	1.03 (0.99, 1.07)	0.158	1.08 (1.02, 1.14)	0.011	1.04 (0.97, 1.11)	0.300
	1000m	low	1.04 (1.01, 1.07)	0.006	1.06 (1.02, 1.10)	0.006	1.03 (0.99, 1.08)	0.171
		high	1.08 (1.05, 1.11)	1.5e-06	1.09 (1.04, 1.14)	1.7e-04	1.08 (1.03, 1.14)	0.002
	1500m	low	1.07 (1.04, 1.10)	5.8e-07	1.08 (1.04, 1.12)	3.0e-05	1.04 (1.00, 1.08)	0.061
		high	1.12 (1.08, 1.15)	1.9e-12	1.16 (1.11, 1.21)	1.3e-10	1.07 (1.02, 1.13)	0.005
	2000m	low	1.06 (1.03, 1.09)	1.9e-05	1.09 (1.04, 1.13)	4.3e-05	1.01 (0.97, 1.06)	0.567
		high	1.11 (1.07, 1.15)	5.6e-10	1.18 (1.13, 1.24)	5.0e-12	1.03 (0.98, 1.09)	0.230
CHICKEN	500m	low	0.98 (0.94, 1.02)	0.275	0.98 (0.93, 1.04)	0.574	0.96 (0.91, 1.02)	0.236
		high	1.04 (1.00, 1.08)	0.031	1.01 (0.96, 1.07)	0.705	1.04 (0.98, 1.11)	0.209
	1000m	low	1.00 (0.97, 1.02)	0.848	1.00 (0.96, 1.04)	0.965	1.01 (0.97, 1.06)	0.528
		high	1.02 (1.00, 1.05)	0.098	1.01 (0.97, 1.05)	0.636	1.04 (1.00, 1.09)	0.061
	1500m	low	1.00 (0.97, 1.03)	0.994	1.00 (0.96, 1.04)	0.943	1.01 (0.97, 1.05)	0.764
		high	1.01 (0.98, 1.04)	0.490	1.01 (0.97, 1.05)	0.562	1.02 (0.98, 1.07)	0.349
	2000m	low	1.00 (0.97, 1.03)	0.841	0.99 (0.94, 1.04)	0.630	1.01 (0.96, 1.06)	0.732
		high	1.03 (1.00, 1.07)	0.053	1.04 (0.99, 1.10)	0.094	1.03 (0.97, 1.08)	0.362
MINK	500m	low	1.04 (0.83, 1.30)	0.731	0.85 (0.60, 1.20)	0.356	1.30 (0.94, 1.80)	0.109
		high	0.92 (0.76, 1.12)	0.426	0.93 (0.70, 1.23)	0.601	0.97 (0.71, 1.33)	0.870
	1000m	low	0.99 (0.89, 1.10)	0.821	0.99 (0.85, 1.16)	0.927	0.97 (0.82, 1.15)	0.744
		high	1.01 (0.90, 1.12)	0.921	0.95 (0.81, 1.11)	0.493	1.11 (0.94, 1.31)	0.211
	1500m	low	1.01 (0.94, 1.08)	0.785	1.00 (0.91, 1.11)	0.942	1.01 (0.90, 1.13)	0.889
		high	1.06 (0.99, 1.14)	0.121	0.97 (0.87, 1.08)	0.593	1.20 (1.08, 1.34)	0.001
	2000m	low	1.01 (0.95, 1.06)	0.797	1.00 (0.93, 1.08)	0.954	1.02 (0.93, 1.11)	0.731
		high	1.06 (1.00, 1.12)	0.049	0.98 (0.90, 1.06)	0.549	1.20 (1.10, 1.31)	2.3e-05

Models were adjusted for: sex, origin, marital status, standardized household income, neighborhood 's socioeconomic position, urbanization degree at neighborhood level, proportion of low educated residents in the neighborhood , ambient particulate matter < 2.5 µm in diameter ($PM_{2.5}$) levels and ambient nitrogen dioxide (NO_2) levels. Models were also adjusted for the presence of goats and rabbits. The referent category are those with zero animals in the respective buffer.

FIGURES



BAG = Basisregistratie Adressen en Gebouwen, the cadastral dataset containing all addresses in the Netherlands used to compute individual residential exposure proxies.

Figure 1: Flow chart of the study population.

RESPIRATORY MORTALITY ENDPOINTS (ICD-10 codes)		EXPOSURE VARIABLES		POTENTIAL CONFOUNDERS
Respiratory system diseases (J00-J99) **Chronic lower respiratory diseases** (J40-J47) **Pneumonia** (J12-J18, A70, A78, I00, A010, A022, A212, A221, A370, A371, A378, A379, A420, A430, A548, A698, B012, B052, B068, B250, B334, B371, B380, B381, B382, B390, B391, B392, B393, B394, B395, B399, B440, B441, B583, B650, B651, B652, B653, B658, B659, B778)	~	Four models, one per buffer size (500m, 1000m, 1550m, 2000m) : cattle + pigs + chicken + mink[1]	+	*basic model:* sex *intermediate model:* *basic* + origin + marital status + standardized household income *full model:* *intermediate* + $PM_{2.5}$ + NO_2 + neighbourhood's socioeconomic position + neighbourhood's urbanization degree + neighbourhood's proportion of low educated residents

$PM_{2.5}$ = ambient Particulate Matter < 2.5 µm in diameter, NO_2 = ambient nitrogen dioxide
1 Exposure variables were categorized into 'none', 'low' (<median number of animals in buffer) and 'high' (≥median numb)

Figure 2: General framework of the models used in the study.

REFERENCES

1. Smit LAM, van der Sman-de Beer F, Opstal-van Winden AWJ, Hooiveld M, Beekhuizen J, Wouters IM, et al. Q fever and pneumonia in an area with a high livestock density: a large population-based study. PLoS One. 2012;7(6):e38843.
2. Cambra-López M, Aarnink AJA, Zhao Y, Calvet S, Torres AG. Airborne particulate matter from livestock production systems: A review of an air pollution problem. Environ Pollut [Internet]. 2010;158(1):1–17. Available from: http://www.sciencedirect.com/science/article/pii/S0269749109003509
3. Jonges M, van Leuken J, Wouters I, Koch G, Meijer A, Koopmans M. Wind-Mediated Spread of Low-Pathogenic Avian Influenza Virus into the Environment during Outbreaks at Commercial Poultry Farms. PLoS One. 2015;10(5):e0125401.
4. Seedorf J, Hartung J, Schröder M, Linkert KH, Phillips VR, Holden MR, et al. Concentrations and Emissions of Airborne Endotoxins and Microorganisms in Livestock Buildings in Northern Europe. J Agric Eng Res [Internet]. 1998;70(1):97–109. Available from: http://www.sciencedirect.com/science/article/pii/S002186349790281X
5. Schulze A, van Strien R, Ehrenstein V, Schierl R, Küchenhoff H, Radon K. Ambient endotoxin level in an area with intensive livestock production. Ann Agric Environ Med. 2006;13(1):87–91.
6. de Rooij MMT, Heederik DJJ, Borlée F, Hoek G, Wouters IM. Spatial and temporal variation in endotoxin and PM10 concentrations in ambient air in a livestock dense area. Environ Res [Internet]. 2017;153:161–70. Available from: http://www.sciencedirect.com/science/article/pii/S0013935116306028
7. Winkel A, Mosquera J, Groot Koerkamp PWG, Ogink NWM, Aarnink AJA. Emissions of particulate matter from animal houses in the Netherlands. Atmos Environ [Internet]. 2015;111:202–12. Available from: http://www.sciencedirect.com/science/article/pii/S135223101500271X
8. Borlée F, Yzermans CJ, van Dijk CE, Heederik D, Smit LAM. Increased respiratory symptoms in COPD patients living in the vicinity of livestock farms. Eur Respir J [Internet]. 2015 Dec 1;46(6):1605 LP – 1614. Available from: http://erj.ersjournals.com/content/46/6/1605.abstract
9. Borlée F, Yzermans CJ, Aalders B, Rooijackers J, Krop E, Maassen CBM, et al. Air Pollution from Livestock Farms Is Associated with Airway Obstruction in Neighboring Residents. Am J Respir Crit Care Med [Internet]. 2017 May 10;196(9):1152–61. Available from: https://doi.org/10.1164/rccm.201701-0021OC
10. Radon K, Schulze A, Ehrenstein V, van Strien RT, Praml G, Nowak D. Environmental Exposure to Confined Animal Feeding Operations and Respiratory Health of Neighboring Residents. Epidemiology [Internet]. 2007;18(3):300–8. Available from: http://www.jstor.org/stable/20486369
11. Rasmussen S, Casey J, Bandeen-Roche K, Schwartz B. Proximity to Industrial Food Animal Production and Asthma Exacerbations in Pennsylvania, 2005–2012. Int J Environ Res Public Health [Internet]. 2017 Mar 31 [cited 2020 Nov 12];14(4):362. Available from: http://www.mdpi.com/1660-4601/14/4/362

12. van Kersen W, Oldenwening M, Aalders B, Bloemsma LD, Borlée F, Heederik D, et al. Acute respiratory effects of livestock-related air pollution in a panel of COPD patients. Environ Int. 2020 Mar;136:105426.
13. O'Connor AM, Auvermann BW, Dzikamunhenga RS, Glanville JM, Higgins JPT, Kirychuk SP, et al. Updated systematic review: associations between proximity to animal feeding operations and health of individuals in nearby communities. Syst Rev [Internet]. 2017;6(1):86. Available from: https://doi.org/10.1186/s13643–017–0465-z
14. Eijrond V, Claassen L, van der Giessen J, Timmermans D. Intensive Livestock Farming and Residential Health: Experts' Views. Int J Environ Res Public Health [Internet]. 2019 Sep 27;16(19):3625. Available from: https://pubmed.ncbi.nlm.nih.gov/31569632
15. Maassen K, Smit L, Wouters I, van Duijkeren E, Janse I, Hagenaars T, et al. Veehouderij en gezondheid omwonenden (Livestock farming and the health of local residents) [Internet]. 2016. Available from: https://www.rivm.nl/publicaties/veehouderij-en-gezondheid-omwonenden
16. de Neeling AJ, van den Broek MJM, Spalburg EC, van Santen-Verheuvel MG, Dam-Deisz WDC, Boshuizen HC, et al. High prevalence of methicillin resistant Staphylococcus aureus in pigs. Vet Microbiol [Internet]. 2007;122(3):366–72. Available from: https://www.sciencedirect.com/science/article/pii/S037811350700065X
17. Overdevest I, Willemsen I, Rijnsburger M, Eustace A, Xu L, Hawkey P, et al. Extended-spectrum β-lactamase genes of Escherichia coli in chicken meat and humans, The Netherlands. Emerg Infect Dis [Internet]. 2011 Jul;17(7):1216–22. Available from: https://pubmed.ncbi.nlm.nih.gov/21762575
18. Oude Munnink BB, Sikkema RS, Nieuwenhuijse DF, Molenaar RJ, Munger E, Molenkamp R, et al. Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. Science. 2021 Jan;371(6525):172–7.
19. Oreshkova N, Molenaar RJ, Vreman S, Harders F, Oude Munnink BB, Hakze-van der Honing RW, et al. SARS-CoV-2 infection in farmed minks, the Netherlands, April and May 2020. Euro Surveill Bull Eur sur les Mal Transm = Eur Commun Dis Bull. 2020 Jun;25(23).
20. Smit LAM, Boender GJ, de Steenhuijsen Piters WAA, Hagenaars TJ, Huijskens EGW, Rossen JWA, et al. Increased risk of pneumonia in residents living near poultry farms: does the upper respiratory tract microbiota play a role? Pneumonia (Nathan Qld). 2017;9:3.
21. Kalkowska DA, Boender GJ, Smit LAM, Baliatsas C, Yzermans J, Heederik DJJ, et al. Associations between pneumonia and residential distance to livestock farms over a five-year period in a large population-based study. PLoS One. 2018;13(7):e0200813.
22. Freidl GS, Spruijt IT, Borlée F, Smit LAM, van Gageldonk-Lafeber AB, Heederik DJJ, et al. Livestock-associated risk factors for pneumonia in an area of intensive animal farming in the Netherlands. PLoS One [Internet]. 2017 Mar 31;12(3):e0174796. Available from: https://doi.org/10.1371/journal.pone.0174796
23. van Dijk CE, Garcia-Aymerich J, Carsin A-E, Smit LAM, Borlée F, Heederik DJ, et al. Risk of exacerbations in COPD and asthma patients living in the neighborhood of livestock

farms: Observational study using longitudinal data. Int J Hyg Environ Health [Internet]. 2016;219(3):278–87. Available from: http://www.sciencedirect.com/science/article/pii/S1438463916000031

24. Smit LAM, Hooiveld M, van der Sman-de Beer F, Opstal-van Winden AWJ, Beekhuizen J, Wouters IM, et al. Air pollution from livestock farms, and asthma, allergic rhinitis and COPD among neighbouring residents. Occup Environ Med [Internet]. 2014 Feb 1;71(2):134 LP – 140. Available from: http://oem.bmj.com/content/71/2/134.abstract
25. World Health Organization. ICD-10 : international statistical classification of diseases and related health problems : tenth revision. 2nd ed. World Health Organization; 2004.
26. Knol FA. Van hoog naar laag, van laag naar hoog: de sociaal-ruimtelijke ontwikkeling van wijken tussen 1971–1995 [in Dutch]. The Hague: Sociaal en Cultureel Planbureau; 1998.
27. de Hoogh K, Chen J, Gulliver J, Hoffmann B, Hertel O, Ketzel M, et al. Spatial PM2.5, NO2, O3 and BC models for Western Europe – Evaluation of spatiotemporal stability. Environ Int [Internet]. 2018;120:81–92. Available from: http://www.sciencedirect.com/science/article/pii/S0160412018309759
28. Malhotra J, Malvezzi M, Negri E, La Vecchia C, Boffetta P. Risk factors for lung cancer worldwide. Eur Respir J. 2016 Sep;48(3):889–902.
29. Stoffel EM, Murphy CC. Epidemiology and Mechanisms of the Increasing Incidence of Colon and Rectal Cancers in Young Adults. Gastroenterology. 2020 Jan;158(2):341–53.
30. Saginala K, Barsouk A, Aluru JS, Rawla P, Padala SA, Barsouk A. Epidemiology of Bladder Cancer. Med Sci (Basel, Switzerland) [Internet]. 2020 Mar 13;8(1):15. Available from: https://pubmed.ncbi.nlm.nih.gov/32183076
31. Sheron N. Alcohol and liver disease in Europe – Simple measures have the potential to prevent tens of thousands of premature deaths. J Hepatol [Internet]. 2016;64(4):957–67. Available from: http://www.sciencedirect.com/science/article/pii/S0168827815007370
32. Statistics Netherlands (CBS). Weidegang van melkvee (pasture grazing of dairy cattle); weidegebied 1997–2015 [Internet]. 1997 [cited 2020 Dec 10]. Available from: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/70736ned/table?ts=1607603670712
33. de Rooij MMT, Smit LAM, Erbrink HJ, Hagenaars TJ, Hoek G, Ogink NWM, et al. Endotoxin and particulate matter emitted by livestock farms and respiratory health effects in neighboring residents. Environ Int [Internet]. 2019;132:105009. Available from: http://www.sciencedirect.com/science/article/pii/S0160412019311365
34. Fontana L, Lee S-J, Capitanelli I, Re A, Maniscalco M, Mauriello MC, et al. Chronic Obstructive Pulmonary Disease in Farmers: A Systematic Review. J Occup Environ Med [Internet]. 2017;59(8). Available from: https://journals.lww.com/joem/Fulltext/2017/08000/Chronic_Obstructive_Pulmonary_Disease_in_Farmers_.14.aspx
35. Guillien A, Soumagne T, Dalphin J-C, Degano B. COPD, airflow limitation and chronic bronchitis in farmers: a systematic review and meta-analysis. Occup Environ Med [Internet]. 2019 Jan 1;76(1):58 LP – 68. Available from: http://oem.bmj.com/content/76/1/58.abstract
36. Post PM, Houthuijs D, Sterk HAM, Marra M, van de Kassteele J, van Pul A, et al. Proximity to livestock farms and exposure to livestock-related particulate matter are associated with lower probability of medication dispensing for obstructive airway diseases. Int J Hyg Environ Health. 2021 Jan;231:113651.
37. Klous G, Smit LAM, Freidl GS, Borlée F, van der Hoek W, IJzermans CJ, et al. Pneumonia risk

of people living close to goat and poultry farms – Taking GPS derived mobility patterns into account. Environ Int [Internet]. 2018;115:150–60. Available from: http://www.sciencedirect.com/science/article/pii/S0160412017317907

38. Post PM, Hogerwerf L, Huss A, Petie R, Boender GJ, Baliatsas C, et al. Risk of pneumonia among residents living near goat and poultry farms during 2014–2016. PLoS One. 2019;14(10):e0223601.

39. Velthof GL, van Bruggen C, Groenestein CM, Haan BJ, Hoogeveen MW, Huijsmans JFM. A model for inventory of ammonia emissions from agriculture in the Netherlands. Atmos Environ. 2012 Jan 31;46:248–55.

40. de Rooij M, Heederik D, van Nunen E, van Schothorst I, Maassen C, Gerard H, et al. Spatial Variation of Endotoxin Concentrations Measured in Ambient PM10 in a Livestock-Dense Area: Implementation of a Land-Use Regression Approach. Environ Health Perspect [Internet]. 2018 Nov 13;126(1):17003. Available from: https://doi.org/10.1289/EHP2252

SUPPLEMENTAL MATERIAL

Contents

*https://doi.org/10.1016/j.envint.2022.107140

S1 Classification of farms in the "Livestock Farming and Neighbouring Residents' Health" (VGO) project

The VGO project analyses included farms with a minimum number of animals (Table S1.1). Table S1.2 shows how the classification of farms changed in our study when we assigned zero to all farms that had less than the minimum number of animals used in the VGO project.

Table S1.1: Minimum number of animals in a farm in the "Livestock Farming and Neighbouring Residents' Health" (VGO) project.

Type animal	Minimum number on VGO project
Cattle	>5
Pigs	>25
Chicken	>250
Goats	≥50
Rabbits	>250
Mink	≥400

Table S1.2: Comparing the number of animals in each exposure category (per buffer size) used in the main analysis (no cut-off for the minimum number of animals in a farm) and in sensitivity analysis using the same the cut-offs as used in the "Livestock Farming and Neighbouring Residents' Health" (VGO) project.

Animal type	VGO analysis / Main analysis	Number of animals					
		500m			1000m		
		None	Low	High	None	Low	High
CATTLE	none	2042288	0	0	662508	0	0
	low	268568	731567	0	187121	1502969	0
	high	0	134767	863655	0	93824	1594423
PIGS	none	3353026	0	0	2280490	0	0
	low	125302	218865	0	268570	611608	0
	high	0	63389	280263	0	134420	745757
CHICKEN	none	3366436	0	0	2141180	0	0
	low	338460	0	0	950699	0	0
	high	114627	110670	110652	197110	379113	372743
MINK	none	4021352	0	0	3963822	0	0
	low	6278	3469	0	22459	16416	0
	high	0	3328	6418	0	11147	27001

Animal type	VGO analysis / Main analysis	Number of animals					
		1500m			2000m		
		None	Low	High	None	Low	High
CATTLE	none	228186	0	0	79920	0	0
	low	100686	1806147	0	48868	1931709	0
	high	0	49873	1855953	0	24575	1955773
PIGS	none	1509321	0	0	990886	0	0
	low	308670	958698	0	336939	1188632	0
	high	0	152784	1111372	0	167879	1356509
CHICKEN	none	1157039	0	0	556734	0	0
	low	1442379	0	0	1524120	218309	0
	high	74597	683475	683355	0	762371	979311
MINK	none	3874409	0	0	3752949	0	0
	low	43349	40230	0	64173	80607	0
	high	0	28534	54323	0	36161	106955

S2 Descriptive statistics

Table S2.1: Descriptive statistics of the exposure proxy variables, as continuous: average, median, minimum and maximum number of animals per buffer size for the period 1999–2003.

Buffer size / Animal type	500m			1000m			1500m			2000m		
	Median	Min[1]	Max[2]	Median	Min[1]	Max[2]	Median	Min[1]	Max[2]	Median	Min[1]	Max[2]
Cattle	59	1	11000	177	1	17500	462	1	22000	932	1	30500
Pigs	334	1	46000	703	1	69500	1169	1	94000	1843	1	122500
Chicken	19	1	663500	39	1	726500	151	1	1030500	1828	1	1403000
Mink	1760	1	41000	2050	1	54500	2050	1	62500	2136	1	66000

Min = minimum, Max = maximum

[1] The minimum was rounded to 1 if the average as <1.

[2] To avoid risk of disclosure the maximum was rounded to a multiple of 500.

Table S2.2: Number of individuals unexposed and exposed to low and high numbers of each type of livestock animal within 500m, 1000, 1500m and 2000m of their residence.

Type animal	Buffer size	Number of animals		
		None	Low	High
CATTLE	500m	2042288 (50.54%)	874498 (21.64%)	1124059 (27.82%)
	1000m	662508 (16.4%)	1216848 (30.11%)	2161489 (53.49%)
	1500m	228186 (5.65%)	1163861 (28.8%)	2648798 (65.55%)
	2000m	79920 (1.98%)	1113365 (27.55%)	2847560 (70.47%)
PIGS	500m	3353026 (82.98%)	305475 (7.56%)	382344 (9.46%)
	1000m	2280490 (56.44%)	708297 (17.53%)	1052058 (26.04%)
	1500m	1509321 (37.35%)	940950 (23.29%)	1590574 (39.36%)
	2000m	990886 (24.52%)	1039689 (25.73%)	2010270 (49.75%)
CHICKEN	500m	3366436 (83.31%)	335019 (8.29%)	339390 (8.4%)
	1000m	2141180 (52.99%)	886973 (21.95%)	1012692 (25.06%)
	1500m	1157039 (28.63%)	1258487 (31.14%)	1625319 (40.22%)
	2000m	556734 (13.78%)	1394646 (34.51%)	2089465 (51.71%)
MINK	500m	4021352 (99.52%)	8222 (0.2%)	11271 (0.28%)
	1000m	3963822 (98.09%)	26691 (0.66%)	50332 (1.25%)
	1500m	3874409 (95.88%)	57969 (1.43%)	108467 (2.68%)
	2000m	3752949 (92.88%)	102880 (2.55%)	185016 (4.58%)

Table S2.3: Cross table of the number of individuals that were alive and that died of respiratory system diseases, chronic lower respiratory diseases and pneumonia that were unexposed and exposed to low and high numbers of each type of livestock animal within 500m, 1000, 1500m and 2000m of their residence.

Type animal	Buffer size	Category	Respiratory System Diseases		Chronic Lower Respiratory Diseases		Pneumonia	
			Alive	Dead	Alive	Dead	Alive	Dead
CATTLE		none	2021827 (50.0%)	20461 (0.5%)	2032501 (50.3%)	9787 (0.2%)	2034650 (50.4%)	7638 (0.2%)
	500m	low	865385 (21.4%)	9113 (0.2%)	870187 (21.5%)	4311 (0.1%)	870990 (21.6%)	3508 (0.1%)
		high	1113502 (27.6%)	10557 (0.3%)	1119103 (27.7%)	4956 (0.1%)	1120016 (27.7%)	4043 (0.1%)
		none	655383 (16.2%)	7125 (0.2%)	659151 (16.3%)	3357 (0.1%)	659770 (16.3%)	2738 (0.1%)
	1000m	low	1204013 (29.8%)	12835 (0.3%)	1210803 (30.0%)	6045 (0.1%)	1211939 (30.0%)	4909 (0.1%)
		high	2141318 (53.0%)	20171 (0.5%)	2151837 (53.2%)	9652 (0.2%)	2153947 (53.3%)	7542 (0.2%)
		none	225646 (5.6%)	2540 (0.1%)	226928 (5.6%)	1258 (0.0%)	227254 (5.6%)	932 (0.0%)
	1500m	low	1151770 (28.5%)	12091 (0.3%)	1158225 (28.7%)	5636 (0.1%)	1159197 (28.7%)	4664 (0.1%)
		high	2623298 (64.9%)	25500 (0.6%)	2636638 (65.2%)	12160 (0.3%)	2639205 (65.3%)	9593 (0.2%)
		none	79038 (2.0%)	882 (0.0%)	79486 (2.0%)	434 (0.0%)	79588 (2.0%)	332 (0.0%)
	2000m	low	1101782 (27.3%)	11583 (0.3%)	1107968 (27.4%)	5397 (0.1%)	1108836 (27.4%)	4529 (0.1%)
		high	2819894 (69.8%)	27666 (0.7%)	2834337 (70.1%)	13223 (0.3%)	2837232 (70.2%)	10328 (0.3%)
PIGS		none	3319493 (82.2%)	33533 (0.8%)	3337116 (82.6%)	15910 (0.4%)	3340399 (82.7%)	12627 (0.3%)
	500m	low	302378 (7.5%)	3097 (0.1%)	304072 (7.5%)	1403 (0.0%)	304221 (7.5%)	1254 (0.0%)
		high	378843 (9.4%)	3501 (0.1%)	380603 (9.4%)	1741 (0.0%)	381036 (9.4%)	1308 (0.0%)
		none	2257053 (55.9%)	23437 (0.6%)	2269493 (56.2%)	10997 (0.3%)	2271548 (56.2%)	8942 (0.2%)
	1000m	low	701380 (17.4%)	6917 (0.2%)	704988 (17.4%)	3309 (0.1%)	705663 (17.5%)	2634 (0.1%)
		high	1042281 (25.8%)	9777 (0.2%)	1047310 (25.9%)	4748 (0.1%)	1048445 (25.9%)	3613 (0.1%)
		none	1494033 (37.0%)	15288 (0.4%)	1502180 (37.2%)	7141 (0.2%)	1503413 (37.2%)	5908 (0.1%)
	1500m	low	931614 (23.1%)	9336 (0.2%)	936580 (23.2%)	4370 (0.1%)	937387 (23.2%)	3563 (0.1%)
		high	1575067 (39.0%)	15507 (0.4%)	1583031 (39.2%)	7543 (0.2%)	1584856 (39.2%)	5718 (0.1%)
		none	980882 (24.3%)	10004 (0.2%)	986314 (24.4%)	4572 (0.1%)	986906 (24.4%)	3980 (0.1%)
	2000m	low	1029423 (25.5%)	10266 (0.2%)	1034889 (25.6%)	4800 (0.1%)	1035784 (25.6%)	3905 (0.1%)
		high	1990409 (49.3%)	19861 (0.5%)	2000588 (49.5%)	9682 (0.2%)	2002966 (49.6%)	7304 (0.2%)
CHICKEN		none	3333009 (82.5%)	33427 (0.8%)	3350523 (82.9%)	15913 (0.4%)	3353805 (83.0%)	12631 (0.3%)
	500m	low	331716 (8.2%)	3303 (0.1%)	333465 (8.2%)	1554 (0.0%)	333754 (8.3%)	1265 (0.0%)
		high	335989 (8.3%)	3401 (0.1%)	337803 (8.4%)	1587 (0.0%)	338097 (8.4%)	1293 (0.0%)
		none	2119355 (52.5%)	21825 (0.5%)	2130822 (52.7%)	10358 (0.3%)	2132966 (52.8%)	8214 (0.2%)
	1000m	low	878380 (21.7%)	8593 (0.2%)	882906 (21.9%)	4067 (0.1%)	883661 (21.9%)	3312 (0.1%)
		high	1002979 (24.8%)	9713 (0.2%)	1008063 (24.9%)	4629 (0.1%)	1009029 (25.0%)	3663 (0.1%)

Type animal	Buffer size	Category	Respiratory System Diseases		Chronic Lower Respiratory Diseases		Pneumonia	
			Alive	Dead	Alive	Dead	Alive	Dead
CHICKEN	1500m	none	1144907 (28.3%)	12132 (0.3%)	1151298 (28.5%)	5741 (0.1%)	1152449 (28.5%)	4590 (0.1%)
		low	1246236 (30.8%)	12251 (0.3%)	1252758 (31.0%)	5729 (0.1%)	1253756 (31.0%)	4731 (0.1%)
		high	1609571 (39.8%)	15748 (0.4%)	1617735 (40.0%)	7584 (0.2%)	1619451 (40.1%)	5868 (0.1%)
	2000m	none	550970 (13.6%)	5764 (0.1%)	554019 (13.7%)	2715 (0.1%)	554533 (13.7%)	2201 (0.0%)
		low	1381028 (34.2%)	13618 (0.3%)	1388346 (34.4%)	6300 (0.2%)	1389345 (34.4%)	5301 (0.1%)
		high	2068716 (51.2%)	20749 (0.5%)	2079426 (51.5%)	10039 (0.2%)	2081778 (51.5%)	7687 (0.2%)
MINK	500m	none	3981400 (98.5%)	39952 (1.0%)	4002377 (99.0%)	18975 (0.5%)	4006240 (99.1%)	15112 (0.4%)
		low	8151 (0.2%)	71 (0.0%)	8193 (0.2%)	29 (0.0%)	8191 (0.2%)	31 (0.0%)
		high	11163 (0.3%)	108 (0.0%)	11221 (0.3%)	50 (0.0%)	11225 (0.3%)	46 (0.0%)
	1000m	none	3924407 (97.1%)	39415 (1.0%)	3945102 (97.6%)	18720 (0.5%)	3948914 (97.7%)	14908 (0.4%)
		low	26466 (0.6%)	225 (0.0%)	26592 (0.7%)	99 (0.0%)	26599 (0.7%)	92 (0.0%)
		high	49841 (1.2%)	491 (0.0%)	50097 (1.2%)	235 (0.0%)	50143 (1.2%)	189 (0.0%)
	1500m	none	3835897 (94.9%)	38512 (0.9%)	3856101 (95.4%)	18308 (0.4%)	3859863 (95.5%)	14546 (0.4%)
		low	57367 (1.4%)	602 (0.0%)	57679 (1.4%)	290 (0.0%)	57748 (1.4%)	221 (0.0%)
		high	107450 (2.7%)	1017 (0.0%)	108011 (2.7%)	456 (0.0%)	108045 (2.7%)	422 (0.0%)
	2000m	none	3715581 (92.0%)	37368 (0.9%)	3735187 (92.4%)	17762 (0.4%)	3738845 (92.5%)	14104 (0.3%)
		low	101861 (2.5%)	1019 (0.0%)	102397 (2.5%)	483 (0.0%)	102495 (2.5%)	385 (0.0%)
		high	183272 (4.5%)	1744 (0.0%)	184207 (4.6%)	809 (0.0%)	184316 (4.6%)	700 (0.0%)

Table S2.4: Number of deaths by specific chronic lower respiratory disease and pneumonia in the study population.

Cause of death (Data are in n (%))	
Chronic lower respiratory diseases	**19 054**
Chronic bronchitis	119 (1%)
Emphysema	1793 (9%)
Other chronic obstructive pulmonary disease	16 906 (89%)
Asthma and status asthmaticus	158 (1%)
Bronchiectasis	78 (0%)
Pneumonia	**15 189**
Viral pneumonia	51 (0%)
Bacterial pneumonia	237 (2%)
Pneumonia, organism unspecified	14 857 (98%)
Pneumonia, other infectious agents	44 (0%)

Data are in n (%)

S6 Results of the negative controls analyses

Analyses were conducted when there was a minimum of 10 exposed cases.

Type animal	Buffer size	Category	Colon cancer		Bladder cancer		Liver cancer	
			HR [95% CI]	P value	HR [95% CI]	P value	HR [95% CI]	P value
CATTLE	500m	low	1.00 [0.96, 1.04]	0.832	1.02 [0.94, 1.10]	0.674	1.08 [0.93, 1.26]	0.322
		high	0.98 [0.93, 1.02]	0.286	1.04 [0.95, 1.14]	0.366	0.97 [0.80, 1.17]	0.758
	1000m	low	0.99 [0.95, 1.03]	0.641	0.96 [0.87, 1.05]	0.324	1.15 [0.96, 1.37]	0.143
		high	0.98 [0.93, 1.03]	0.389	0.91 [0.82, 1.01]	0.077	1.19 [0.97, 1.48]	0.102
	1500m	low	0.99 [0.92, 1.06]	0.690	0.89 [0.78, 1.02]	0.101	1.24 [0.94, 1.65]	0.127
		high	0.97 [0.90, 1.04]	0.381	0.88 [0.76, 1.02]	0.092	1.27 [0.93, 1.74]	0.125
	2000m	low	0.98 [0.88, 1.09]	0.722	1.02 [0.81, 1.27]	0.889	1.21 [0.78, 1.87]	0.405
		high	0.97 [0.87, 1.09]	0.601	0.98 [0.78, 1.24]	0.890	1.20 [0.75, 1.90]	0.444
PIGS	500m	low	1.00 [0.94, 1.06]	0.905	0.92 [0.81, 1.04]	0.166	0.94 [0.73, 1.21]	0.634
		high	1.00 [0.94, 1.06]	0.876	0.83 [0.72, 0.95]	0.006	1.05 [0.81, 1.36]	0.691
	1000m	low	1.04 [0.99, 1.08]	0.096	1.06 [0.97, 1.15]	0.193	0.81 [0.68, 0.96]	0.017
		high	1.06 [1.01, 1.11]	0.022	0.91 [0.83, 1.01]	0.079	0.87 [0.72, 1.06]	0.170
	1500m	low	1.03 [0.99, 1.07]	0.217	1.03 [0.95, 1.11]	0.523	0.97 [0.83, 1.14]	0.738
		high	1.09 [1.04, 1.14]	3.6e-04	0.95 [0.86, 1.05]	0.303	0.91 [0.75, 1.10]	0.327
	2000m	low	0.99 [0.95, 1.03]	0.733	0.97 [0.89, 1.05]	0.446	0.97 [0.82, 1.14]	0.700
		high	1.05 [1.00, 1.11]	0.046	0.93 [0.84, 1.03]	0.184	0.86 [0.70, 1.06]	0.165
CHICKEN	500m	low	0.98 [0.92, 1.04]	0.439	0.91 [0.81, 1.03]	0.148	0.98 [0.77, 1.25]	0.884
		high	1.05 [0.99, 1.11]	0.119	0.95 [0.83, 1.07]	0.397	1.08 [0.85, 1.38]	0.538
	1000m	low	1.00 [0.96, 1.04]	0.939	0.95 [0.88, 1.04]	0.259	0.88 [0.74, 1.04]	0.124
		high	1.01 [0.97, 1.06]	0.499	0.99 [0.91, 1.08]	0.838	1.02 [0.86, 1.21]	0.827
	1500m	low	0.99 [0.95, 1.03]	0.531	0.92 [0.85, 1.00]	0.062	0.91 [0.78, 1.07]	0.258
		high	1.00 [0.95, 1.04]	0.859	0.98 [0.90, 1.07]	0.663	0.94 [0.79, 1.12]	0.476
	2000m	low	0.98 [0.93, 1.02]	0.308	0.92 [0.84, 1.02]	0.102	0.86 [0.71, 1.04]	0.111
		high	1.02 [0.97, 1.07]	0.518	0.94 [0.85, 1.05]	0.266	0.87 [0.71, 1.07]	0.185
MINK	500m	low	0.79 [0.54, 1.16]	0.233	N/A	N/A	N/A	N/A
		high	1.17 [0.88, 1.56]	0.273	N/A	N/A	N/A	N/A
	1000m	low	0.82 [0.69, 0.98]	0.030	1.14 [0.83, 1.55]	0.421	N/A	N/A
		high	1.15 [0.99, 1.34]	0.073	1.08 [0.77, 1.51]	0.650	N/A	N/A
	1500m	low	0.93 [0.83, 1.04]	0.188	1.00 [0.80, 1.25]	0.981	1.12 [0.72, 1.74]	0.602
		high	1.16 [1.05, 1.29]	0.005	0.96 [0.75, 1.22]	0.743	1.23 [0.80, 1.90]	0.339
	2000m	low	0.97 [0.89, 1.05]	0.435	0.97 [0.81, 1.16]	0.727	1.35 [0.99, 1.84]	0.060
		high	1.05 [0.97, 1.14]	0.233	1.02 [0.86, 1.23]	0.799	1.23 [0.88, 1.72]	0.235

	Alcoholic liver disease		All cause mortality		Non-accidental mortality	
	HR [95% CI]	P value	HR [95% CI]	P value	HR [95% CI]	P value
	1.02 [0.88, 1.19]	0.756	0.99 [0.99, 1.00]	0.129	0.99 [0.98, 1.00]	0.066
	0.87 [0.73, 1.04]	0.127	1.00 [0.99, 1.01]	0.474	1.00 [0.99, 1.00]	0.305
	1.02 [0.86, 1.21]	0.808	0.99 [0.99, 1.00]	0.223	0.99 [0.98, 1.00]	0.149
	0.87 [0.72, 1.07]	0.181	0.99 [0.98, 1.00]	0.080	0.99 [0.98, 1.00]	0.064
	0.97 [0.76, 1.25]	0.814	1.00 [0.99, 1.01]	0.996	1.00 [0.99, 1.02]	0.814
	0.90 [0.68, 1.18]	0.433	1.00 [0.98, 1.01]	0.582	1.00 [0.98, 1.01]	0.678
	0.99 [0.67, 1.47]	0.956	0.99 [0.97, 1.01]	0.325	0.99 [0.96, 1.01]	0.237
	0.99 [0.65, 1.50]	0.972	0.99 [0.97, 1.01]	0.474	0.99 [0.97, 1.01]	0.396
	0.75 [0.58, 0.98]	0.034	0.99 [0.98, 1.01]	0.397	1.00 [0.98, 1.01]	0.480
	0.97 [0.75, 1.24]	0.779	0.98 [0.96, 0.99]	1.4e-04	0.98 [0.96, 0.99]	3.3e-04
	0.84 [0.71, 0.99]	0.043	1.01 [1.00, 1.02]	0.033	1.01 [1.00, 1.02]	0.020
	0.82 [0.68, 1.00]	0.045	1.00 [0.99, 1.01]	0.698	1.00 [0.99, 1.01]	0.687
	0.96 [0.83, 1.12]	0.640	1.02 [1.01, 1.02]	2.4e-04	1.02 [1.01, 1.02]	1.7e-04
	0.92 [0.77, 1.11]	0.393	1.01 [1.00, 1.02]	0.034	1.01 [1.00, 1.02]	0.028
	1.00 [0.86, 1.17]	0.975	1.01 [1.00, 1.02]	0.005	1.01 [1.00, 1.02]	0.006
	0.85 [0.70, 1.04]	0.107	1.02 [1.01, 1.03]	8.2e-05	1.02 [1.01, 1.03]	2.2e-04
	0.88 [0.69, 1.13]	0.316	0.98 [0.97, 1.00]	0.009	0.99 [0.97, 1.00]	0.019
	1.17 [0.94, 1.47]	0.169	1.01 [1.00, 1.02]	0.157	1.01 [1.00, 1.02]	0.160
	0.91 [0.77, 1.06]	0.221	0.99 [0.98, 1.00]	0.025	0.99 [0.98, 1.00]	0.033
	0.99 [0.84, 1.17]	0.900	1.00 [0.99, 1.00]	0.386	1.00 [0.99, 1.01]	0.432
	1.07 [0.91, 1.24]	0.410	1.00 [0.99, 1.01]	0.911	1.00 [0.99, 1.01]	0.957
	0.99 [0.84, 1.17]	0.904	1.00 [0.99, 1.01]	0.596	1.00 [0.99, 1.01]	0.711
	0.96 [0.80, 1.16]	0.697	1.00 [0.99, 1.01]	0.622	1.00 [0.99, 1.01]	0.779
	1.05 [0.86, 1.28]	0.626	1.00 [0.99, 1.01]	0.452	1.00 [0.99, 1.01]	0.669
	N/A	N/A	0.94 [0.87, 1.01]	0.082	0.94 [0.87, 1.01]	0.086
	N/A	N/A	1.00 [0.94, 1.06]	0.885	1.00 [0.94, 1.06]	0.951
	N/A	N/A	0.99 [0.96, 1.02]	0.542	0.99 [0.96, 1.02]	0.610
	N/A	N/A	0.99 [0.96, 1.03]	0.611	0.99 [0.95, 1.02]	0.474
	1.10 [0.72, 1.69]	0.662	1.01 [0.99, 1.03]	0.279	1.01 [0.99, 1.04]	0.209
	0.69 [0.40, 1.17]	0.170	1.01 [0.99, 1.04]	0.249	1.01 [0.99, 1.04]	0.297
	0.91 [0.64, 1.30]	0.599	1.01 [0.99, 1.03]	0.203	1.01 [1.00, 1.03]	0.154
	0.59 [0.38, 0.91]	0.017	1.01 [0.99, 1.03]	0.276	1.01 [0.99, 1.03]	0.375

CHAPTER 3

RESIDENTIAL PROXIMITY TO CROPS AND AGRICULTURAL PESTICIDE USE AND CAUSE-SPECIFIC MORTALITY: A PROSPECTIVE CENSUS-BASED COHORT STUDY IN THE NETHERLANDS

Mariana Simões, Anke Huss, Maartje Brouwer, Esmeralda Krop, Nicole Janssen, Roel Vermeulen

Published: Sci Total Environ. 2022 Apr;817:152932. doi: 10.1016/j.scitotenv.2022.152932

ABSTRACT

Background: There is continued concern about residential proximity to agricultural pesticide use and possible adverse health effects. Studies on this subject have been scarce with inconsistent results. We explored associations between residential proximity to specific crops, pesticide use and cause-specific mortality in a prospective census-based cohort study in The Netherlands.
Methods: Selecting inhabitants aged >30 living in less urbanized areas, at the same address for nine years up to baseline (2004) from a national register-based cohort, we followed ~3.1 million individuals for cause-specific mortality until 2012. We estimated the area of specific crop groups cultivated within buffers of 50m, 100m and 250m around each individual's residence and the amount of fungicides, herbicides and insecticides used within the same buffers for the period 1995–2003. The association between these exposure proxies and 25 primary causes of death was investigated using Cox proportional hazards regression, adjusting for individual and area-level confounders.
Results: Residential proximity to crops was associated with decreased mortality risks overall. In contrast to the overall trend an increased risk was observed for chronic lower respiratory diseases and proximity to maize cultivation. We found no evidence of an association between the amount of pesticides used and cause-specific mortality.
Conclusions: In a large prospective census-based cohort study in The Netherlands we found evidence of an increased risk of chronic lower respiratory diseases in relation to maize cultivation which was not reflected in general pesticide use, hinting to specific pesticides or practices in maize cultivation that may lead to the observed increased risk.

INTRODUCTION

It is well established that pesticide use in intensive agriculture management impacts ecosystems and (human) health. Exposure to pesticides has been linked to adverse outcomes such as neurodegenerative and respiratory diseases, reduced fertility, and various forms of cancer[1]. While these associations have been extensively studied in occupational settings, epidemiological studies on the health of residents living near agricultural plots where pesticides are applied are still scarce. Results across these were heterogeneous, but provided indications of increased occurrence of conditions ranging from skin and respiratory irritations to diseases such as Parkinson's, leukemia, and autism [2]. Previous research suggests that people living in agricultural settings are subject to higher exposure to pesticides than people living in urban areas and that residential pesticide exposure, though characteristically low-dose, has a longer duration than in typically higher-dose occupational settings[3,4]. The relatively large affected population, as compared to the occupationally exposed population, that furthermore includes vulnerable populations (children, elderly, subjects with co-morbidities), warrants the study of potential health effects of residential pesticide exposures.

We studied the possible associations between residential proximity to specific crops and estimated average use of pesticides on these crops, and cause-specific mortality using a prospective census-based cohort study with linkage to mortality records. The Netherlands provides a unique possibility to study such an association, as the high population density and intensive agricultural land use result in 46% of all residences being located within 500m of an agricultural plot. Furthermore, historical data on pesticide usage, crop maps, and a digital population registry are available since 1995.

METHODS

Study population

The Dutch Environmental Longitudinal Study (DUELS) administrative cohort includes the whole Dutch population and comprises data from several databases from Statistics Netherlands (Centraal Bureau voor de Statistiek, CBS) including mortality, individual characteristics, residential history, and neighbourhood characteristics[5]. In this study, we included adults aged ≥30 years on 01–01–2004, registered in the population register (GBA – Gemeentelijke Basisadministratie Persoonsgegevens); registration in GBA is obligatory in the Netherlands. We excluded persons who lived within 500m of the bor-

der with Germany or Belgium (for whom we were unable to compute the exposure), persons who changed address in the ten years prior to enrolment, and persons living in more urbanized areas (≥1500 addresses per km^2, at neighborhood level) (Figure 1).

Mortality endpoints

We selected 25 primary causes of death (endpoints), according to the International Classification of Diseases, 10th revision (ICD-10)[6]: 18 specific causes of death, six broadly defined groups of causes of death and any cause of death (Figure 2). This selection was based on diseases for which there were at least some prior reports about possible links with pesticide exposure and supplemented with external causes of mortality (accidents and intentional self-harm) to evaluate the difference between all-cause mortality and non-external mortality[7–12]. Reliability of causes of death statistics in the Netherlands was shown to be high (> 90% for major causes of death such as cancers and acute myocardial infarction and about 85% for respiratory diseases)[13].

Exposure: land use and pesticide use on crops

Residential exposure to agricultural pesticides was assessed for the period 1995 to 2003 by determining (a) types of crops cultivated in the vicinity of residences based on land use maps, and (b) amount of insecticides, fungicides, and herbicides likely used based on national pesticide surveys among farmers on their annual pesticide use per crop. Data and methods used to compute the exposure proxies are outlined below and described in detail in the supplementary material (S1).

Land use variables

We geocoded all residences using the Basisregistratie Adressen en Gebouwen (BAG)[14] and computed the area of specific crops, in hectares (ha), within squared areas around the residences. These areas are named buffers throughout the paper and roughly correspond to radii of 50m, 100m, 250m and 500m. The area of crops within the buffers was computed using the Landelijk Grondgebruik Nederland (LGN)[15] raster maps with a resolution of 25 × 25 m that are available every three to five years. We used the LGN maps from 1995–1997, 1999–2000, and 2003–2004 and selected seven crop groups: maize, grains, potatoes, beets, fruit (apple and pear trees), flower bulbs, and a group 'other crops' (see Table S1.1 for a detailed definition of the crops), excluding grass. We considered only crops grown in open fields, excluding thus crops grown in greenhouses. Combining those crop groups together, we created an additional group 'all crops'.

For the years in-between available LGN maps, we calculated average areas of maize, fruit, flower bulbs, and 'other crops', which are considered to be "stable crops" over the years. Potatoes, beets and grains are often cultivated in an annual rotation scheme, that is, each year one of these three crops is grown in a field followed by another of these crops the following year. We used the yearly Dutch agricultural census data provided by CBS from 1995 to 2003 to estimate the "probable" area of these "rotation crops" per region for the years that no LGN map was available[16–18]. Finally, by averaging the areas across the exposure period (1995 to 2004) we obtained four land use buffers reflecting the average area (ha) of a specific crop cultivated within 50, 100, 250 and 500m for each cohort members' residence.

Additional exposure variables reflecting land use

We additionally computed two metrics that we used to evaluate consistency of trends. *Donut variables:* We computed the area of cultivated crops around the residence as "donuts with holes" of 50–100m, 100–250m, and 250–500m. We further dichotomized these donuts into "presence" or "absence" of a specific crop as a binary variable.
Distance variable: We computed the shortest distance (in meters) to the edge of nearest crop. For unavailable LGN years we averaged the shortest distance between two available years for the stable crops and considered the averaged minimum distance to any of the three rotation crops as the "probable" distance to cultivated potatoes, cereals, and beets. As with the buffers, we averaged the computed distances across the study period. In the analyses, distance was used as a categorical variable with the classes <50m, 50–100m, 100–250m, 250–500, and ≥500m (referent).

Pesticide use variables

Using the average amount, in kilograms (kg), of annual pesticides use in each crop group as reported by a sample of farmers in 1995, 1998, 2000 and 2004 (see Table S1.1 for information on which crops were taken into account for computing the average amount of pesticides used in each crop group), we estimated the amount of insecticides, herbicides, fungicides, and the total amount of these three pesticide classes that were likely used within our buffers[19]. For the years between available data sets, we used the average of the amounts reported between two available years. We averaged amounts across the exposure period (1995 to 2004) within each buffer.

Statistical analysis

A recently conducted exposure assessment study in The Netherlands found that there was significant difference in the concentrations of several pesticides in air and house dust between people living within 250m and living beyond 500m from flower bulb crops[20]. However, there was only a weak gradient in concentrations, especially in house dust, from small distances from the field up to 250m. In this study we compared residents living within 500m of a (specific) crop ('exposed') to residents that do not have (specific) crop fields within 500m of their residences ('unexposed') in three steps (Figure 2). First, we explored mortality gradients across 50m, 100m and 250m buffers using the land use variables. Second, we identified 'findings' and 'noteworthy observations' by applying predefined criteria for the interpretation of the results (see below). Third, for the noteworthy observations we further evaluated the results of the 500m buffer analysis. As a secondary analysis, we explored mortality gradients across 50m, 100m and 250m buffers using the pesticide use variables.

Main analysis (land use variables)

For the main analysis, we used the 50m, 100m and 250m crop-specific land use buffers. In each model the 'complementary donut' (the surface that remains until 500m) of the considered crop was used as a confounder (Figure 2).
We used Cox proportional hazards regression to explore the association between the endpoints of interest and the exposure variables (Hazard Ratio (HR) and 95% confidence interval (CI) per 1 hectare increase in crop area). Models used one-year age strata and time-in-study as time-scale, similarly to what was done by Fischer et al. (2015) This approach yields the same results as using age as the time-scale with the advantage that running such models is computationally faster. We specified three models with increasing degrees of covariate adjustment (see Table 1 for the categories used in the confounder variables):

- *basic model*, adjusted for sex
- *individual model*, basic model further adjusted for origin, marital status, standardized household income (an individual socioeconomic indicator adjusted for differences in household size and composition), and the presence (yes/no) of other crops within 500m of the individual's residence (except for when *all crops* was the exposure)
- *full model*, individual model further adjusted for urbanization degree at neighborhood level as defined in the "Wijk- en buurtkaart" (neighborhood map) from 1999 and 2003, the proportion of low educated residents in the neighborhood in 2007

and socioeconomic position (SEP) at four-digit postcode level from 2002 (a social status indicator derived every four years by the Netherlands Institute for Social Research that takes income level, unemployment rate and educational level into account)[21]

Each individual was assigned nine years of "exposure" period [01–01–1995 to 31–12-2003], followed by 1 year of lag period [01–01–2004 to 31–12–2004], to allow for a latency period. We used a latency period to exclude deaths caused by an accidental acute exposure, although this is unlikely to affect results, given the long exposure period assessed and that many of the mortality endpoints featured in this study are chronic diseases. The follow-up period was from 01–01–2005 to 31–12–2012. Follow-up was terminated at the end of the follow-up period, at the time of death or when individuals were lost to follow-up, whatever came first.

To account for multiple testing, we computed the Benjamini-Hochberg false discovery rate (FDR) adjusted P values, considering 25 tests (one for each outcome, crops were considered independent) and used a threshold of $q<0.1$.

Complementary analyses

To support the main analysis results, we ran models using the donuts, as continuous (area of cultivated crop, in hectares) and binary (presence of crop in buffer or donut, yes/no) variables, and the distance, as a categorical variable (Figure 2).

Criteria for interpretation

The results of each exposure-endpoint combination were evaluated using the following *a priori* defined criteria:

- *Consistency among buffer models*: did the associations between the buffers have the same direction?
- *Consistency between the three land use based models*: did associations of the buffers, donut, and distance models have the same general direction?
- *Statistical significance*: in the full model, was there a result with a Q-value lower than 0.1 in at least one of the buffers?

We considered a *finding* if results met all abovementioned criteria. In addition, we considered as *noteworthy observations* associations that met only two criteria.

Sensitivity analyses

We performed two sensitivity analyses, using the full model of the main (buffer) analysis. First, we restricted the analyses to non-urban residents (that is, people living in neighborhoods with <1000 addresses per km^2), to assess potential bias from a semi-urban environment. Second, we excluded people that worked for at least one year in agriculture in the period 1999–2003, in order to exclude a possible influence of occupational exposure. For this analysis we linked the DUELS cohort to microdata on employment and self-employment available from CBS.

Additional analyses with pesticide use variables

For completeness and transparency, we performed similar analyses as described above but using amount (kg) of insecticides, herbicides, fungicides, and their sum ("total pesticides") used on the area of cultivated land within the buffers around residences (as a continuous variable).

Software

The exposure metrics and statistical analyses were performed using STATA/MP 14, ArcGIS 10.4 and R version 3.4.1 (2017–06–30).

RESULTS

A total of 3 160 231 persons contributed to our analysis (Figure 1). During the follow-up period 16 154 (0.5%) cohort members were lost to follow-up (mainly due to emigration) and 353 730 (11.2%) died. Table 1 describes the demographic characteristics of the study and exposed populations. About 81% (n=2 560 479) of the individuals were exposed to at least one type of crop, that is, resided within 500m of a crop (more descriptive statistics in S2).

Overall, there was no clear evidence of associations between living closer to crops and higher risk of death from a range of causes of death. In effect, 10 findings emerged, all showing decreased Hazard Ratios (HRs). We further identified 56 noteworthy observations, most (~79%) also with HRs below unity. Figure 3 shows the results for the main analyses with findings and noteworthy observations indicated with ‡ and †, respectively. We noted that 11 noteworthy observations met the 3 criteria when we included the 500m buffer on the evaluation of the results, lending strong support that these

could be indications of associations (Figure 4). Of these, only one association showed increased risk of mortality: living near maize and death from chronic lower respiratory diseases. All results from the main and complementary analyses can be found in S3 and S4. Sensitivity analyses restricting to people living in non-urban areas and excluding people that worked for at least one year in the agriculture did not materially change estimates (S6).
Correlation between the amount of insecticides, herbicides, fungicides, and the total amount of pesticides was high, as well as correlations between them and area of "all crops" (S2). Results from pesticide use information analyses (S5) yielded no evidence of associations with our mortality endpoints, except an indication for decreased risk of lung cancer mortality in people living near crops where insecticides were applied.

DISCUSSION

We used a nationwide prospective census-based cohort study to explore possible associations of presence of crops and estimated use of pesticides near residences, with cause-specific mortality. We observed an overall slightly lower mortality risk when investigating proximity to specific crops and no evidence of associations with the amount of pesticides used. In contrast to this general picture, a few associations emerged that indicated possible associations, namely living close to maize and higher chronic lower respiratory diseases mortality risk.
We assured independence of self-reported exposure data and precluded selection bias by including the entire Dutch population (aged 30+). Even after selecting individuals living in less urbanized areas who had not moved in the ten years prior to enrolment, we obtained a large study population (~3.1 million). The eight year follow-up (2005–2012) could account for the long period of onset and worsening of disease that usually precedes mortality. We had access to rather unique historical datasets pertaining to pesticide use and information on land use with national coverage and across several years that matched the long timeframe evaluated in our study (1995–2003). These data enabled exposure proxies assignment to each cohort member. To the best of our knowledge, this makes it one of the largest studies on the topic to date.
Residential exposure to pesticides was not measured but rather investigated based on the area and proximity of crops to residences. Although using crop acreage as proxy for residential exposure provides lower specificity than biological or environmental measurements, previous studies have shown that it can be useful in estimating pesticide levels in residential homes located near crops[3]. By assigning exposure to a registered place of residence, exposure misclassification might have occurred. We

did not consider other relevant locations for exposure, mobility of the individuals nor their presence at the residence during spraying events, which may influence exposure levels. However, the main alternative exposure location is probably the workplace and our sensitivity analysis excluding agricultural workers showed no major differences in effect estimates. Other sources of exposure such as domestic use of pesticides or nutritional exposure were not considered but are not likely to differ materially within the short distances analyzed here, limiting residual confounding. We could not account for wind speed and direction, which affect spread of pesticides applied in fields, but it is difficult to determine what the best non-symmetrical buffer could be. Prevailing wind in the Netherlands is West to South West, but lower speed Eastern winds are also important for stable lower spread of pesticides at short distances. We also did not have information on pesticide application methods nor on adoption of risk mitigation measures, such as presence of buffer zone, direction and height of release, sprayer speed and spraying pressure applied. These techniques can affect exposure levels and, although most are featured in the products' labels, it is difficult to determine whether they were systematically implemented in the period covered in our study (1995–2004). This period precedes the adoption of the Directive on Sustainable Use of pesticides (SUD, Directive 2009/128/EC) in the European Union, a framework designed to reduce the risks and impacts of pesticides on human health and the environment. This directive encompasses measures such as training of users, inspection of pesticide application equipment, prohibition of aerial spraying, limitation of pesticide use in sensitive areas, and information and awareness raising about pesticide risks. Regulation (EC) No 1107/2009 also took effect in 2009, and regulates the placing of pesticides on the market. According to this regulation, pesticides may enter the market only if they do not pose immediate or delayed risk to human and animal health and the environment. However, approval may be subject to conditions and restrictions (such as risk mitigation measures and monitoring during and after use). These important legislation requirements were not enforced prior to 2009, meaning that nowadays residential exposure may be lower than that of the study period. We observed overall similarity of risk estimates in residents living close to potato, beet, and grain crops, likely due to assumptions regarding rotation schemes for these crops; results should not be interpreted independently. Table 2 shows each crop's top ten most used active ingredients in 1995, 1998, 2000 and 2004. It was not possible to identify clear active ingredient or chemical group candidates for explaining potential effects of a specific crop. Not only were several active ingredients' usage introduced or discontinued over the long exposure window assessed here, but active ingredients have also different modes of action, toxicity levels and environmental persistence, making it difficult to disentangle their effects.

We observed many effect estimates below unity, indicating possible residual confounding. If residential proximity to crops had no influence on mortality, we would expect a similar proportion of risk estimates above and below unity. Because we made use of administrative databases, we could not account for individual level lifestyle factors such as smoking, alcohol use, physical activity, or Body Mass Index (BMI). It remains unclear why generally living closer to crops is associated with lower mortality or why some missing confounders would be associated with our exposure proxies that were based on very short distances. Indicative however is that lung cancer, which is driven predominantly by smoking (about 90% of cases are attributed to smoking [22]), showed HRs below unity. This may suggest that smoking behaviors are differential even at this scale. As chronic lower respiratory diseases are also associated with smoking, the increased risk we observed is therefore all the more noteworthy. Interestingly, negative associations have previously been reported for Dutch rural populations, such as associations between non-accidental mortality and air pollution from agricultural sources and between indicators of livestock farm emissions and respiratory health[23,24]. Finally, there are two potential sources of confounding that we did not take into account: one was whether people had pre-existing health conditions at baseline; the other was that people may choose residences both further away and nearer to more rural areas for health reasons.

We found no associations when using exposure proxies based on pesticide use, except for a negative association between insecticide use and lung cancer mortality. One reason for this general lack of associations could be a dilution of the overall effect on mortality when using a variable that averages effects across a broad range of pesticides (herbicides, insecticides, fungicides). By using specific crops as exposure proxies, we retain a higher level of specificity of pesticides used as they are often specific to crops and farming methods.

We applied an objective procedure to address possible false positives arising from the high number of tests performed, with the possible limitation of overlooking a signal. We identified 11 noteworthy observations for which including the 500m buffer analysis lent support for possible associations. Particularly, the increased risk of death from chronic lower respiratory diseases in the presence of maize suggests that something about maize crops is associated with this mortality endpoint. Further analyses strengthened this result. First, the association was consistent across different regions of the Netherlands indicating specificity of this association to maize and not other geospatially related factors (S7). Second, after controlling for the presence of livestock animals the estimates became (slightly) stronger, with results meeting the original interpretation criteria (S8). Lastly, in a validation effort, we applied the same approach using data from the Dutch Occupational and Environmental Health Cohort Study (AMI-

GO) [25] where we were able to further control for smoking, alcohol use, BMI, having a job, educational level, and air pollution at individual level (S9). We observed an Odds Ratio of 1.09 (95% Confidence Interval 0.89, 1.33) for COPD incidence (which accounted for 98% of chronic lower respiratory diseases deaths in our study population, S2) in the 250m maize buffer.

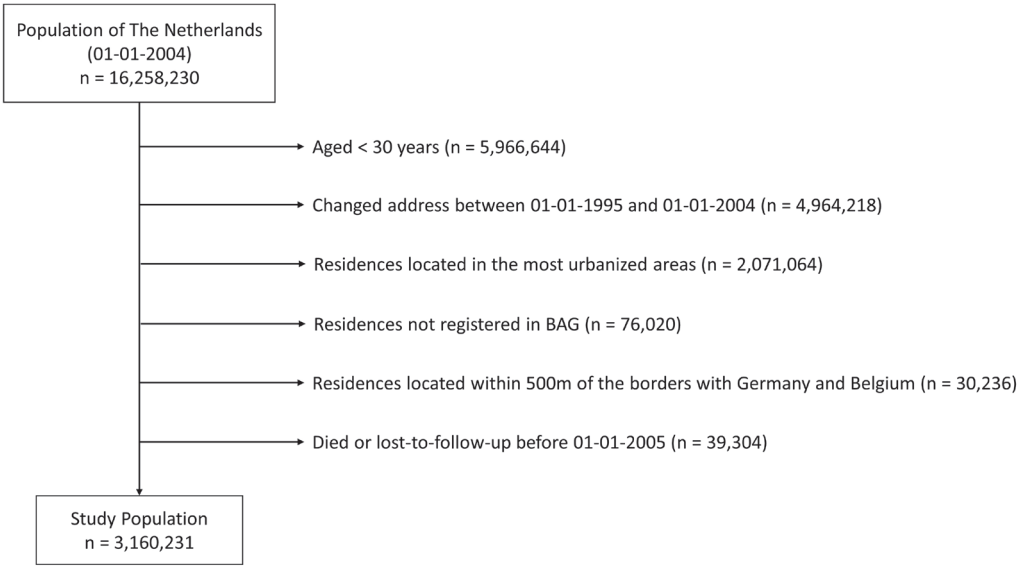
Various studies have reported associations between occupational exposure to pesticides and respiratory health problems ranging from cough and wheezing to COPD, reduced lung function, and asthma [11,26]. One study indicated an association between exposure to atrazine (the most used active ingredient in maize in 1995 and 1998, see Table 2) and increased wheezing among farmer pesticide applicators[27]. Few studies that were conducted in the general population suggested increased risks for (exacerbation of) asthma, bronchitis, and impaired lung function, with many focusing on children as a particularly vulnerable group[28].

In conclusion, we observed a potential association between living near maize and death from chronic lower respiratory diseases, supported by additional analyses. The increased risk was not reflected in general pesticide use hinting to specific pesticides or practices in maize cultivation that may lead to the observed increased risk.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Health, Welfare and Sport (VWS), in the context of the Policy Advisory on Plant Protection Products. We would like to thank Statistics Netherlands (CBS) for providing remote access to a secured environment for statistical analyses and Rob Vijftigschild (CBS) for his assistance regarding the pesticide usage data.

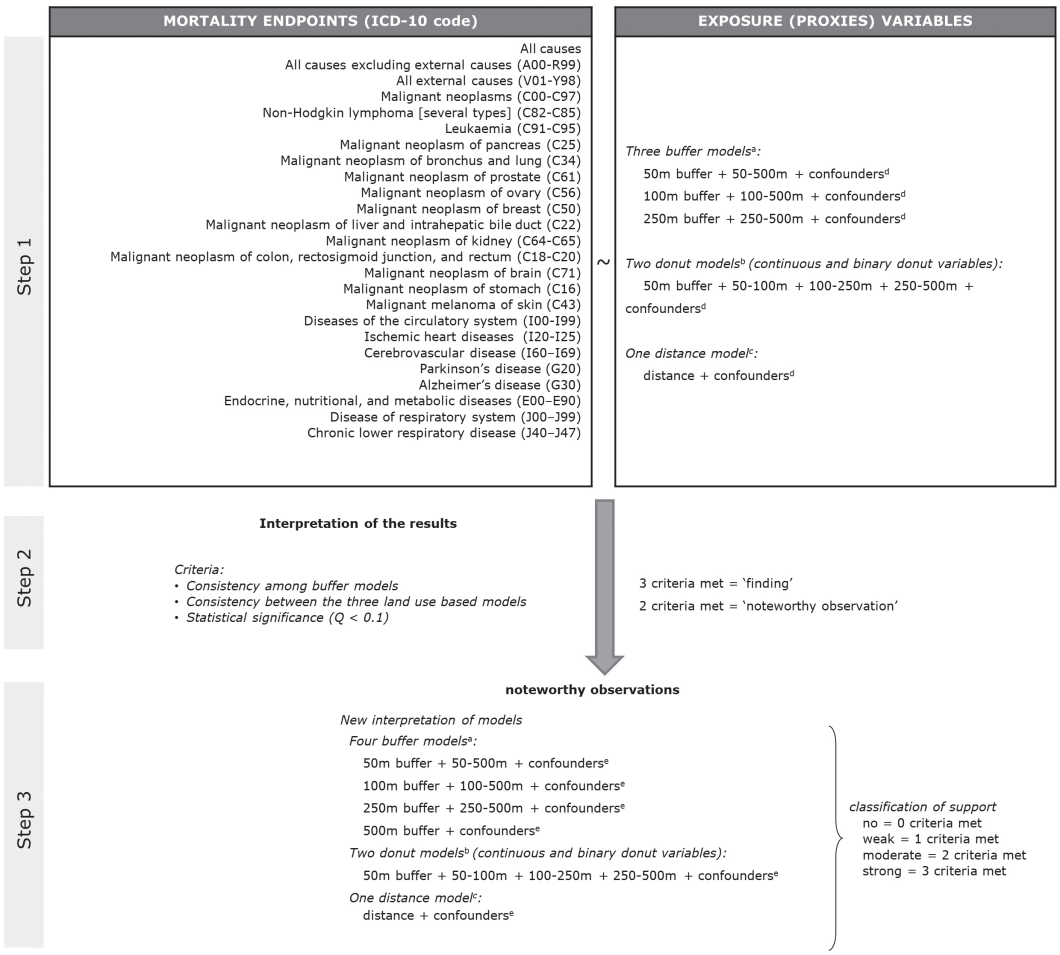
FIGURES



BAG = Basisregistratie Adressen en Gebouwen, the cadastral dataset containing all addresses in the Netherlands used to compute individual residential exposure proxies.

Figure 1: Exclusion criteria.

Figure 2: General framework of analyses using the land use exposure proxies.



Step 1: main analysis with buffer land use exposure (proxy) variables, complemented with analysis using the donut and distance land use exposure (proxy) variables.

Step 2: evaluation of the results from Step 1 by applying the a priori defined criteria for interpretation; identification of findings and noteworthy observations.

Step 3: evaluation of the noteworthy observations identified in Step 2 by including the results of the 500m buffer analysis in the interpretation and applying the a priori defined criteria once more; classification into "no", "weak", "moderate" and "strong" support of an association when including the 500m buffer analysis.

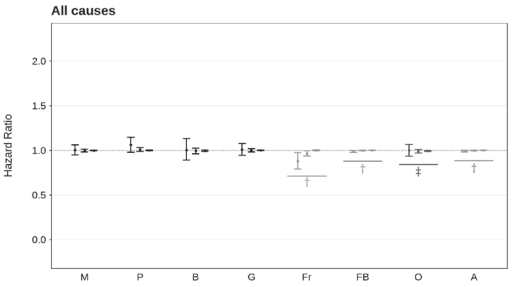
a Buffer variables reflecting land use (area of cultivated crop, in hectares).

b Donut variables reflecting land use; continuous variables = area of cultivated crop, in hectares; binary variables = presence of crop in buffer or donut, yes/no.

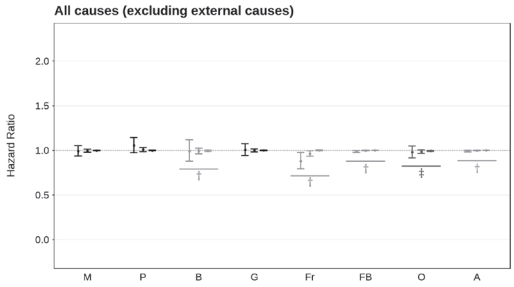
c Distance variable reflecting land use, in meters; categorical variable with classes "<50m", "50–100m", "100–250m", "250–500m" and ">500m" (referent).

d Confounders:
- basic model: age + sex
- individual model: age + sex + ethnicity + marital status + standardized household income + other crops
- full model: age + sex + ethnicity + marital status + standardized household income + other crops + urbanization degree at neighbourhood level + neighbourhood socioeconomic position + proportion of low educated residents in the neighbourhood

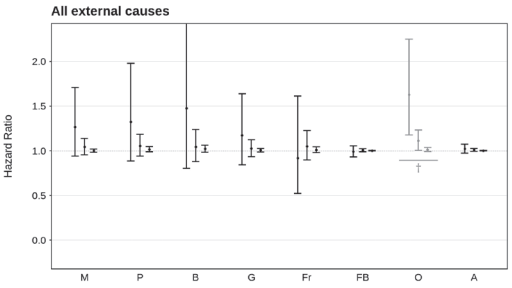
e Full model.



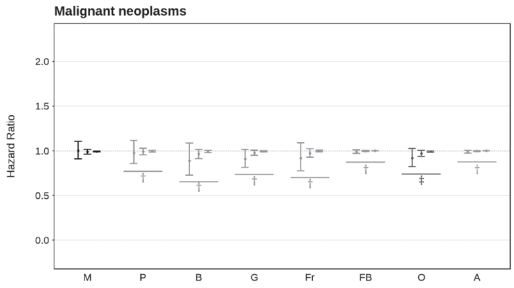
	M	P	B	G	Fr	FB	O	A
50m	1.00 (0.98)	1.06 (0.64)	1.00 (0.97)	1.01 (0.98)	0.88 (0.20)	0.98 (0.13)	1.00 (0.97)	0.99 (0.39)
100m	1.00 (0.90)	1.01 (0.88)	0.99 (0.91)	1.00 (0.99)	0.96 (0.27)	1.00 (0.18)	0.99 (0.61)	1.00 (0.24)
250m	1.00 (0.31)	1.00 (0.78)	0.99 (0.57)	1.00 (0.67)	1.00 (0.97)	1.00 (0.95)	**0.99 (0.00)**	1.00 (0.63)



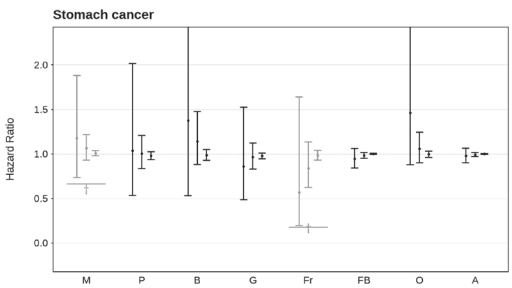
	M	P	B	G	Fr	FB	O	A
50m	0.99 (0.92)	1.05 (0.64)	0.99 (0.95)	1.00 (0.98)	0.88 (0.20)	0.98 (0.13)	0.98 (0.79)	0.99 (0.39)
100m	0.99 (0.87)	1.01 (0.88)	0.99 (0.91)	1.00 (0.99)	0.96 (0.27)	0.99 (0.18)	0.98 (0.60)	1.00 (0.24)
250m	1.00 (0.29)	1.00 (0.65)	0.99 (0.55)	1.00 (0.67)	1.00 (0.97)	1.00 (0.95)	**0.99 (0.00)**	1.00 (0.63)



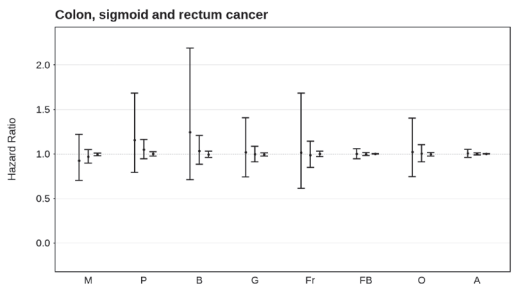
	M	P	B	G	Fr	FB	O	A
50m	1.27 (0.84)	1.32 (0.64)	1.48 (0.77)	1.18 (0.73)	0.92 (0.93)	0.99 (0.94)	**1.63 (0.09)**	1.02 (0.66)
100m	1.04 (0.87)	1.05 (0.88)	1.04 (0.91)	1.02 (0.97)	1.05 (0.88)	1.00 (0.91)	1.11 (0.39)	1.01 (0.56)
250m	1.00 (0.97)	1.02 (0.65)	1.02 (0.73)	1.01 (0.75)	1.01 (0.97)	1.00 (0.95)	1.01 (0.51)	1.00 (0.63)



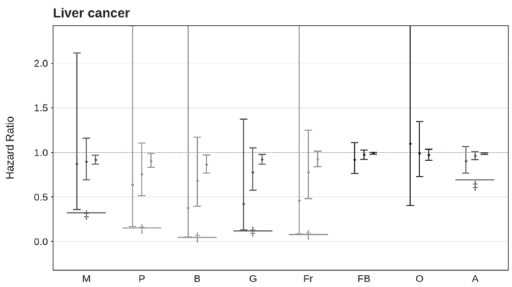
	M	P	B	G	Fr	FB	O	A
50m	1.00 (0.98)	0.98 (0.90)	0.89 (0.77)	0.91 (0.73)	0.92 (0.76)	0.99 (0.72)	0.92 (0.58)	0.99 (0.47)
100m	0.99 (0.87)	0.99 (0.88)	0.96 (0.85)	0.98 (0.70)	0.97 (0.63)	1.00 (0.64)	0.97 (0.49)	1.00 (0.44)
250m	0.99 (0.19)	0.99 (0.65)	0.99 (0.57)	1.00 (0.67)	1.00 (0.97)	1.00 (0.95)	**0.99 (0.06)**	1.00 (0.63)



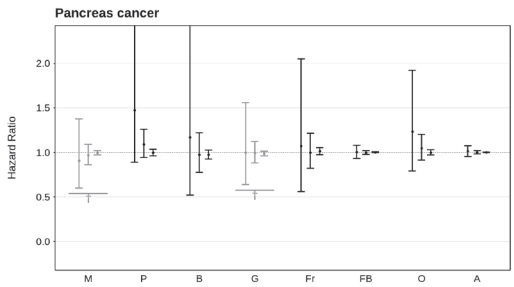
	M	P	B	G	Fr	FB	O	A
50m	1.18 (0.92)	1.03 (0.95)	1.38 (0.80)	0.86 (0.89)	0.57 (0.76)	0.94 (0.79)	1.46 (0.58)	0.98 (0.84)
100m	1.06 (0.87)	1.00 (0.98)	1.14 (0.85)	0.96 (0.97)	0.84 (0.63)	0.98 (0.64)	1.06 (0.83)	0.99 (0.79)
250m	1.01 (0.97)	0.98 (0.70)	0.98 (0.81)	0.98 (0.67)	0.98 (0.97)	1.00 (0.95)	0.99 (0.86)	1.00 (0.67)



	M	P	B	G	Fr	FB	O	A
50m	0.92 (0.92)	1.15 (0.78)	1.24 (0.77)	1.02 (0.98)	1.02 (0.99)	1.00 (0.98)	1.02 (0.97)	1.00 (0.93)
100m	0.97 (0.87)	1.05 (0.88)	1.03 (0.92)	0.99 (0.99)	0.98 (0.98)	1.00 (0.95)	1.00 (0.99)	1.00 (0.95)
250m	0.99 (0.97)	1.00 (0.94)	0.99 (0.81)	0.99 (0.67)	1.00 (0.97)	1.00 (0.95)	0.99 (0.86)	1.00 (0.88)



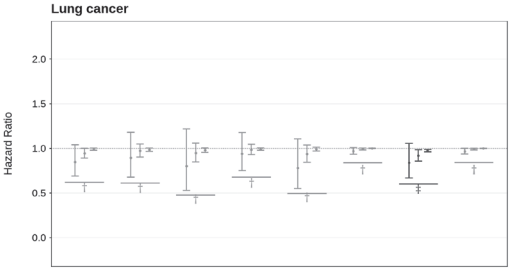
	M	P	B	G	Fr	FB	O	A
50m	0.87 (0.92)	0.64 (0.79)	0.38 (0.77)	0.42 (0.73)	0.46 (0.78)	0.92 (0.80)	1.10 (0.97)	0.90 (0.66)
100m	0.89 (0.87)	0.75 (0.88)	0.68 (0.85)	0.78 (0.69)	0.77 (0.63)	0.97 (0.64)	0.99 (0.99)	0.96 (0.44)
250m	**0.92 (0.04)**	0.90 (0.20)	0.86 (0.14)	**0.92 (0.08)**	0.92 (0.86)	0.99 (0.84)	0.97 (0.54)	**0.99 (0.08)**



	M	P	B	G	Fr	FB	O	A
50m	0.91 (0.92)	1.47 (0.64)	1.17 (0.90)	1.00 (0.99)	1.07 (0.96)	1.00 (0.98)	1.23 (0.66)	1.01 (0.91)
100m	0.97 (0.90)	1.09 (0.88)	0.97 (0.94)	0.99 (0.99)	1.00 (0.98)	1.00 (0.95)	1.05 (0.83)	1.00 (0.97)
250m	0.99 (0.97)	1.00 (0.92)	0.97 (0.73)	0.99 (0.67)	1.01 (0.97)	1.00 (0.95)	1.00 (0.93)	1.00 (0.91)

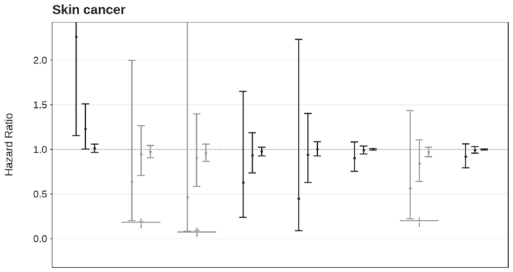
Figure 3: Results obtained from the full model analyses using the land use buffers (main analysis).

Lung cancer



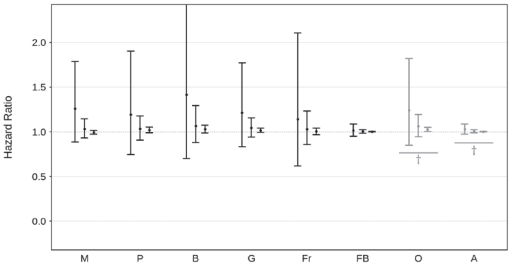
50m	0.84 (0.84)	0.89 (0.78)	0.80 (0.77)	0.94 (0.89)	0.78 (0.65)	0.97 (0.53)	0.84 (0.58)	0.96 (0.39)
100m	0.94 (0.73)	0.97 (0.88)	0.95 (0.85)	0.98 (0.97)	0.94 (0.63)	0.99 (0.64)	0.92 (0.39)	0.99 (0.24)
250m	0.99 (0.39)	0.98 (0.65)	0.98 (0.56)	0.99 (0.67)	0.99 (0.97)	1.00 (0.95)	**0.97 (0.00)**	1.00 (0.21)

Skin cancer



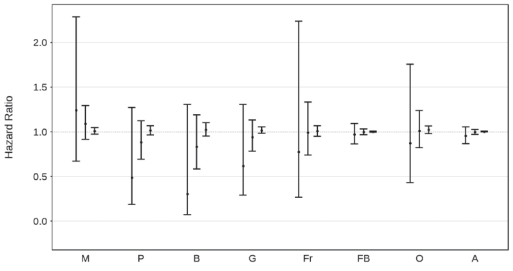
50m	2.26 (0.50)	0.64 (0.78)	0.46 (0.77)	0.63 (0.73)	0.45 (0.76)	0.90 (0.72)	0.56 (0.62)	0.92 (0.66)
100m	1.23 (0.73)	0.94 (0.88)	0.90 (0.91)	0.93 (0.97)	0.94 (0.98)	0.99 (0.95)	0.84 (0.61)	0.99 (0.86)
250m	1.01 (0.97)	0.97 (0.78)	0.96 (0.81)	0.97 (0.67)	1.00 (0.97)	1.00 (0.95)	0.97 (0.52)	1.00 (0.84)

Breast cancer



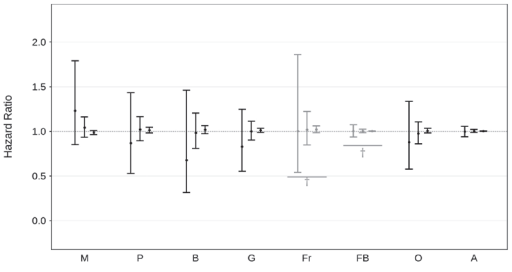
50m	1.26 (0.92)	1.19 (0.78)	1.41 (0.77)	1.21 (0.73)	1.14 (0.86)	1.01 (0.86)	1.24 (0.62)	1.03 (0.66)
100m	1.03 (0.90)	1.03 (0.88)	1.06 (0.91)	1.04 (0.97)	1.03 (0.98)	1.00 (0.95)	1.06 (0.64)	1.01 (0.79)
250m	0.99 (0.97)	1.02 (0.65)	1.03 (0.73)	1.02 (0.67)	1.00 (0.97)	1.00 (0.95)	1.02 (0.26)	1.00 (0.69)

Ovary cancer



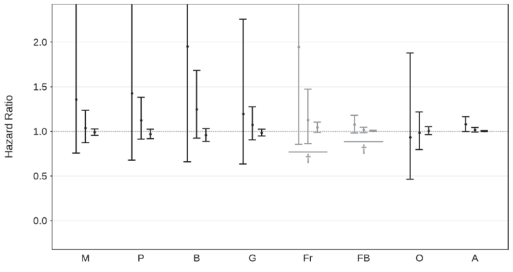
50m	1.24 (0.92)	0.49 (0.64)	0.30 (0.77)	0.62 (0.73)	0.77 (0.86)	0.97 (0.86)	0.87 (0.89)	0.96 (0.66)
100m	1.09 (0.87)	0.88 (0.88)	0.83 (0.85)	0.94 (0.97)	0.99 (0.98)	1.00 (0.95)	1.01 (0.99)	1.00 (0.92)
250m	1.01 (0.97)	1.01 (0.83)	1.02 (0.81)	1.02 (0.67)	1.01 (0.97)	1.00 (0.95)	1.02 (0.54)	1.00 (0.70)

Prostate cancer



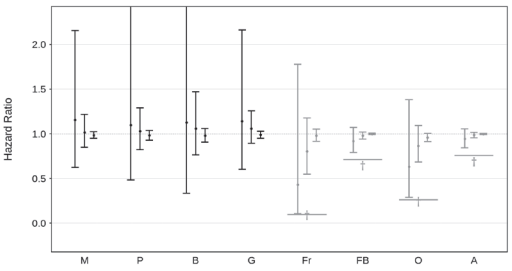
50m	1.23 (0.92)	0.87 (0.81)	0.68 (0.77)	0.83 (0.73)	1.00 (0.99)	1.00 (0.98)	0.88 (0.79)	0.99 (0.93)
100m	1.04 (0.88)	1.02 (0.88)	0.98 (0.97)	1.00 (0.99)	1.02 (0.98)	1.00 (0.95)	0.97 (0.96)	1.00 (0.89)
250m	0.98 (0.81)	1.01 (0.81)	1.02 (0.81)	1.01 (0.67)	1.02 (0.97)	1.00 (0.95)	1.01 (0.86)	1.00 (0.63)

Kidney cancer



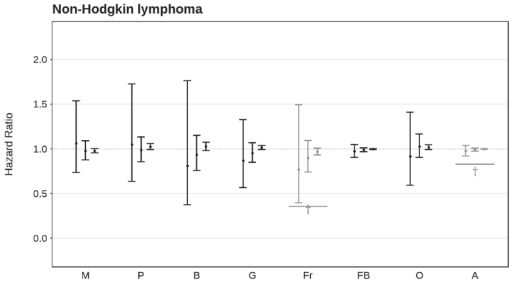
50m	1.36 (0.92)	1.42 (0.78)	1.95 (0.77)	1.19 (0.89)	1.94 (0.63)	1.08 (0.53)	0.93 (0.97)	1.08 (0.43)
100m	1.04 (0.90)	1.12 (0.88)	1.25 (0.85)	1.07 (0.97)	1.13 (0.72)	1.01 (0.72)	0.98 (0.99)	1.02 (0.59)
250m	0.99 (0.97)	0.97 (0.65)	0.95 (0.73)	0.98 (0.67)	1.04 (0.86)	1.00 (0.84)	1.01 (0.86)	1.00 (0.67)

Brain cancer



50m	1.16 (0.92)	1.10 (0.92)	1.13 (0.95)	1.14 (0.96)	0.43 (0.76)	0.92 (0.72)	0.63 (0.62)	0.94 (0.66)
100m	1.01 (0.94)	1.03 (0.88)	1.06 (0.93)	1.06 (0.97)	0.80 (0.63)	0.98 (0.64)	0.86 (0.61)	0.98 (0.64)
250m	0.98 (0.97)	0.98 (0.81)	0.98 (0.81)	0.99 (0.75)	0.98 (0.97)	1.00 (0.95)	0.96 (0.32)	1.00 (0.67)

Non-Hodgkin lymphoma



50m	1.06 (0.92)	1.05 (0.92)	0.81 (0.84)	0.87 (0.89)	0.77 (0.81)	0.97 (0.86)	0.91 (0.89)	0.97 (0.69)
100m	0.98 (0.90)	0.98 (0.88)	0.93 (0.91)	0.95 (0.97)	0.90 (0.63)	0.99 (0.64)	1.03 (0.96)	0.99 (0.59)
250m	0.98 (0.29)	1.02 (0.65)	1.02 (0.73)	1.01 (0.67)	0.97 (0.86)	1.00 (0.84)	1.02 (0.51)	1.00 (0.63)

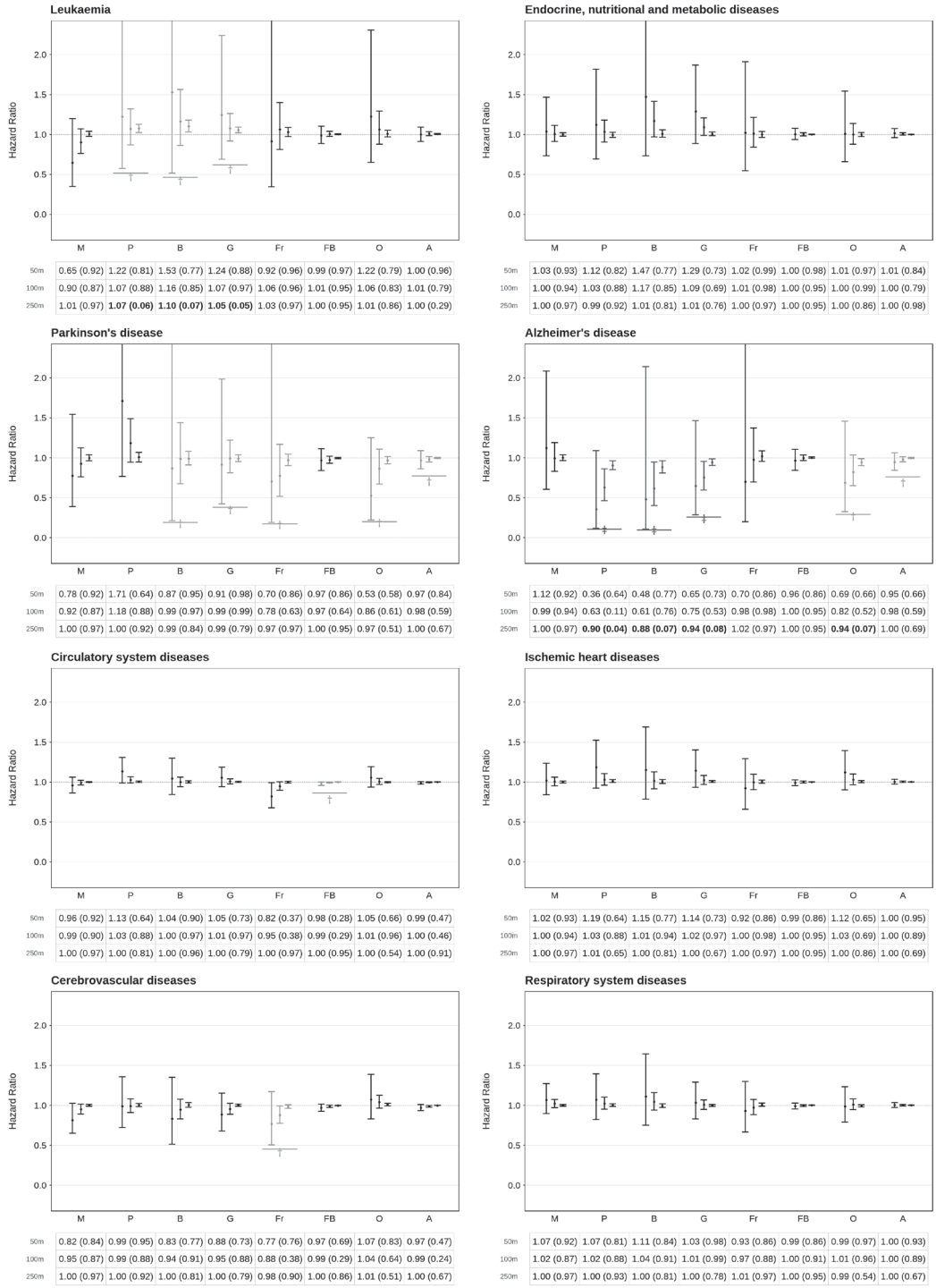
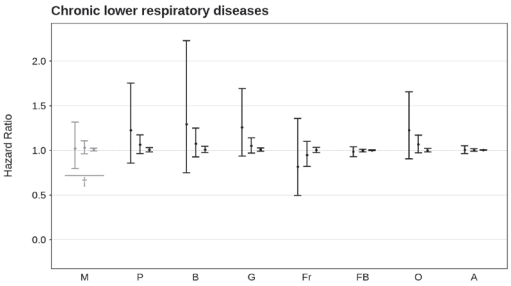


Figure 3: Results obtained from the full model analyses using the land use buffers (main analysis). (cont.)



	M	P	B	G	Fr	FB	O	A
50m	1.02 (0.93)	1.23 (0.74)	1.29 (0.77)	1.26 (0.73)	0.82 (0.81)	0.98 (0.86)	1.22 (0.59)	1.01 (0.93)
100m	1.03 (0.87)	1.06 (0.88)	1.08 (0.85)	1.05 (0.97)	0.95 (0.83)	1.00 (0.95)	1.07 (0.61)	1.00 (0.80)
250m	1.01 (0.97)	1.01 (0.83)	1.01 (0.81)	1.01 (0.67)	1.00 (0.97)	1.00 (0.95)	1.00 (0.94)	1.00 (0.67)

The plots display the Hazard Ratios per increase in 1 hectare of area of (specific) crop and their 95% Confidence Intervals. The tables below each plot show the corresponding Hazard Ratios and adjusted P values for each buffer.
M=Maize, P=Potatoes, B=Beets, G=Grains, Fr=Fruit, FB=Flower bulbs, O='Other crops', A='All crops'

‡ Finding: all three interpretation criteria met

† Noteworthy finding: two out of the three interpretation criteria met

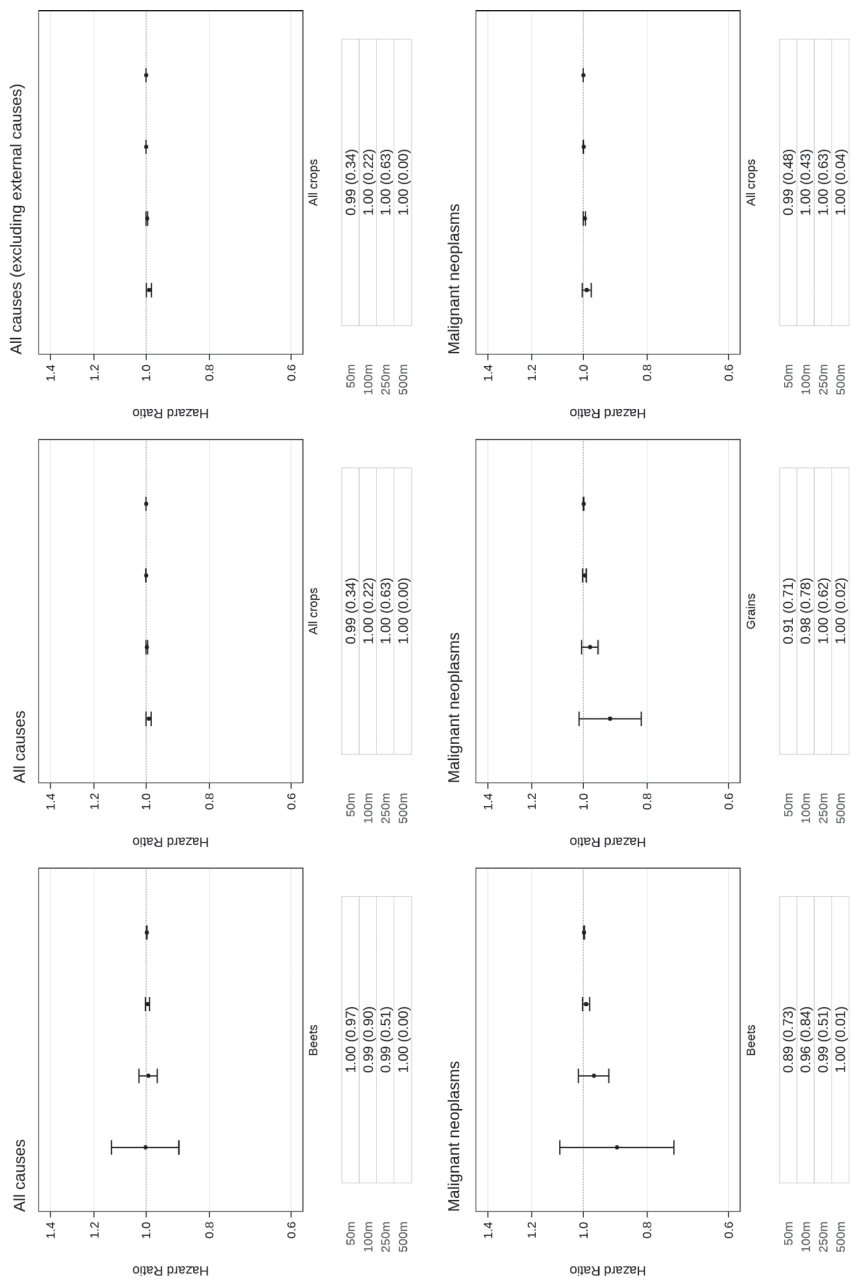
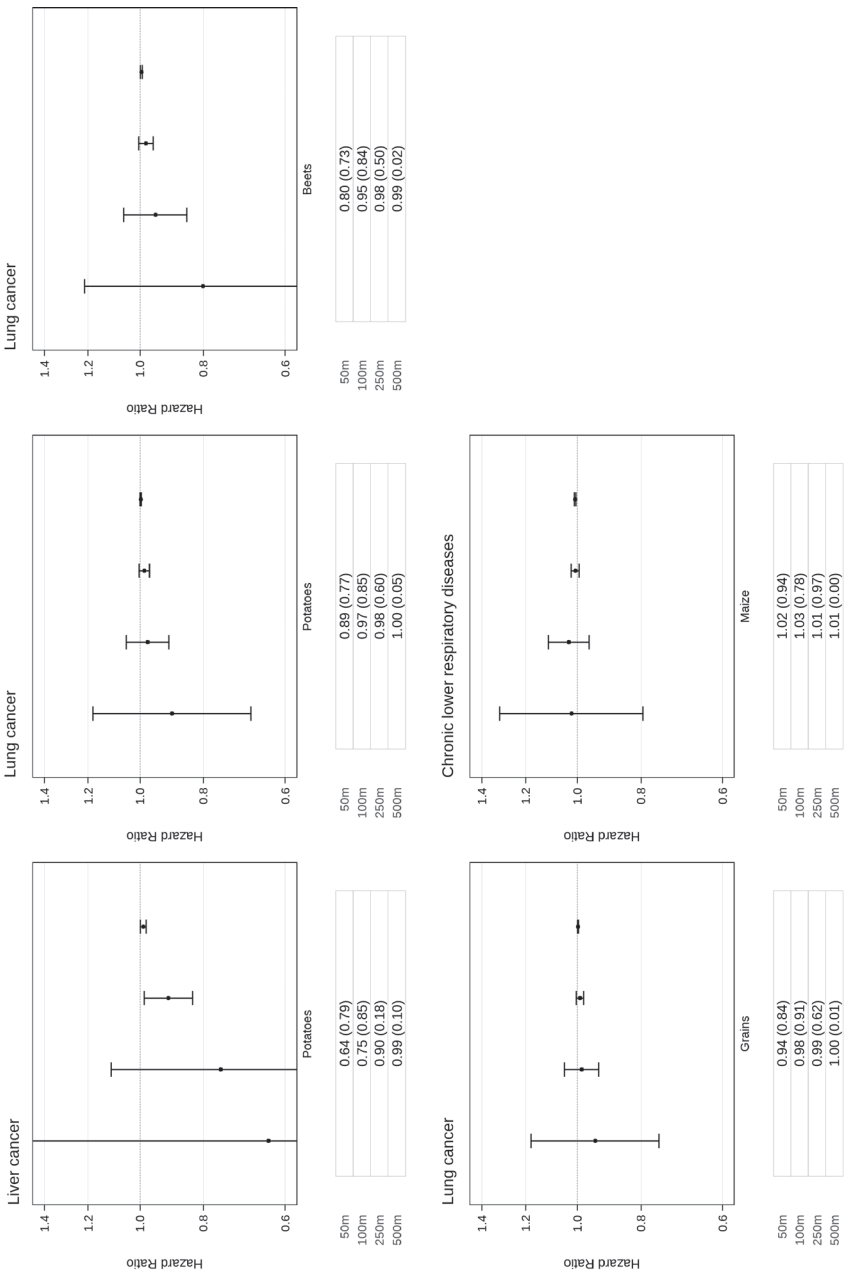


Figure 4: Results obtained from the full model analyses of the 11 noteworthy findings for which including the 500m buffer analysis in the interpretation provided strong support of an association.



The plots display the Hazard Ratios per increase in 1 hectare of area of (specific) crop and their 95% Confidence Intervals. The tables below each plot show the corresponding Hazard Ratios and FDR adjusted P values for each buffer.
M=Maize, P=Potatoes, B=Beets, G=Grains, A='All crops'

TABLES

Table 1: Demographic characteristics at baseline for the study population and for the exposed population (living within 500m of a (specific) crop).

Characteristic	Study Population N=3 160 231	Population exposed in 500m			
		Maize N=1 952 844 (61.8%)	Potatoes N=1 763 090 (55.8%)	Beets N=1 762 973 (55.8%)	
AGE [MEAN (SD)]	56.4 (13.0)	55.8 (12.9)	55.9 (12.9)	55.9 (12.9)	
SEX [N (%)]					
Female	1 636 224 (51.8%)	998 936 (51.2%)	901 554 (51.1%)	901 495 (51.1%)	
Male	1 524 007 (48.2%)	953 908 (48.8%)	861 536 (48.9%)	861 478 (48.9%)	
ORIGIN [N (%)]					
Dutch	2 895 820 (91.6%)	1 803 319 (92.3%)	1 624 804 (92.2%)	1 624 693 (92.2%)	
Western	222 813 (7.1%)	130 356 (6.7%)	121 718 (6.9%)	121 712 (6.9%)	
Non-Western	41 598 (1.3%)	19 169 (1.0%)	16 568 (0.9%)	16 568 (0.9%)	
CIVIL STATUS [N (%)]					
Married/partner	2 460 560 (77.9%)	1 544 100 (79.1%)	1 390 506 (78.9%)	1 390 427 (78.9%)	
Widowed	279 995 (8.9%)	161 887 (8.3%)	147 504 (8.4%)	147 493 (8.4%)	
Divorced	151 099 (4.8%)	81 848 (4.2%)	75 372 (4.3%)	75 363 (4.3%)	
Single	268 495 (8.5%)	164 952 (8.4%)	149 660 (8.5%)	149 642 (8.5%)	
Unknown	82 (0.0%)	57 (0.0%)	48 (0.0%)	48 (0.0%)	
STANDARDIZED HOUSEHOLD INCOME [N (%)]					
<1 percentile	21 769 (0.7%)	14 991 (0.8%)	13 870 (0.8%)	13 870 (0.8%)	
1-<5 percentile	32 888 (1.0%)	21 661 (1.1%)	19 751 (1.1%)	19 748 (1.1%)	
5-<10 percentile	64 038 (2.0%)	40 972 (2.1%)	37 479 (2.1%)	37 469 (2.1%)	
10-<25 percentile	348 853 (11.0%)	215 575 (11.0%)	196 010 (11.1%)	195 999 (11.1%)	
25-<50 percentile	787 914 (24.9%)	484 923 (24.8%)	441 507 (25.0%)	441 472 (25.0%)	
50-<75 percentile	895 094 (28.3%)	558 642 (28.6%)	504 861 (28.6%)	504 833 (28.6%)	
75-<90 percentile	596 649 (18.9%)	368 858 (18.9%)	330 468 (18.7%)	330 454 (18.7%)	
90-<95 percentile	206 749 (6.5%)	125 576 (6.4%)	111 974 (6.4%)	111 964 (6.4%)	
95-<99 percentile	161 374 (5.1%)	95 642 (4.9%)	84 230 (4.8%)	84 228 (4.8%)	
99–100 percentile	42 385 (1.3%)	24 656 (1.3%)	21 526 (1.2%)	21 526 (1.2%)	
Unknown	2 518 (0.1%)	1348 (0.1%)	1414 (0.1%)	1410 (0.1%)	
URBANIZATION DEGREE AT NEIGHBOURHOOD LEVEL [N (%)]					
<500 addresses per km2	1 115 598 (35.3%)	914 830 (46.8%)	852 708 (48.4%)	852 591 (48.4%)	
500–1000 addresses per km2	1 031 621 (32.6%)	637 661 (32.7%)	546 815 (31.0%)	546 815 (31.0%)	
1000–1500 addresses per km2	1 013 012 (32.1%)	400 353 (20.5%)	363 567 (20.6%)	363 567 (20.6%)	

	Population exposed in 500m				
	Grains N=1 763 090 (55.8%)	**Fruits N=645 178 (20.4%)**	**Flower bulbs N=848 690 (26.9%)**	**'Other crops' N=2 004 662 (63.4%)**	**'All crops' N=2 560 479 (81.0%)**
	55.9 (12.9)	56.0 (12.9)	55.9 (12.9)	56.0 (12.9)	56.0 (12.9)
	901 554 (51.1%)	331 585 (51.4%)	435 245 (51.3%)	1 027 789 (51.3%)	1 316 424 (51.4%)
	861 536 (48.9%)	313 593 (48.6%)	413 445 (48.7%)	976 873 (48.7%)	1 244 055 (48.6%)
	1 624 804 (92.2%)	590 709 (91.6%)	780 841 (92.0%)	1 846 515 (92.1%)	2 357 490 (92.1%)
	121 718 (6.9%)	48 024 (7.4%)	59 584 (7.0%)	136 962 (6.8%)	174 633 (6.8%)
	16 568 (0.9%)	6445 (1.0%)	8265 (1.0%)	21 185 (1.1%)	28 356 (1.1%)
	1 390 506 (78.9%)	509 213 (78.9%)	668 286 (78.7%)	1 578 769 (78.8%)	2 012 413 (78.6%)
	147 504 (8.4%)	54 076 (8.4%)	69 823 (8.2%)	167 567 (8.4%)	217 129 (8.5%)
	75 372 (4.3%)	28 406 (4.4%)	37 745 (4.4%)	88 582 (4.4%)	114 527 (4.5%)
	149 660 (8.5%)	53 465 (8.3%)	72 817 (8.6%)	169 693 (8.5%)	216 341 (8.4%)
	48 (0.0%)	18 (0.0%)	19 (0.0%)	51 (0.0%)	69 (0.0%)
	13 870 (0.8%)	4 534 (0.7%)	6 621 (0.8%)	14 905 (0.7%)	18 868 (0.7%)
	19 751 (1.1%)	6 738 (1.0%)	9 393 (1.1%)	21 607 (1.1%)	27 566 (1.1%)
	37 479 (2.1%)	13 104 (2.0%)	17 449 (2.1%)	41 387 (2.1%)	52 668 (2.1%)
	196 010 (11.1%)	66 631 (10.3%)	89 266 (10.5%)	221 004 (11.0%)	281 748 (11.0%)
	441 507 (25.0%)	158 771 (24.6%)	206 076 (24.3%)	496 738 (24.8%)	635 220 (24.8%)
	504 861 (28.6%)	183 781 (28.5%)	241 394 (28.4%)	571 363 (28.5%)	729 697 (28.5%)
	330 468 (18.7%)	125 399 (19.4%)	164 571 (19.4%)	379 958 (19.0%)	484 368 (18.9%)
	111 974 (6.4%)	43 658 (6.8%)	57 439 (6.8%)	130 312 (6.5%)	166 891 (6.5%)
	84 230 (4.8%)	33 691 (5.2%)	44 528 (5.2%)	100 215 (5.0%)	128 376 (5.0%)
	21 526 (1.2%)	8305 (1.3%)	11 180 (1.3%)	25 556 (1.3%)	33 041 (1.3%)
	1414 (0.1%)	566 (0.1%)	773 (0.1%)	1617 (0.1%)	2036 (0.1%)
	852 708 (48.4%)	327 765 (50.8%)	422 376 (49.8%)	865 183 (43.2%)	1 062 635 (41.5%)
	546 815 (31.0%)	192 643 (29.9%)	245 306 (28.9%)	657 700 (32.8%)	844 491 (33.0%)
	363 567 (20.6%)	124 770 (19.3%)	181 008 (21.3%)	481 779 (24.0%)	653 353 (25.5%)

Table 1: Demographic characteristics at baseline for the study population and for the exposed population (living within 500m of a (specific) crop). (cont.)

Characteristic	Study Population N=3 160 231	Population exposed in 500m			
		Maize N=1 952 844 (61.8%)	Potatoes N=1 763 090 (55.8%)	Beets N=1 762 973 (55.8%)	
NEIGHBOURHOOD'S SOCIOECONOMIC POSITION [N (%)]					
1st quintile	238 333 (7.5%)	113 077 (5.8%)	118 329 (6.7%)	118 329 (6.7%)	
2nd quintile	618 665 (19.6%)	372 508 (19.1%)	366 165 (20.8%)	366 124 (20.8%)	
3rd quintile	863 510 (27.3%)	584 617 (29.9%)	516 560 (29.3%)	516 560 (29.3%)	
4th quintile	848 470 (26.8%)	549 053 (28.1%)	473 550 (26.9%)	473 550 (26.9%)	
5th quintile	568 566 (18.0%)	315 132 (16.1%)	272 222 (15.4%)	272 222 (15.4%)	
Unknown	22 687 (0.7%)	18 457 (0.9%)	16 264 (0.9%)	16 188 (0.9%)	
PROPORTION OF LOW EDUCATED RESIDENTS IN NEIGHBOURHOOD [N (%)]					
1st quintile	390 595 (12.4%)	219 099 (11.2%)	173 092 (9.8%)	173 092 (9.8%)	
2nd quintile	672 915 (21.3%)	413 278 (21.2%)	358 125 (20.3%)	358 125 (20.3%)	
3rd quintile	752 191 (23.8%)	483 987 (24.8%)	450 998 (25.6%)	450 957 (25.6%)	
4th quintile	746 164 (23.6%)	471 065 (24.1%)	434 486 (24.6%)	434 410 (24.6%)	
5th quintile	597 476 (18.9%)	364 819 (18.7%)	345 796 (19.6%)	345 796 (19.6%)	
Unknown	890 (0.0%)	596 (0.0%)	593 (0.0%)	593 (0.0%)	

sd = standard deviation

	Population exposed in 500m				
	Grains N=1 763 090 (55.8%)	**Fruits N=645 178 (20.4%)**	**Flower bulbs N=848 690 (26.9%)**	**'Other crops' N=2 004 662 (63.4%)**	**'All crops' N=2 560 479 (81.0%)**
	118 329 (6.7%)	24 340 (3.8%)	30 784 (3.6%)	123 354 (6.2%)	164 846 (6.4%)
	366 165 (20.8%)	121 036 (18.8%)	154 968 (18.3%)	390 990 (19.5%)	500 888 (19.6%)
	516 560 (29.3%)	193 232 (30.0%)	249 977 (29.5%)	573 059 (28.6%)	728 673 (28.5%)
	473 550 (26.9%)	192 471 (29.8%)	258 251 (30.4%)	556 387 (27.8%)	696 550 (27.2%)
	272 222 (15.4%)	110 842 (17.2%)	149 552 (17.6%)	345 908 (17.3%)	448 546 (17.5%)
	16 264 (0.9%)	3257 (0.5%)	5158 (0.6%)	14 964 (0.7%)	20 976 (0.8%)
	173 092 (9.8%)	67 524 (10.5%)	91 146 (10.7%)	207 357 (10.3%)	287 487 (11.2%)
	358 125 (20.3%)	149 394 (23.2%)	202 109 (23.8%)	415 876 (20.7%)	550 107 (21.5%)
	450 998 (25.6%)	158 094 (24.5%)	213 308 (25.1%)	496 378 (24.8%)	626 177 (24.5%)
	434 486 (24.6%)	162 620 (25.2%)	208 647 (24.6%)	505 004 (25.2%)	616 997 (24.1%)
	345 796 (19.6%)	107 444 (16.7%)	133 336 (15.7%)	379 393 (18.9%)	478 937 (18.7%)
	593 (0.0%)	102 (0.0%)	144 (0.0%)	654 (0.0%)	774 (0.0%)

Table 2: Top 10 most used active ingredients in 1995, 1998, 2000 and 2004 on each of the seven crop groups studied.

Crop	1995			
	Active ingredient	**Type pesticide**	**Chemical group**	**Amount (kg) used per hectare**
MAIZE	atrazine	herbicide	triazine	0.70
	bentazon	herbicide	benzothiadiazinone	0.46
	pyridate	herbicide	phenyl-pyridazine	0.26
	metolachlor	herbicide	chloroacetamide	0.25
POTATOES	paraquat-dichloride	herbicide	quarternary ammonium compound	0.43
	maneb	fungicide	dithio-carbamates and relatives	2.87
	pencycuron	fungicide	phenylurea	0.47
	monolinuron	herbicide	urea	0.13
	fentin acetate	fungicide	tri-phenyl tin compounds	0.65
	cymoxanil	fungicide	cyanoacetamide-oxime	0.22
	mancozeb	fungicide	dithio-carbamates and relatives	3.05
	fluazinam	fungicide	2,6-dinitro-anilines	0.57
	esfenvalerate	insecticide	pyrethroids pyrethrins	0.01
	metribuzin	herbicide	triazinone	0.10
BEETS	phenmedipham	herbicide	phenyl-carbamate	0.23
	ethofumesate	herbicide	benzofuran	0.28
	metamitron	herbicide	triazinone	0.82
	chloridazon	herbicide	pyridazinone (inhibition of photosynthesis at photo-system ii)	0.84
	desmedipham	herbicide	phenyl-carbamate	0.01
	triallate	herbicide	thiocarbamate	0.09
	clopyralid	herbicide	pyridine carboxylic acid	0.01
	parathion-ethyl	insecticide	organophosphates	0.03
	fluazifop-p-butyl	herbicide	aryloxyphenoxypropionate	0.01

	1998			
	Active ingredient	**Type pesticide**	**Chemical group**	**Amount (kg) used per hectare**
	atrazine	herbicide	triazine	0.48
	sulcotrion	herbicide	triketone	0.17
	pyridate	herbicide	phenyl-pyridazine	0.26
	bentazon	herbicide	benzothiadiazinone	0.18
	metolachlor	herbicide	chloroacetamide	0.38
	bromoxynil	herbicide	nitrile	0.02
	dicamba	herbicide	benzoic acid (synthetic auxins)	0.02
	fluazinam	fungicide	2,6-dinitro-anilines	0.96
	pencycuron	fungicide	phenylurea	0.73
	cymoxanil	fungicide	cyanoacetamide-oxime	0.23
	mancozeb	fungicide	dithio-carbamates and relatives	4.22
	chlorothalonil	fungicide	chloronitriles (phthalonitriles)	0.98
	propamocarb hydrochloride	fungicide	carbamates	0.98
	maneb	fungicide	dithio-carbamates and relatives	1.59
	fentin acetate	fungicide	tri-phenyl tin compounds	0.42
	metribuzin	herbicide	triazinone	0.20
	deltamethrin	insecticide	pyrethroids pyrethrins	0.01
	ethofumesate	herbicide	benzofuran	0.44
	phenmedipham	herbicide	phenyl-carbamate	0.31
	metamitron	herbicide	triazinone	1.17
	desmedipham	herbicide	phenyl-carbamate	0.02
	chloridazon	herbicide	pyridazinone (inhibition of photosynthesis at photosystem ii)	0.33
	haloxyfop-p-methyl	herbicide	aryloxyphenoxypropionate	0.01
	triallate	herbicide	thiocarbamate	0.07
	clopyralid	herbicide	pyridine carboxylic acid	0.01
	triflusulfuron-methyl	herbicide	sulfonylurea	0.01
	lenacil	herbicide	uracil	0.01

Table 2: Top 10 most used active ingredients in 1995, 1998, 2000 and 2004 on each of the seven crop groups studied. (cont.)

Crop	1995			
	Active ingredient	**Type pesticide**	**Chemical group**	**Amount (kg) used per hectare**
GRAINS	mcpa	herbicide	phenoxy-carboxylic-acid	0.47
	mecoprop-p	herbicide	aryloxyalkanoic acid	0.39
	fenpropimorph	fungicide	morpholines	0.19
	fluroxypyr	herbicide	pyridine carboxylic acid	0.05
	propiconazole	fungicide	triazoles	0.11
	epoxiconazole	fungicide	triazoles	0.06
	metsulfuron-methyl	herbicide	sulfonylurea	0.01
	dimethoate	insecticide	organophosphates	0.09
	isoproturon	herbicide	urea	0.37
	bifenox	herbicide	diphenylether (ppo)	0.10
FRUITS	captan	fungicide	phthalimides	16.28
	fenoxycarb	insecticide	fenoxycarb	0.08
	pirimicarb	insecticide	carbamates	0.20
	tolylfluanid	fungicide	sulfamides	2.46
	glyphosate	herbicide	glycine	0.74
	copper oxychloride	fungicide	inorganic compound	1.75
	difenoconazole	fungicide	triazoles	0.09
	diuron	herbicide	urea	0.42
	nitrothal isopropyl	fungicide	unclassified	0.54
	simazin	herbicide	triazine	0.29
FLOWER BULBS	metoxuron	herbicide	urea	1.31
	metamitron	herbicide	triazinone	1.15
	prochloraz	fungicide	imidazoles	0.65
	chlorpropham	herbicide	carbamates	1.39
	chlorothalonil	fungicide	chloronitriles (phthalonitriles)	1.42
	chloridazon	herbicide	pyridazinone (inhibition of photosynthesis at photo-system ii)	0.82
	propoxur	insecticide	carbamates	0.82
	carbendazim	fungicide	benzimidazoles	0.39
	esfenvalerate	insecticide	pyrethroids pyrethrins	0.03
	maneb	fungicide	dithio-carbamates and relatives	6.48

	1998			
	Active ingredient	**Type pesticide**	**Chemical group**	**Amount (kg) used per hectare**
	mcpa	herbicide	phenoxy-carboxylic-acid	0.38
	epoxiconazole	fungicide	triazoles	0.07
	fenpropimorph	fungicide	morpholines	0.19
	kresoxim-methyl	fungicide	oximino-acetates	0.08
	fluroxypyr	herbicide	pyridine carboxylic acid	0.06
	isoproturon	herbicide	urea	0.51
	mecoprop-p	herbicide	aryloxyalkanoic acid	0.34
	metsulfuron-methyl	herbicide	sulfonylurea	0.02
	propiconazole	fungicide	triazoles	0.09
	dicamba	herbicide	benzoic acid (synthetic auxins)	0.02
	captan	fungicide	phthalimides	11.83
	glyphosate	herbicide	glycine	0.88
	tolylfluanid	fungicide	sulfamides	3.08
	fenoxycarb	insecticide	fenoxycarb	0.10
	copper oxychloride	fungicide	inorganic compound	1.93
	triadimenol	fungicide	triazoles	0.06
	amitraz	insecticide	amitraz	0.51
	pirimicarb	insecticide	carbamates	0.08
	pyrimethanil	fungicide	anilino-pyrimidines	0.31
	dithianon	fungicide	quinones (anthra-quinones)	1.12
	metamitron	herbicide	triazinone	1.29
	chlorpropham	herbicide	carbamates	1.71
	prochloraz	fungicide	imidazoles	0.80
	carbendazim	fungicide	benzimidazoles	0.41
	asulam	herbicide	carbamates	0.88
	chlorothalonil	fungicide	chloronitriles (phthalonitriles)	1.12
	maneb	fungicide	dithio-carbamates and relatives	4.90
	chloridazon	herbicide	pyridazinone (inhibition of photosynthesis at photosystem ii)	0.77
	esfenvalerate	insecticide	pyrethroids pyrethrins	0.03
	simazin	herbicide	triazine	0.10

Table 2: Top 10 most used active ingredients in 1995, 1998, 2000 and 2004 on each of the seven crop groups studied. (cont.)

Crop	1995			
	Active ingredient	**Type pesticide**	**Chemical group**	**Amount (kg) used per hectare**
OTHER CROPS	chlorpropham	herbicide	carbamates	0.45
	maneb	fungicide	dithio-carbamates and relatives	2.15
	propachlor	herbicide	chloroacetamide	1.38
	propyzamide	herbicide	benzamide (microtubule assembly inhibition)	0.80
	pirimicarb	insecticide	carbamates	0.10
	bentazon	herbicide	benzothiadiazinone	0.34
	metoxuron	herbicide	urea	1.49
	propazin	herbicide	triazine	1.01
	pendimethalin	herbicide	dinitroaniline	0.45
	oxydemeton-methyl	insecticide	organophosphates	0.15

Crop	2000			
	Active ingredient	**Type pesticide**	**Chemical group**	**Amount (kg) used per hectare**
MAIZE	sulcotrion	herbicide	triketone	0.20
	terbutylazine	herbicide	triazine	0.22
	pyridate	herbicide	phenyl-pyridazine	0.14
	nicosulfuron	herbicide	sulfonylurea	0.03
	bromoxynil	herbicide	nitrile	0.05
	dicamba	herbicide	benzoic acid (synthetic auxins)	0.01
POTATOES	fluazinam	fungicide	2,6-dinitro-anilines	1.06
	cymoxanil	fungicide	cyanoacetamide-oxime	0.36
	mancozeb	fungicide	dithio-carbamates and relatives	4.43
	pencycuron	fungicide	phenylurea	0.70
	metribuzin	herbicide	triazinone	0.11
	maneb	fungicide	dithio-carbamates and relatives	0.78
	imazalil	fungicide	imidazoles	0.14
	lambda-cyhalothrin	insecticide	pyrethroids pyrethrins	0.01
	chlorothalonil	fungicide	chloronitriles (phthalonitriles)	0.69
	paraquat-dichloride	herbicide	quarternary ammonium compound	0.12

1998			
Active ingredient	**Type pesticide**	**Chemical group**	**Amount (kg) used per hectare**
maneb	fungicide	dithio-carbamates and relatives	3.24
zineb	fungicide	dithio-carbamates and relatives	3.36
bentazon	herbicide	benzothiadiazinone	0.26
propachlor	herbicide	chloroacetamide	1.80
propyzamide	herbicide	benzamide (microtubule assembly inhibition)	0.77
pendimethalin	herbicide	dinitroaniline	0.35
asulam	herbicide	carbamates	0.58
deltamethrin	insecticide	pyrethroids pyrethrins	0.01
chlorpropham	herbicide	carbamates	0.55
metoxuron	herbicide	urea	1.08

2004			
Active ingredient	**Type pesticide**	**Chemical group**	**Amount (kg) used per hectare**
nicosulfuron	herbicide	sulfonylurea	0.05
sulcotrion	herbicide	triketone	0.16
terbutylazine	herbicide	triazine	0.14
dimethenamid-p	herbicide	chloroacetamide	0.22
pyridate	herbicide	phenyl-pyridazine	0.06
mesotrione	herbicide	triketone	0.03
bromoxynil	herbicide	nitrile	0.01
dicamba	herbicide	benzoic acid (synthetic auxins)	0.02
bentazon	herbicide	benzothiadiazinone	0.05
fluazinam	fungicide	2,6-dinitro-anilines	0.90
diquat dibromide	herbicide	bipyridylium	0.37
cymoxanil	fungicide	cyanoacetamide-oxime	0.42
mancozeb	fungicide	dithio-carbamates and relatives	4.66
metribuzin	herbicide	triazinone	0.17
rimsulfuron	herbicide	sulfonylurea	0.01
pencycuron	fungicide	phenylurea	0.49
paraquat-dichloride	herbicide	quarternary ammonium compound	0.17
lambda-cyhalothrin	insecticide	pyrethroids pyrethrins	0.01
linuron	herbicide	urea	0.18

Table 2: Top 10 most used active ingredients in 1995, 1998, 2000 and 2004 on each of the seven crop groups studied. (cont.)

Crop	2000			
	Active ingredient	**Type pesticide**	**Chemical group**	**Amount (kg) used per hectare**
BEETS	metamitron	herbicide	triazinone	1.43
	phenmedipham	herbicide	phenyl-carbamate	0.30
	ethofumesate	herbicide	benzofuran	0.36
	desmedipham	herbicide	phenyl-carbamate	0.02
	clopyralid	herbicide	pyridine carboxylic acid	0.03
	haloxyfop-p-methyl	herbicide	aryloxyphenoxypropionate	0.01
	triflusulfuron-methyl	herbicide	sulfonylurea	0.01
	glyphosate	herbicide	glycine	0.18
	triallate	herbicide	thiocarbamate	0.06
GRAINS	mcpa	herbicide	phenoxy-carboxylic-acid	0.43
	epoxiconazole	fungicide	triazoles	0.10
	fluroxypyr	herbicide	pyridine carboxylic acid	0.10
	kresoxim-methyl	fungicide	oximino-acetates	0.05
	fenpropimorph	fungicide	morpholines	0.24
	metsulfuron-methyl	herbicide	sulfonylurea	0.02
	isoproturon	herbicide	urea	0.36
	mecoprop-p	herbicide	aryloxyalkanoic acid	0.23
	propiconazole	fungicide	triazoles	0.06
	lambda-cyhalothrin	insecticide	pyrethroids pyrethrins	0.00
FRUITS	tolylfluanid	fungicide	sulfamides	4.30
	dithianon	fungicide	quinones (anthra-quinones)	1.23
	fenoxycarb	insecticide	fenoxycarb	0.12
	pirimicarb	insecticide	carbamates	0.17
	glyphosate	herbicide	glycine	1.04
	triadimenol	fungicide	triazoles	0.03
	captan	fungicide	phthalimides	4.58
	imidacloprid	insecticide	neonicotinoids	0.04
	mcpa	herbicide	phenoxy-carboxylic-acid	0.57
	carbendazim	fungicide	benzimidazoles	0.40

2004			
Active ingredient	**Type pesticide**	**Chemical group**	**Amount (kg) used per hectare**
phenmedipham	herbicide	phenyl-carbamate	0.30
ethofumesate	herbicide	benzofuran	0.44
metamitron	herbicide	triazinone	1.27
desmedipham	herbicide	phenyl-carbamate	0.03
chloridazon	herbicide	pyridazinone (inhibition of photosynthesis at photosystem ii)	0.42
triflusulfuron-methyl	herbicide	sulfonylurea	0.02
s-metolachlor	herbicide	chloroacetamide	0.23
glyphosate	herbicide	glycine	0.22
clopyralid	herbicide	pyridine carboxylic acid	0.01
haloxyfop-p-methyl	herbicide	aryloxyphenoxypropionate	0.01
mcpa	herbicide	phenoxy-carboxylic-acid	0.49
epoxiconazole	fungicide	triazoles	0.09
fluroxypyr	herbicide	pyridine carboxylic acid	0.11
metsulfuron-methyl	herbicide	sulfonylurea	0.00
fenpropimorph	fungicide	morpholines	0.18
isoproturon	herbicide	urea	0.87
picoxystrobin	fungicide	methoxy-acrylates	0.12
kresoxim-methyl	fungicide	oximino-acetates	0.05
lambda-cyhalothrin	insecticide	pyrethroids pyrethrins	0.00
mecoprop-p	herbicide	aryloxyalkanoic acid	0.18
captan	fungicide	phthalimides	8.04
tolylfluanid	fungicide	sulfamides	4.42
dithianon	fungicide	quinones (anthra-quinones)	0.94
triadimenol	fungicide	triazoles	0.07
glyphosate	herbicide	glycine	0.93
pirimicarb	insecticide	carbamates	0.24
fenoxycarb	insecticide	fenoxycarb	0.10
imidacloprid	insecticide	neonicotinoids	0.04
thiram	fungicide	carbamates	4.39
thiram	fungicide	dithio-carbamates and relatives	4.39

Table 2: Top 10 most used active ingredients in 1995, 1998, 2000 and 2004 on each of the seven crop groups studied. (cont.)

Crop	2000			
	Active ingredient	**Type pesticide**	**Chemical group**	**Amount (kg) used per hectare**
FLOWER BULBS	metamitron	herbicide	triazinone	2.12
	asulam	herbicide	carbamates	1.63
	chlorpropham	herbicide	carbamates	1.61
	prochloraz	fungicide	imidazoles	0.41
	fluazinam	fungicide	2,6-dinitro-anilines	0.62
	maneb	fungicide	dithio-carbamates and relatives	3.94
	chloridazon	herbicide	pyridazinone (inhibition of photosynthesis at photo-system ii)	0.39
	chlorothalonil	fungicide	chloronitriles (phthalonitriles)	0.73
	esfenvalerate	insecticide	pyrethroids pyrethrins	0.05
	mancozeb	fungicide	dithio-carbamates and relatives	4.46
OTHER CROPS	propachlor	herbicide	chloroacetamide	2.18
	bentazon	herbicide	benzothiadiazinone	0.50
	propyzamide	herbicide	benzamide (microtubule assembly inhibition)	0.79
	metoxuron	herbicide	urea	1.49
	thiometon	insecticide	organophosphates	0.36
	pendimethalin	herbicide	dinitroaniline	0.68
	asulam	herbicide	carbamates	0.65
	dimethoate	insecticide	organophosphates	0.31
	maneb	fungicide	dithio-carbamates and relatives	2.26
	zineb	fungicide	dithio-carbamates and relatives	2.38

2004			
Active ingredient	**Type pesticide**	**Chemical group**	**Amount (kg) used per hectare**
metoxuron	herbicide	urea	1.84
metamitron	herbicide	triazinone	1.64
asulam	herbicide	carbamates	2.65
chlorpropham	herbicide	carbamates	1.43
fluazinam	fungicide	2,6-dinitro-anilines	0.84
mancozeb	fungicide	dithio-carbamates and relatives	7.14
tebuconazole	fungicide	triazoles	0.30
chloridazon	herbicide	pyridazinone (inhibition of photosynthesis at photosystem ii)	0.70
prochloraz	fungicide	imidazoles	0.27
glyphosate	herbicide	glycine	0.61
pendimethalin	herbicide	dinitroaniline	0.84
propyzamide	herbicide	benzamide (microtubule assembly inhibition)	0.99
mancozeb	fungicide	dithio-carbamates and relatives	5.94
vinclozolin	fungicide	dicarboximides	0.24
chlorpropham	herbicide	carbamates	0.49
metoxuron	herbicide	urea	1.36
deltamethrin	insecticide	pyrethroids pyrethrins	0.01
bentazon	herbicide	benzothiadiazinone	0.50
dimethoate	insecticide	organophosphates	0.31
carbetamide	herbicide	carbamates	0.40

REFERENCES

1. Kim K-H, Kabir E, Jahan SA. Exposure to pesticides and the associated human health effects. Sci Total Environ. 2017 Jan;575:525–35.
2. Dereumeaux C, Fillol C, Quenel P, Denys S. Pesticide exposures for residents living close to agricultural lands: A review. Environ Int. 2020 Jan;134:105210.
3. Ward MH, Lubin J, Giglierano J, Colt JS, Wolter C, Bekiroglu N, et al. Proximity to crops and residential to agricultural herbicides in Iowa. Environ Health Perspect. 2006;114(6):893–7.
4. Lu C, Fenske RA, Simcox NJ, Kalman D. Pesticide exposure of children in an agricultural community: evidence of household proximity to farmland and take home exposure pathways. Environ Res. 2000 Nov;84(3):290–302.
5. Fischer PH, Marra M, Ameling CB, Hoek G, Beelen R, de Hoogh K, et al. Air Pollution and Mortality in Seven Million Adults: The Dutch Environmental Longitudinal Study (DUELS). Environ Health Perspect. 2015 Jul;123(7):697–704.
6. World Health Organization. ICD-10 : international statistical classification of diseases and related health problems : tenth revision. 2nd ed. World Health Organization; 2004.
7. Bassil KL, Vakil C, Sanborn M, Cole DC, Kaur JS, Kerr KJ. Cancer health effects of pesticides: systematic review. Can Fam Physician. 2007 Oct;53(10):1704–11.
8. Wahab A, Hod R, Ismail NH, Omar N. The effect of pesticide exposure on cardiovascular system: a systematic review. Int J Community Med Public Heal Vol 3, No 1 January 2016. 2017;
9. Brouwer M, Huss A, van der Mark M, Nijssen PCG, Mulleners WM, Sas AMG, et al. Environmental exposure to pesticides and the risk of Parkinson's disease in the Netherlands. Environ Int [Internet]. 2017;107:100–10. Available from: http://www.sciencedirect.com/science/article/pii/S0160412017300752.
10. Mnif W, Hassine AIH, Bouaziz A, Bartegi A, Thomas O, Roig B. Effect of endocrine disruptor pesticides: a review. Int J Environ Res Public Health. 2011 Jun;8(6):2265–303.
11. Mamane A, Baldi I, Tessier J-F, Raherison C, Bouvier G. Occupational exposure to pesticides and respiratory health. Eur Respir Rev. 2015;24(136):306–19.
12. Yan D, Zhang Y, Liu L, Yan H. Pesticide exposure and risk of Alzheimer's disease: a systematic review and meta-analysis. Sci Rep. 2016 Sep;6:32222.
13. Harteloh P, de Bruin K, Kardaun J. The reliability of cause-of-death coding in The Netherlands. Eur J Epidemiol. 2010 Aug;25(8):531–8.
14. Nationaal Georegister (NGR). Dutch cadastral key-registry of buildings and addresses (BAG) [Internet]. 2016 [cited 2016 Jun 11]. Available from: https://data.overheid.nl/data/dataset/57437-bag-verblijfsobjecten.
15. Hazeu GW, Bregt AK, de Wit AJW, Clevers JGPW. A Dutch multi-date land use database: Identification of real and methodological changes. Int J Appl Earth Obs Geoinf [Internet]. 2011;13(4):682–9. Available from: http://www.sciencedirect.com/science/article/pii/S0303243411000560.
16. Statistics Netherlands (CBS). Landbouwtelling (Agricultural census statistics) 1980 – 2000 [Internet]. 2000 [cited 2018 Aug 24]. Available from: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/7316SLLB/table?ts=1536750050999.

17. Statistics Netherlands (CBS). Landbouwtelling (Agricultural census statistics) 2000–2017 [Internet]. 2018 [cited 2018 Aug 24]. Available from: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/80781NED/table?fromstatweb.
18. Brouwer M, Kromhout H, Vermeulen R, Duyzer J, Kramer H, Hazeu G, et al. Assessment of residential environmental exposure to pesticides from agricultural fields in the Netherlands. J Expo Sci Environ Epidemiol [Internet]. 2017 Mar 22;28:173. Available from: http://dx.doi.org/10.1038/jes.2017.3.
19. Statistics Netherlands (CBS). StatLine: Gebruik gewasbeschermingsmiddelen in de landbouw; gewas en toepassing '95-'12 [Internet]. Statistics Netherlands (CBS). 2012 [cited 2018 Aug 9]. Available from: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/82886NED/table?ts=1536749058879.
20. Gooijer YM, Hoftijser GW, Lageschaar LCC, Figueiredo DM, Huss A, Krop EJM, et al. Research on exposure of residents to pesticides in the Netherlands [Internet]. 2019. Available from: https://www.bestrijdingsmiddelen-omwonenden.nl/english.
21. Knol FA. Van hoog naar laag, van laag naar hoog: de sociaal-ruimtelijke ontwikkeling van wijken tussen 1971–1995 [in Dutch]. The Hague: Sociaal en Cultureel Planbureau; 1998.
22. Loyola E, Lebedeva E, Graen L, Mauer-Stender K, Fedkina N, Rakovac I. European tobacco use: Trends report 2019. 2019.
23. Fischer PH, Marra M, Ameling CB, Velders GJM, Hoogerbrugge R, de Vries W, et al. Particulate air pollution from different sources and mortality in 7.5 million adults – The Dutch Environmental Longitudinal Study (DUELS). Sci Total Environ. 2020 Feb;705:135778.
24. Smit LAM, Hooiveld M, van der Sman-de Beer F, Opstal-van Winden AWJ, Beekhuizen J, Wouters IM, et al. Air pollution from livestock farms, and asthma, allergic rhinitis and COPD among neighbouring residents. Occup Environ Med [Internet]. 2014 Feb 1;71(2):134 LP – 140. Available from: http://oem.bmj.com/content/71/2/134.abstract.
25. Slottje P, Yzermans CJ, Korevaar JC, Hooiveld M, Vermeulen RCH. The population-based Occupational and Environmental Health Prospective Cohort Study (AMIGO) in The Netherlands. BMJ Open. 2014 Nov;4(11):e005858.
26. Ratanachina J, De Matteis S, Cullinan P, Burney P. Pesticide exposure and lung function: a systematic review and meta-analysis. Occup Med (Chic Ill) [Internet]. 2020 Mar 12;70(1):14–23. Available from: https://doi.org/10.1093/occmed/kqz161.
27. HOPPIN JA, UMBACH DM, LONDON SJ, ALAVANJA MCR, SANDLER DP. Chemical Predictors of Wheeze among Farmer Pesticide Applicators in the Agricultural Health Study. Am J Respir Crit Care Med [Internet]. 2002 Mar 1;165(5):683–9. Available from: https://doi.org/10.1164/ajrccm.165.5.2106074.
28. Mamane A, Raherison C, Tessier J-F, Baldi I, Bouvier G. Environmental exposure to pesticides and respiratory health. Eur Respir Rev [Internet]. 2015 Sep 1;24(137):462 LP – 473. Available from: http://err.ersjournals.com/content/24/137/462.abstract.

SUPPLEMENTAL MATERIAL

Contents

*https://doi.org/10.1016/j.scitotenv.2022.152932

S1 Extended description of Materials and Methods

Land use variables

We computed the area, in hectares, of specific crops in the vicinity of all Dutch residences (n = 8,006,906), and linked these data to the DUELS administrative cohort. We used the Landelijk Grondgebruik Nederland (LGN) raster maps, with a resolution of 25 x 25 m, to obtain the type of crop cultivation registered per raster cell during specific years. These maps are available every three to five years, and are constructed using satellite imagery complemented with data from geographic information systems (GIS) that were available at the time such as topographic maps, aerial photography, and vector spatial databases. One LGN map uses several satellite images obtained from sections of the Netherlands taken in over a couple of years, meaning that one LGN map contains data from multiple years. For example the map LGN5 was built based on satellite imagery from the West of the Netherlands obtained in 2003 and from the East of the country in 2004.[1,2] In this study we used the maps from 1995/1997 (LGN3+), 1999/2000 (LGN4) and 2003/2004 (LGN5), from which we selected seven crop groups: maize, grains, potatoes, beets, fruit, flower bulbs and a group 'other crops' (excluding grass). An additional crop, 'all crops', was defined by combining all the previous crops together. For each raster cell on an LGN map, we computed the area covered by a specific crop using a moving average over a block of 2, 4, 10 and 20 cells north, south, east and west to the cell (Figure S1.1). This created squares with a surface area of 1.56 ha, 5.06 ha, 27.56 ha and 105.06 ha, respectively, that we refer to as the 50, 100, 250 and

500m buffers throughout the paper. We considered that the first two buffer distances were more likely to capture potentially higher exposure due to spray drift of pesticide droplets. This drift is highest within the first meters of an application and has an exponential decrease in concentration with distance, but can be detected up to 100 meters away from a field edge (depending on application technique and meteorological conditions).[3,4] The remaining two buffers are considered to capture secondary emission processes such as volatilization. This crop information was assigned to all residences in the Netherlands using the Basisregistraties Adressen en Gebouwen (BAG), a high-resolution cadastral dataset containing all buildings and addresses in The Netherlands.[5] We used the BAG data from February 2016 for the geolocation of the residences (as coordinates, which generally pertain to points within the building footprint). Residences closer than 500 m to the border between The Netherlands and Germany and Belgium were not included, as no information about land use across national borders was available (n= 39,017). Similarly to a previous study conducted in the Netherlands, assumptions on potential crop rotation were made for the periods in between the LGN maps. This is described in more detail in Brouwer et al.[6] In summary, we considered that fruit crops are a stable crop over the years and that this also applies to the location of flower bulb and maize crop fields while potato, grains and beet crops have regular annual rotation schemes. We computed the average area of the "stable crops" for the years in between LGN maps and summed the area of the "rotation crops" together. We then used the publicly available yearly Dutch agricultural census data provided by CBS from 1995 to 2003.[7,8] This data set provides information on the total area of specific crops for 66 agricultural regions that are relatively homogenous regarding the type of soil and agricultural land use and was used to calculate the proportion of potatoes, cereals and beet crops per region for each year. We summed the area of potatoes, cereals and beets form the calculated buffers into a "rotation crops" variable for the available LGN years and averaged these areas for the years in between. Then, by applying the calculated proportion of crops per year, we obtained the "probable" area of cultivated potatoes, cereals and beets for the unavailable LGN years. Finally, we linked the computed areas of specific crops within the four buffers to the residences in the Dutch Environmental Longitudinal Study (DUELS) cohort, thus assigning this pesticide exposure proxy to each individual in the cohort. By averaging the areas of crops across the exposure period of this study (1995 to 2004), we obtained four land use variables per crop that we used as the main exposure metrics in the analyses. In short, the land use variables reflect the average area (ha) of (specific) crops cultivated within 50, 100, 250 and 500m buffers around the cohort members' residences for the period 1995 to 2003.

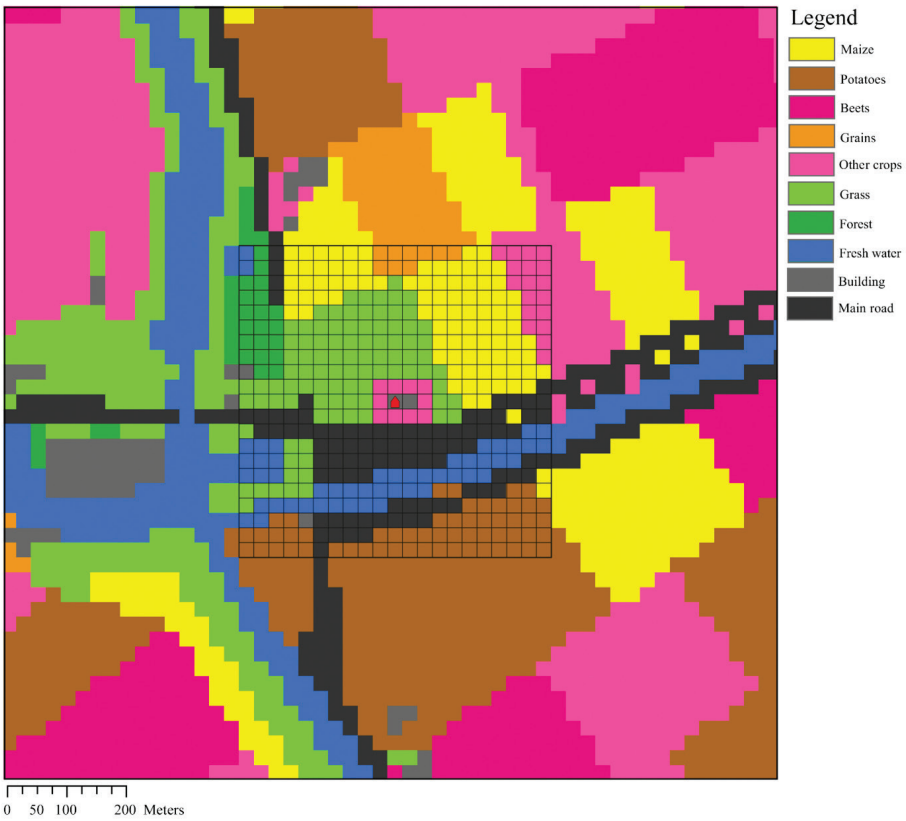


Figure S1.1: Detail from the Landelijk Grondgebruik Nederland raster map from 1999/2000 (LGN4) and an example of a 250m squared buffer around a house (red).

Pesticide use variables

We computed the amounts of insecticides, herbicides, fungicides and the total amount of these three pesticide classes that were used within the buffers we previously defined. For this we used the average amount of specific pesticides used per hectare of crop (kg/ha) as self-reported in national farmer surveys from 1995, 1998, 2000 and 2004.[9] These data are collected every three to four years using a survey covering a representative sample of Dutch farmers. Linkage between pesticide use data from the farmers' surveys (FS) and crop data from LGN maps was done for individual years and took into account the amounts of pesticides used on each type of crop (Figure S1.2). By multiplying the area by amount per hectare we obtained the average amount (kg) of pesticides used within the buffers. For the years in between available surveys, we

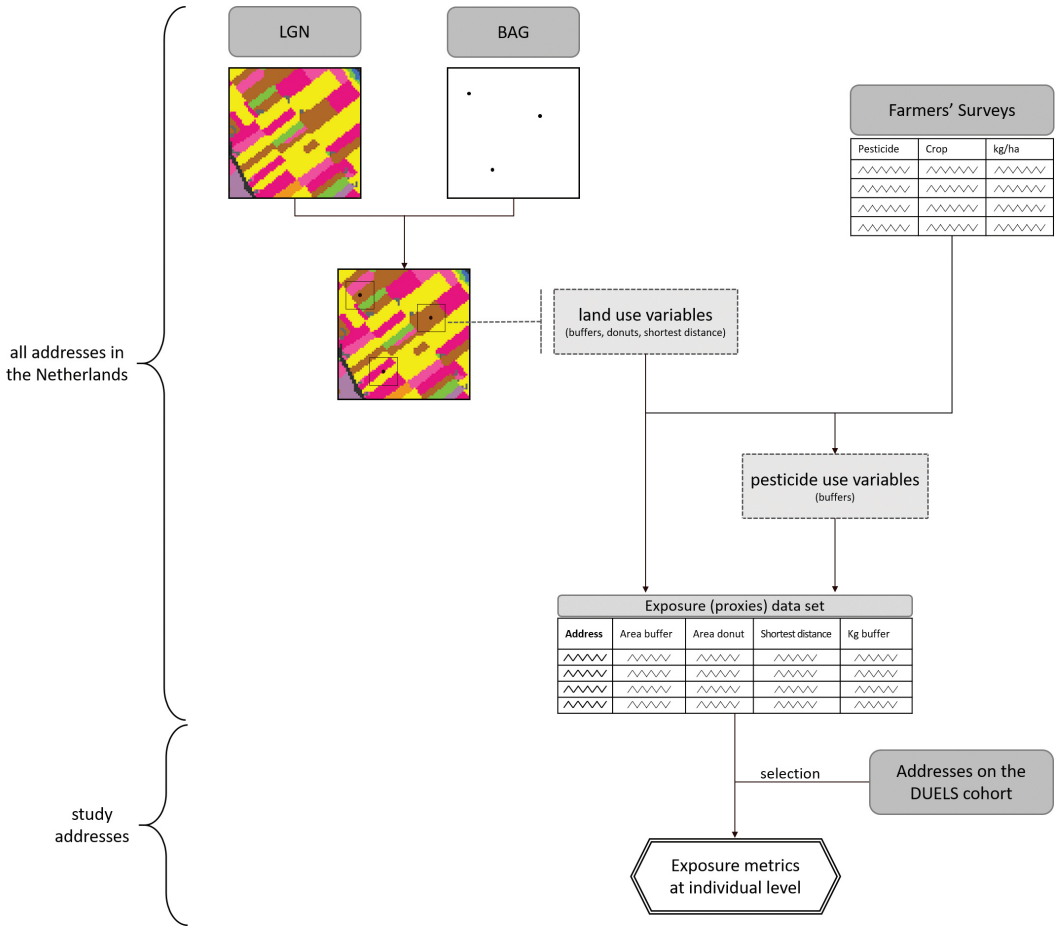
averaged the kg/ha between two surveys. Then, by summing the all the amounts obtained for all crop types present within the buffers we obtained the total amount of insecticides, herbicides and fungicides that were used within the buffers around each residence. The sum of these amounts resulted in the total amount of pesticides used. Finally, we averaged these amounts across the exposure period (1995 to 2004), obtaining four pesticide use variables per pesticide class within 50, 100, 250 and 500m buffers around the cohort members' residences for the period 1995 to 2003.

Data		Study Year	Computation of	
LGN data set	Farmers' Survey (FS)		Land use (LU) variables (per crop type)	Pesticide use (PU) variables (per pesticide class)
LGN3+ (1995/1997)	FS_{95}	1995	$LU_{95} = Area_{LGN3+}$	$PU_{95} = \sum LU_{95}^{crop} * FS_{95}^{crop}$
		1996	$LU_{96} = Area_{LGN3+}$	$PU_{96} = \sum LU_{96}^{crop} * mean(FS_{95}^{crop} * FS_{98}^{crop})$
		1997	$LU_{97} = Area_{LGN3+}$	$PU_{97} = \sum LU_{97}^{crop} * mean(FS_{95}^{crop} * FS_{98}^{crop})$
	FS_{98}	1998	$LU_{98} = mean(Area_{LGN3+}, Area_{LGN4})$	$PU_{98} = \sum LU_{98}^{crop} * FS_{98}^{crop}$
LGN4 (1999/2000)		1999	$LU_{99} = Area_{LGN4}$	$PU_{99} = \sum LU_{99}^{crop} * mean(FS_{98}^{crop} * FS_{00}^{crop})$
	FS_{00}	2000	$LU_{00} = Area_{LGN4}$	$PU_{00} = \sum LU_{00}^{crop} * FS_{00}^{crop}$
		2001	$LU_{01} = mean(Area_{LGN4}, Area_{LGN5})$	$PU_{01} = \sum LU_{01}^{crop} * mean(FS_{00}^{crop} * FS_{04}^{crop})$
		2002	$LU_{02} = mean(Area_{LGN4}, Area_{LGN5})$	$PU_{02} = \sum LU_{02}^{crop} * mean(FS_{00}^{crop} * FS_{04}^{crop})$
LGN5 (2003/2004)		2003	$LU_{03} = Area_{LGN5}$	$PU_{03} = \sum LU_{03}^{crop} * mean(FS_{00}^{crop} * FS_{04}^{crop})$
	FS_{04}		⬇	⬇
			$\overline{LU}_{95-03} = \frac{\sum_{1995}^{2003} LU}{9}$	$\overline{PU}_{95-03} = \frac{\sum_{1995}^{2003} PU}{9}$

Computations for the land use variables were performed for each buffer size (50, 100, 250 and 500m) and crop group (maize, grains, potatoes, beets, fruit, flower bulbs, and 'other crops'). Summing these crops resulted in the group 'all crops'. Computations for the pesticide use variables were performed for each buffer size and pesticides class (insecticides, herbicides, and fungicides). Combining these classes together resulted in the "Total pesticides" variable. LGN = Landelijk Grondgebruik Nederland.

Figure S1.2: Data and formulas used for the computation of the land use and pesticide use variables.

Figure S1.3 gives an overview of all data used in this study and how they were integrated to generate exposure proxies at the individual level. In summary, by overlaying the BAG points (addresses) over the LGN map, we were able to extract the area (buffers) of crop(s) around each residence of the Netherlands. This overlay also allowed the computation of the shortest distance variable, while the donut variables were computed by subtracting the area of a donut by the area of the smaller adjacent buffer. By multiplying the buffer areas by the amount of pesticides reported to be used per hectare of a specific crop (farmers' surveys) we obtained the amount (kg) of pesticides used within the buffers around the residences. Finally, we selected the addresses from the DUELS cohort hence obtaining land use and pesticide use variables for each cohort member.



LGN = Landelijk Grondgebruik Nederland, BAG = Basisregistraties Adressen en Gebouwen, DUELS = Dutch Environmental Longitudinal Study

Figure S1.3: Overview of the data used in the study.

In this study, we used the crop groups provided in the LGN maps, except for grass and greenhouse crops. A definition of these crops is provided in Table S1.1. Pesticide usage data from farmers' surveys pertained more specific crop groups. Table S1.1 also shows how we corresponded the pesticide usage information to the LGN crop groups. We used the average of the pesticide usage (kg/ha) of the specific crops to obtain the average pesticide usage in a LGN crop croup.

Table S1.1: Correspondence between LGN crop groups and pesticide usage information from farmers' surveys

Crop name	LGN definition of crop	Pesticide usage information from farmer's surveys (CBS)	
MAIZE	Agricultural plots with maize crop	Maize silage	
POTATOES	Agricultural plots with potatoes crops with no distinction between seed potatoes, potatoes for consumption and potatoes for starch	Seed potatoes Potatoes for consumption	Potatoes for starch
BEETS	Agricultural plots with sugar beets and fodder beets crops; excluding red beets which falls in the 'other crops' group as an horticultural crop	Sugar beets	
GRAINS	Agricultural plots with any grain crop: wheat, barley, oats, rye, etc., with no distinction made between summer grains or winter grains	Winter wheat Summer barley	Summer wheat
FRUIT	Orchards: tall fruit trees without distinction according to the type of fruit	Apples Pears	
FLOWER BULBS	Plots with flower bulbs, with no distinction between the type of flower bulb, nor between spring or autumn bulbs.	Hyacinths Tulips Daffodils	Irises Gladioles Lilies
OTHER CROPS	Agricultural plots with crops not falling within the preceding groups: horticultural crops, cabbages crops, hemp, rapeseed, etc. Excluding grass and greenhouse crops.	Flowers in open field Chicory Field beans Rapeseed Flax Cauliflower Green peas Kidney beans Grass seed Onions from seeds Onions from bulb Strawberries Asparagus Leak	Salsify Lettuce Head cabbages Brussel sprouts Green beans Carrots, first year Carrots, second year Chicory root Rose bushes Ornamental conifers Perenniel plants Forest and hedge plants Avenue and park trees Fruit trees (for garden centers)

S2 Descriptive data

Table S2.1: Descriptive statistics of the land use buffers for the exposed population within each buffer and per crop.

Crop	Buffer size	Exposed population within buffer		
		N (%)	Median (IQR)	Min – Max
MAIZE	500m	204 197 (6.46%)	0.13 (0.22)	0.014 – 1.498
	1000m	396 609 (12.55%)	0.29 (0.57)	0.014 – 4.876
	1500m	1 053 692 (33.34%)	0.93 (2.1)	0.014 – 23.323
	2000m	1 952 844 (61.79%)	3.1 (6.85)	0.014 – 69.463
POTATOES	500m	197 714 (6.26%)	0.06 (0.13)	8e-04 – 1.976
	1000m	379756 (12.02%)	0.12 (0.35)	5e-04 – 6.1059
	1500m	959895 (30.37%)	0.41 (1.15)	5e-04 – 27.0463
	2000m	1 763090 (55.79%)	1.21 (3.7)	5e-04 – 81.2856
BEETS	500m	197 712 (6.26%)	0.04 (0.09)	1e-04 – 1.1847
	1000m	379 749 (12.02%)	0.09 (0.25)	1e-04 – 3.6255
	1500m	959 858 (30.37%)	0.31 (0.9)	1e-04 – 18.5047
	2000m	1 762 973 (55.79%)	0.94 (3.07)	1e-04 – 56.7433
GRAINS	500m	197 714 (6.26%)	0.07 (0.15)	0.003 – 2.244
	1000m	379 756 (12.02%)	0.16 (0.41)	0.003 – 7.442
	1500m	959 895 (30.37%)	0.57 (1.57)	0.003 – 38.484
	2000m	1 763 090 (55.79%)	1.83 (5.27)	0.003 – 111.182
FRUITS	500m	218 200 (6.9%)	0.11 (0.18)	0.014 – 1.512
	1000m	427 945 (13.54%)	0.22 (0.44)	0.014 – 4.935
	1500m	1 106 089 (35%)	0.68 (1.53)	0.014 – 23.73
	2000m	2 004 662 (63.43%)	2.01 (5.03)	0.014 – 86.221
FLOWER BULBS	500m	47 581 (1.51%)	0.13 (0.24)	0.014 – 1.498
	1000m	104 218 (3.3%)	0.25 (0.58)	0.014 – 4.873
	1500m	300 520 (9.51%)	0.63 (1.87)	0.014 – 25.709
	2000m	645 178 (20.42%)	1.21 (4.55)	0.014 – 84.472
OTHER CROPS	500m	70 694 (2.24%)	0.56 (1.55)	0.014 – 13.48
	1000m	148 660 (4.7%)	1.12 (3.59)	0.014 – 43.85
	1500m	409 670 (12.96%)	3.19 (11.77)	0.014 – 231.333
	2000m	848 690 (26.86%)	6.46 (29)	0.014 – 760.278
ALL CROPS	500m	457 819 (14.49%)	0.24 (0.51)	0.014 – 14.978
	1000m	814 087 (25.76%)	0.62 (1.49)	0.014 – 48.723
	1500m	1 712 419 (54.19%)	2.72 (6.9)	0.014 – 257.042
	2000m	2 560 479 (81.02%)	11.8 (28.1)	0.014 – 852.857

Data are in n (%). IQR = Interquartile range, Min = Minimum, Max = Maximum

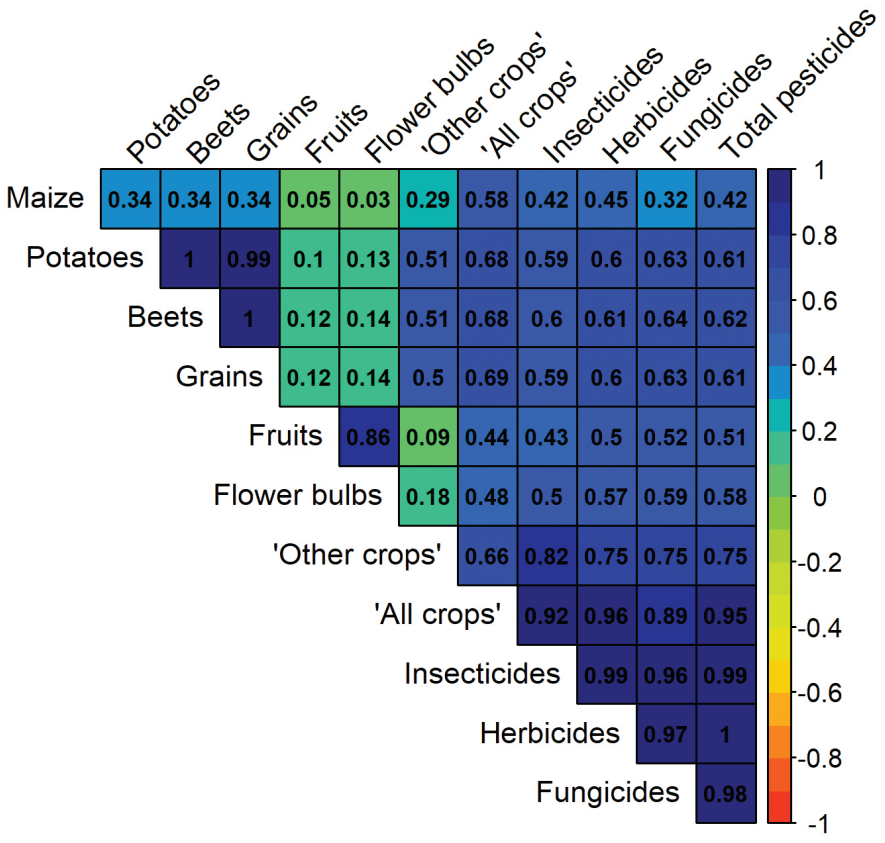


Figure S2.1: Spearman correlation matrix area of specific crop and amount of pesticides within buffers of 250m around the residences of the exposed population (i.e., individuals with least one crop within 500m of their residence).

Table S2.2: Number of deaths from any cause and from thirteen (specific) causes: overall and per exposed to specific crops.

		All causes	All causes (excluding external causes)	All external causes	Malignant neoplasms	Stomach cancer	Colon, sigmoid, and rectum cancer	
OVERALL		353 730	342 271	11 459	121 146	4448	14 937	
MAIZE	500m	20 807	20 092	715	7070	296	879	
	1000m	39 924	38 581	1343	13 702	562	1698	
	1500m	106 295	102 731	3564	37 195	1414	4664	
	2000m	206 219	199 564	6655	71 647	2680	8864	
POTATOES	500m	19 268	18 615	653	6777	271	805	
	1000m	37 752	36 516	1236	13 220	515	1589	
	1500m	98 774	95 573	3201	34 442	1356	4150	
	2000m	18 8191	182 145	6046	65 341	2501	7911	
BEETS	500m	19 268	18 615	653	6777	271	805	
	1000m	37 750	36 514	1236	13 220	515	1589	
	1500m	98 768	95 568	3200	34 440	1356	4150	
	2000m	18 8179	182 134	6045	65 338	2501	7910	
GRAINS	500m	19 268	18 615	653	6777	271	805	
	1000m	37 752	36 516	1236	13 220	515	1589	
	1500m	98 774	95 573	3201	34 442	1356	4150	
	2000m	18 8191	182 145	6046	65 341	2501	7911	
FRUITS	500m	4844	4681	163	1723	54	207	
	1000m	10 768	10 412	356	3763	130	436	
	1500m	31 015	30 051	964	10 798	397	1256	
	2000m	68 530	66 361	2169	23 662	860	2795	
FLOWER BULBS	500m	6800	6555	245	2449	90	290	
	1000m	14 712	14 199	513	5193	195	625	
	1500m	41 391	40 044	1347	14 562	534	1754	
	2000m	89 057	86 160	2897	30 837	1148	3696	
OTHER CROPS	500m	20 451	19 717	734	7222	314	861	
	1000m	41 079	39 674	1405	14 542	579	1733	
	1500m	111 726	108 033	3693	39 482	1534	4806	
	2000m	213 869	206 936	6933	74 518	2796	9150	
ALL CROPS	500m	45 386	43 804	1582	15 753	628	1942	
	1000m	81 354	78 605	2749	28 330	1106	3488	
	1500m	176 061	170 286	5775	61 820	2314	7582	
	2000m	275 612	266 724	8888	95 561	3538	11 752	

	Liver cancer	Pancreas cancer	Lung cancer	Skin cancer	Breast cancer	Ovary cancer	Prostate cancer
	1754	7019	28 763	2078	9232	3004	7929
	89	416	1630	117	509	183	498
	163	799	3248	226	1005	348	897
	512	2181	8998	623	2800	917	2359
	1003	4138	17 241	1187	5440	1761	4562
	83	386	1570	109	503	181	502
	179	752	3089	212	998	361	941
	496	1949	8264	595	2596	878	2275
	914	3722	15 712	1122	4980	1617	4265
	83	386	1570	109	503	181	502
	179	752	3089	212	998	361	941
	496	1949	8264	595	2596	878	2275
	914	3722	15 712	1122	4980	1617	4265
	83	386	1570	109	503	181	502
	179	752	3089	212	998	361	941
	496	1949	8264	595	2596	878	2275
	914	3722	15712	1122	4980	1617	4265
	21	107	404	25	121	53	131
	52	233	906	59	260	99	278
	176	647	2673	167	769	286	723
	357	1405	5772	378	1740	603	1516
	34	143	543	39	173	69	187
	70	309	1222	88	365	133	385
	219	852	3540	251	1050	384	976
	455	1801	7415	527	2266	793	2025
	90	415	1721	121	553	178	499
	190	848	3449	253	1111	369	975
	555	2276	9559	678	2959	996	2580
	1078	4301	18 052	1298	5689	1838	4816
	195	907	3667	261	1153	402	1136
	368	1667	6675	496	2088	714	1956
	905	3577	14 861	1087	4672	1541	4016
	1370	5522	22 893	1673	7278	2355	6203

Table S2.3: Number of deaths from twelve specific causes: overall and per exposed to specific crops.

		Kidney cancer	Non-Hodgkin's lymphoma	Brain cancer	Leukaemia	Endocrine, nutritional, and metabolic	Parkinson's disease	
OVERALL		2992	8397	2748	3405	9440	2742	
MAIZE	500m	178	502	165	198	559	151	
	1000m	352	946	328	407	1064	309	
	1500m	928	2514	862	1062	2800	818	
	2000m	1830	4931	1692	2025	5448	1555	
POTATOES	500m	179	469	171	215	509	141	
	1000m	336	895	324	388	974	271	
	1500m	842	2334	794	972	2587	731	
	2000m	1629	4445	1494	1828	4949	1391	
BEETS	500m	179	469	171	215	509	141	
	1000m	336	895	324	388	974	271	
	1500m	842	2333	794	972	2587	731	
	2000m	1629	4444	1494	1828	4949	1391	
GRAINS	500m	179	469	171	215	509	141	
	1000m	336	895	324	388	974	271	
	1500m	842	2334	794	972	2587	731	
	2000m	1629	4445	1494	1828	4949	1391	
FRUITS	500m	44	119	28	59	118	33	
	1000m	97	264	72	124	255	73	
	1500m	266	786	217	311	772	219	
	2000m	576	1626	552	693	1728	485	
FLOWER BULBS	500m	58	187	48	78	165	51	
	1000m	133	362	104	158	345	107	
	1500m	364	1038	295	416	1016	307	
	2000m	762	2106	701	898	2201	648	
OTHER CROPS	500m	180	480	157	206	483	129	
	1000m	372	984	339	400	1003	296	
	1500m	965	2670	913	1076	2876	812	
	2000m	1828	5089	1685	2038	5620	1590	
ALL CROPS	500m	391	1092	357	448	1168	327	
	1000m	707	1922	661	813	2070	613	
	1500m	1530	4211	1435	1700	4616	1352	
	2000m	2380	6607	2216	2659	7306	2082	

	Alzheimer's disease	Circulatory system diseases	Ischemic heart diseases	Cerebrovascular diseases	Respiratory system diseases	Chronic lower respiratory diseases
	3405	107 817	31 059	23 656	34 266	16 366
	165	6413	1842	1390	2182	1051
	311	12 142	3511	2645	4054	1941
	883	32 286	9357	7058	10 452	5069
	1814	62 777	18 228	13 707	20 252	9742
	138	5913	1813	1234	1809	909
	273	11 573	3485	2432	3597	1739
	844	30 187	8851	6554	9361	4528
	1679	57 355	16 677	12 502	18 175	8786
	138	5913	1813	1234	1809	909
	273	11 572	3485	2431	3596	1739
	844	30 185	8850	6553	9360	4528
	1679	57 349	16 675	12 501	18 174	8786
	138	5913	1813	1234	1809	909
	273	11 573	3485	2432	3597	1739
	844	30 187	8851	6554	9361	4528
	1679	57 355	16 677	12 502	18 175	8786
	37	1483	441	300	449	206
	88	3336	995	682	1015	473
	264	9547	2827	2035	2917	1359
	603	21 028	6190	4511	6447	3014
	69	2052	619	432	638	303
	141	4489	1336	957	1385	636
	379	12 628	3763	2736	3831	1765
	848	27163	8004	5878	8357	3895
	178	6213	1881	1328	1941	977
	357	12 542	3690	2765	3859	1905
	986	33 879	9881	7465	10 564	5134
	2038	64 806	18 920	14 102	20 632	10 012
	368	13 883	4110	2969	4482	2185
	666	24 762	7231	5360	7913	3817
	1533	53 483	15 621	11 664	16 805	8105
	2554	83 890	24 352	18 348	26 655	12 813

Table S2.4: Number of deaths by specific chronic lower respiratory disease in the study population.

Cause of death (ICD-10)	
Chronic lower respiratory diseases	**16 366**
Bronchitis, not specified as acute or chronic (J40)	53 (0.3%)
Unspecified chronic bronchitis (J42)	49 (0.3%)
Emphysema (J42)	1540 (9.4%)
Other chronic obstructive pulmonary disease (J44)	14 523 (88.7%)
Asthma and Status asthmaticus (J45, J46)	131 (0.8%)
Bronchiectasis (J47)	70 (0.4%)

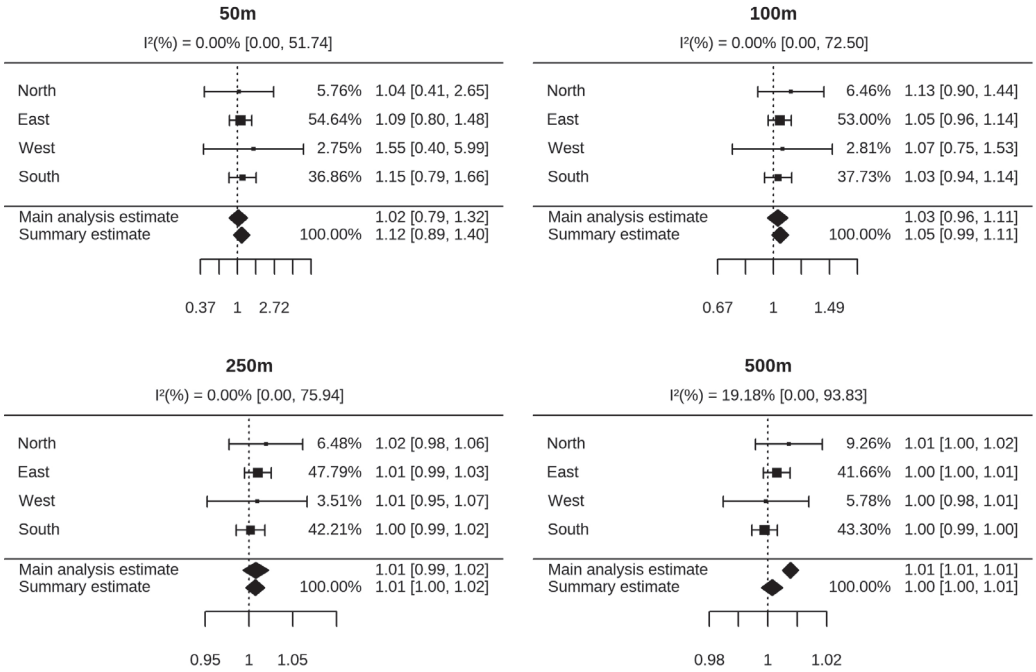
Data are in n (%).

S7 – Stratified analysis by NUTS 1 region

Because crop types are relatively region-specific in the Netherlands, we assessed whether the effect of living near maize on lower chronic respiratory diseases mortality seen for the whole country was driven by specific regions. We therefore conducted stratified analysis by the four major socio-economic regions (according to the Nomenclature des Unités Territoriales Statistiques, NUTS 1), followed by a random effects meta-analysis to assess heterogeneity (I^2 statistic) of regional estimates using the fully adjusted. We considered there was low heterogeneity when $I^2 < 75\%$. Figure S7.1 shows the results of the regional estimates from the stratified analyses (Hazard Ratio and 95% Confidence Interval), corresponding summary Hazard Ratio and heterogeneity from the meta-analysis.

S8 – Analysis controlling for the presence of livestock animals

We also hypothesized that the association between residential proximity to maize crops and chronic lower respiratory diseases mortality could be confounded by co-occurring agricultural activities, namely (intensive) livestock farming.
The yearly data sets from Geographic Information System of Agricultural Companies (GIAB) from 1999 to 2003, which are based on the Agricultural Census and curated by the Netherlands Enterprise Agency (RVO), contain information on the location of farming companies and the number and types of animals they have. We used these data sets to extract the number of cows, pigs, poultry, equines (horses and ponies), sheep, goats, rabbits and fur animals present within 500m of each residence. We then ran the full model for all buffers (50m, 100m, 250m and 500m), controlling for all the animal variables in the model. Table S8.1 shows the results of these analyses as well as the results of the main analysis for easier comparison of estimates.



Models were adjusted for sex, ethnicity, marital status, standardized household income, the presence of other crops within 500m of the residence (except for when 'all crops' was the exposure), neighborhood social economic position, urbanization degree at neighborhood level and proportion of low educated residents in the neighborhood.

Figure S7.1: Forest plots of the stratified analysis by the four major socio-economic regions in the Netherlands, including the results of the subsequent random-effects meta-analysis and I2 statistic, for the association between maize and chronic lower respiratory diseases mortality in the 50m, 100m, 250m and 500m buffers.

Table S8.1: Results of the analyses controlling for the number of livestock animals within 500m.

Buffer size	Main analysis		Analysis controlling for presence of livestock animals	
	HR [95% CI]	P value	HR [95% CI]	P value
50m	1.02 [0.79, 1.32]	0.868	1.30 [1.04, 1.63]	0.020
100m	1.03 [0.96, 1.11]	0.412	1.11 [1.04, 1.17]	0.001
250m	1.01 [0.99, 1.02]	0.319	1.02 [1.01, 1.03]	5.20E-05
500m	1.01 [1.01, 1.01]	1.95E-08	1.01 [1.00, 1.01]	7.40E-07

S9 – Analysis using the AMIGO cohort

Meeting two of the a priori defined criteria for interpretation, we considered the association between maize and lower chronic respiratory diseases to be a noteworthy observation, supported by consistent results of the 500m buffer analysis. Since we had access to the Environmental and Health Prospective Cohort Study (AMIGO) cohort, where information on self-reported chronic obstructive respiratory disease (COPD) was available, and were able to link it to information on land use, we replicated the analyses for this associations as described below.

The ongoing AMIGO study is population-based cohort of 14 829 adults (>30 years old) that were recruited in 2011 and 2012 via the Dutch network of general practitioners.[10] Cohort members filled in online questionnaires at baseline about occupational, environmental and lifestyle determinants, residential history and self-reported health (including symptoms and diagnosed diseases by a doctor). In our validation effort, we used data from these baseline questionnaires to perform a similar analysis to that of our study. We computed area of maize within buffers of 250m and 500m around the residences using the Basisregistratie Gewaspercelen (BRP) polygon agricultural land use maps from 2009, 2010 and 2011 and averaged these areas to obtain one variable per buffer size.[11] We investigated the association between the area of maize within the two buffers and self-reported COPD and built two models:

(1) A model controlling for the same possible confounders that were taken into account in the full model of the main analysis: age, sex, origin, civil status, SES, urbanization degree and presence of other crops within 500m.

(2) A model additionally controlling for additional potential confounders: body mass index (BMI), having a paid job, education, smoking, alcohol use and estimated particulate matter with ≤10µm in diameter (PM_{10}) and particulate matter with ≤2.5µm in diameter ($PM_{2.5}$).

After excluding 25 cohort members that resided within 500m of the borders with Germany and Belgium, we imputed the missing values for the outcome and the confounders using the Multivariate Imputation by Chained Equations (MICE) method. Table S9.1 shows the demographic characteristics for the population before imputation (N=14 804). We imputed five data sets with 20 iterations each. Age at diagnosis of COPD could only be answered when the participant previously answered having been diagnosed with COPD, resulting in "bona fide" missing values for this variable. Because there is no method to impute such bona fide missing values in R's MICE package as yet, after the imputation we further used the subset of diagnosed participants and ran another round of imputation (five datasets, 50 iterations), calculated the mean age of the imputed ages and assigned it as the imputed value. This was done for each of the five imputed data sets.

Table S9.1: Demographic characteristics of the Environmental and Health Prospective Cohort Study (AMIGO) cohort (N=14 804).

	Total population N = 14 804	Exposed to maize in 500m n = 3590
AGE		
Mean (SD)	50.7 (9.4)	50.8 (9.1)
Missing	0 (0%)	0 (0.0%)
SEX		
Male	6543 (44.2%)	1624 (45.2%)
Female	8261 (55.8%)	1966 (54.8%)
Missing	0 (0%)	0 (0%)
COUNTRY OF ORIGIN		
The Netherlands	14 104 (95.3%)	3470 (96.7%)
Other	700 (4.7%)	120 (3.3%)
Missing	0 (0%)	0 (0%)
BMI		
Mean (SD)	26.1 (4.4)	26.2 (4.4)
Missing	54 (0.4%)	13 (0.4%)
PM_{10}		
Mean (SD)	24.5 (0.9)	24.2 (0.6)
Missing	64 (0.4%)	16 (0.4%)
$PM_{2.5}$		
Mean (SD)	16.5 (0.7)	16.6 (0.7)
Missing	64 (0.4%)	16 (0.4%)
SES		
Mean (SD)	39.4 (6.9)	41.2 (5.3)
Missing	135 (0.9%)	70 (1.9%)
PAID JOB		
No	4761 (32.2%)	1141 (31.8%)
Yes	9660 (65.3%)	2328 (64.8%)
Missing	383 (2.6%)	121 (3.4%)
CIVIL STATUS		
Single	12 229 (82.6%)	3063 (85.3%)
Together	2551 (17.2%)	520 (14.5%)
Missing	24 (0.2%)	7 (0.2%)
EDUCATION LEVEL		
Low	439 (30.7%)	1145 (31.9%)
Medium	421 (31.2%)	1251 (34.8%)
High	504 (35.2%)	1072 (29.9%)
Missing	440 (3%)	122 (3.4%)

→

Table S9.1: Demographic characteristics of the Environmental and Health Prospective Cohort Study (AMIGO) cohort (N=14 804). (cont.)

	Total population N = 14 804	Exposed to maize in 500m n = 3590
SMOKING		
Never	6729 (45.5%)	1671 (46.5%)
Ever	5731 (38.7%)	1400 (39%)
Current	2321 (15.7%)	513 (14.3%)
Missing	23 (0.2%)	6 (0.2%)
ALCOHOL USE		
Never	847 (5.7%)	205 (5.7%)
Ever	1092 (7.4%)	275 (7.7%)
Current	12 837 (86.7%)	3102 (86.4%)
Missing	28 (0.2%)	8 (0.2%)
URBANIZATION DEGREE		
≥2500	1250 (8.4%)	3 (0.1%)
1500–2500	2979 (20.1%)	179 (5%)
1000–1500 addresses per km²	3096 (20.9%)	508 (13.2%)
500–1000 addresses per km²	3611 (24.4%)	914 (25.5%)
<500 addresses per km²	3809 (25.7%)	1975 (55%)
Missing	59 (0.4%)	11 (0.3%)
COPD DIAGNOSIS		
No	13 849 (93.5%)	3366 (93.8%)
Yes	553 (3.7%)	127 (3.5%)
Missing	402 (2.7%)	97 (2.7%)
AGE AT COPD DIAGNOSIS		
Mean (SD)	37.7 (18.4)	38.3 (18.4)
Missing	408 (2.8%)	97 (2.7%)

Data are in n (%) unless stated otherwise. SD = standard deviation, BMI = Body Mass Index, PM_{10} = particulate matter with ≤10µm in diameter, $PM_{2.5}$ = particulate matter with ≤2.5µm in diameter, SES = Socioeconomic Status, COPD = Chronic Obstructive Pulmonary Disease

We then selected participants living in neighborhoods with less than 1500 addresses per km2, that did not change addresses three years prior to baseline, and whose COPD diagnose occurred after 2009. These selections resulted in data sets with differing number of participants because urbanization degree and age at COPD diagnosis were imputed variables; on average, the obtained data sets had 9560 participants. Table S9.2 shows a cross table of cases of COPD and having maize within 250m and in 500m of the residence, per data set imputed.

Table S9.2: Cross table of cases of chronic obstructive respiratory disease (COPD) and exposure to maize within 250m and 500m of the residence.

Imputed data set	No. of participants	COPD diagnosis	Maize within 250m of the residence		Maize within 500m of the residence	
			Yes	No	Yes	No
#1	9555	Yes	21	35	35	21
		No	3454	6045	6310	3189
#2	9557	Yes	20	35	34	21
		No	3454	6048	6313	3189
#3	9560	Yes	20	35	34	21
		No	3459	6046	6317	3188
#4	9564	Yes	20	35	34	21
		No	3461	6048	6320	3189
#5	9566	Yes	20	35	34	21
		No	3458	6053	6318	3193

Data are in n. No. = number of, COPD = Chronic Obstructive Pulmonary Disease

Table S9.3 shows the odds ratio (OR) and 95% confidence interval for the association between having maize within 250m and in 500m of the residence and COPD when (1) controlling for the same potential confounders as used in the full model of the main analysis with the DUELS cohort and (2) when adjusting for additional potential confounders. The OR is >1 in 250m when we control for the same confounders as in the analysis with the DUELS cohort and remains so when we adjust for all the confounders available in AMIGO.

Table S9.3: Odds ratio and 95% Confidence interval for the association between the maize and chronic obstructive respiratory disease (COPD).

Buffer size	OR [95% CI]			
	Adjusted for age and sex	Adjusted for potential confounders considered in the main analysis[a]	Adjusted for potential confounders considered in the main analysis and lifestyle factors available in the AMIGO data set[b]	Adjusted for all possible confounders available in the AMIGO data set[c]
250m	1.00 [0.84, 1.20]	1.08 [0.88, 1.32]	1.09 [0.89, 1.33]	1.09 [0.89, 1.33]
500m	0·99 [0·94, 1·05]	0·99 [0·93, 1·06]	1·00 [0·93, 1·07]	0·99 [0·93, 1·07]

OR = Odds Ratio, CI = confidence interval, AMIGO = Environmental and Health Prospective Cohort Study

a Adjusted for age, sex, origin, civil status, socioeconomic position, urbanization degree, and presence of other crops within 500m

b Adjusted for age, sex, origin, civil status, socioeconomic position, urbanization degree, presence of other crops within 500m, BMI, smoking, and alcohol use

c Adjusted for age, sex, origin, civil status, socioeconomic position, urbanization degree, presence of other crops within 500m, BMI, smoking, alcohol use, having a paid job, education, and estimated PM_{10} and $PM_{2.5}$

REFERENCES

1. Hazeu GW, Bregt AK, de Wit AJW, Clevers JGPW. A Dutch multi-date land use database: Identification of real and methodological changes. Int J Appl Earth Obs Geoinf 2011; 13(4): 682–9. DOI: 10.1016/j.jag.2011.04.004
2. Hazeu, GW. Landelijk Grondgebruiksbestand Nederland (LGN5); vervaardiging, nauwkeurigheid en gebruik. Wageningen: Alterra – Centrum Geo-informatie; 2005. 92 p. Alterra-rapport : 1213. Available from: https://edepot.wur.nl/17654
3. Wolters A, Linnemann V, van de Zande JC, Vereecken H. Field experiment on spray drift: Deposition and airborne drift during application to a winter wheat crop. Sci Total Environ 2008; 405(1–3): 269–77.
4. Rautmann D, Streloke M, Winkler R. New basic drift values in the authorization procedure for plant protection products. Workshop on Risk Assessment and Risk Mitigation Measures (WORMM), vol. 13. 2001. 3–141 p.
5. Nationaal Georegister (NGR). Dutch cadastral key-registry of buildings and addresses (BAG). 2016 [cited 2016 Jun 11]. Available from: https://data.overheid.nl/data/dataset/57437-bag-verblijfsobjecten
6. Brouwer M, Kromhout H, Vermeulen R, et al. Assessment of residential environmental exposure to pesticides from agricultural fields in the Netherlands. J Expo Sci Environ Epidemiol 2017; 28(2): 173–181.
7. Statistics Netherlands (CBS). Landbouwtelling (Agricultural census statistics) 1980 – 2000. 2000 [cited 2018 Aug 24]. Available from: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/7316SLLB/table?ts=1536750050999
8. Statistics Netherlands (CBS). Landbouwtelling (Agricultural census statistics) 2000–2017. 2018 [cited 2018 Aug 24]. Available from: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/80781NED/table?fromstatweb
9. Statistics Netherlands (CBS). StatLine: Gebruik gewasbeschermingsmiddelen in de landbouw; gewas en toepassing '95-'12. Statistics Netherlands (CBS). 2012 [cited 2018 Aug 9]. Available from: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/82886NED/table?ts=1536749058879
10. Slottje P, Yzermans CJ, Korevaar JC, Hooiveld M, Vermeulen RCH. The population-based Occupational and Environmental Health Prospective Cohort Study (AMIGO) in the Netherlands. BMJ Open 2014; 4: e005858. DOI: 10.1136/bmjopen-2014–005858
11. Dutch Ministry of Interior and Kingdom Relations. Dataportaal van de Nederlandse overheid — BRP Gewaspercelen. 2016 [cited 2016 Jun 24]. Available from: https://data.overheid.nl/data/dataset/basisregistratie-gewaspercelen-brp

CHAPTER 4

SELF-REPORTED PSYCHOLOGICAL DISTRESS AND SELF-PERCEIVED HEALTH IN RESIDENTS LIVING NEAR PESTICIDE-TREATED AGRICULTURAL LAND: A CROSS-SECTIONAL STUDY IN THE NETHERLANDS

Mariana Simões, Anke Huss, Nicole Janssen, Roel Vermeulen

Occup Environ Med. 2022 Feb;79(2):127–133. doi: 10.1136/oemed-2021–107544

ABSTRACT

Objectives: There is rising concern regarding possible health effects from exposure to pesticides in residents living near agricultural land. Some studies indicated increased risks of reporting symptoms of anxiety and depression among agricultural workers but less is known about mental and perceived health of rural residents. We aimed to study possible associations between self-reported psychological distress (SPD) and self-perceived health (SPH) in residents near pesticide-treated agricultural land.

Methods: Using the Public Health Monitor national survey from 2012, we selected 216 932 participants who lived in rural and semi-urban areas of the Netherlands and changed addresses at most once in the period 2009–2012. Psychological distress was assessed via the Kessler Psychological Distress scale (K10) and participants were asked to assess their own health. We estimated the area of specific crop groups cultivated within buffers of 50m, 100m, 250m and 500m around each individual's residence for the period 2009–2012. Association between these exposure proxies and the outcomes was investigated using logistic regression, adjusting for individual, lifestyle and area-level confounders.

Results: Overall, results showed statistically non-significant Odd ratios (OR) across all buffer sizes for both SPD and SPH, except for the association between SPH and 'all crops' (total area of all considered crop groups) with OR [95% Confidence interval] ranging from 0.77 [0.63, 0.93] in 50m to 1.00 [1.00, 1.00] in 500m. We observed that most OR were below unity for SPH.

Conclusions: This study provides no evidence that residential proximity to pesticide treated-crops is associated with psychological distress or poorer perceived health.

Keywords: Psychological distress, perceived health, residential pesticide exposure, general population, cross-sectional study

INTRODUCTION

Psychological distress (PD), a mental health disorder usually characterized by depression and anxiety, has been shown to be a leading cause of disability and an important burden to society given its heavy impact on the quality of life, higher risk of premature death and absenteeism costs.[1–4] Some studies indicate increased risks of reporting symptoms of PD among agricultural workers.[5] Literature on residential pesticide exposure and PD is, nevertheless, still scarce.[3] Whereas some studies found indications of increased risks of depression among residents living near agricultural land,[6–8] others were unable to obtain similar results.[9,10] Differences in outcome and exposure assessment and in study design hamper comparability of results. The link between residential exposure to pesticides and PD remains therefore unclear.

Self-perceived health (SPH) is a more extensive measurement of health, constituting an important predictor of morbidity and mortality.[11,12] Encompassing both physical and mental health, it has been shown to be inversely associated with PD.[13,14] Concurrently, while poorer SPH has been reported to be associated with exposure to pesticides in farmers in one study,[15] research on this association among residents is strikingly lacking, given the growing public concern about the possible health effects of residential exposure to pesticides.

This study aims to contribute to the body of literature on the association between residential pesticide exposure and PD. Using a large nationwide survey of the Dutch adult population, we explored the associations between residential proximity to crops where pesticides are applied and self-reported psychological distress (SPD) and self-perceived health (SPH).

METHODS

Study population

The Public Health Monitor 2012 (Gezondheidsmonitor) is a national health survey conducted by the 28 regional Public Health Services, Statistics Netherlands (CBS) and the National Institute for Public Health and the Environment (RIVM) that includes information regarding perceived health and lifestyle from citizen aged ≥19 years. Elderly (≥65 years) were oversampled by design; response rates were 45–50%.

We excluded persons not registered in the Netherlands; who changed addresses more than once in the period 2009–2012; who lived in the most urbanized areas of The Netherlands (≥1500 addresses/km^2 at neighborhood level, since urban populations rarely

live in proximity to crops and differ in lifestyle and living environment factors compared to the more rural populations); and who lived within 1000m from the border (for which we were unable to compute exposure). We included a total of 216 932 participants.

Outcomes: Phycological distress and perceived health

Outcomes of interest were self-reported phycological distress and self-perceived (general) health. The first was assessed via the Kessler Psychological Distress scale (K10),[16] classifying participants into "well" and "low to severe risk of psychological distress", at a cut-off value of >19 of the K10 score. The K10 has been validated for the Dutch population.[17] For the latter, participants were also asked to assess their own health based on a simple question: "In general, would you say that your health is...". Participants could answer one of five options that we later dichotomized into "good to very good" and "moderate to very poor".

Exposure: crop area around residences

We used residential proximity to crops as proxy for agricultural pesticide exposure. First, we geocoded all residences using the Basisregistratie Adressen en Gebouwen (BAG)[18]. Second, for computational reasons we rasterized the annual land use polygon maps from 2009 to 2012 (Basisregistratie Gewaspercelen)[19] and computed area of specific crops around participants' residences, in hectares (ha), using a moving average. This resulted in squared buffers that roughly correspond to radii of 50m, 100m, 250m and 500m.We assessed 13 crop groups representing 88–89% of the Netherlands' open field cultivated land, excluding grassland[19]: maize, winter wheat, summer barley, summer wheat, other cereals, potatoes for consumption, potatoes for starch, seed potatoes, beets, ornamental plants and tree nurseries, vegetables, fruit and flower bulbs. Summing the area of all 13 crop groups, we created the group 'all crops'. Third, we averaged the areas across the exposure period (2009–2012) and obtained four land use buffers reflecting the average area (ha) of a specific crop cultivated within 50, 100, 250 and 500m for each participants residence. We considered that the first two buffer distances to capture direct spray drift of pesticide droplets. This drift is highest within the first few meters of application and has an exponential decrease in concentration with distance, but can be detected up to 100 meter away from a field edge (depending on application technique and meteorological conditions).[20,21] The highest environmental exposure to pesticides would therefore occur within these two buffers. Pesticides can however be detected at larger distances due to secondary emission processes such as volatilization.[21] The two larger

buffers are considered to capture this secondary drift. A recent Dutch exposure assessment study observed a high contrast in pesticide concentration in air and house dust between residences within 250m and beyond 500m from flower bulb crops and that gradients in concentrations within 250m distance from fields were weak.[22] We therefore explored odds ratio (OR) gradients across 50m, 100m, 250m and 500m buffers, using the area (in ha) of (specific) crop within a buffer and adjusting the model for the remaining area of that crop up to 500m (buffers and complementary donuts were thus used continuous variables in the analyses). Our referent ("unexposed") group consisted of participants with zero hectares of (specific) crop within 500m of their residences.

Statistical analyses

Imputation

The data set comprised of 147 886 (68%) complete cases, with missing values in both outcome and potential confounders variables (Table 1) in a non-monotone pattern. We used multiple imputation by chained equations (MICE) to impute missing values in 20 datasets with 10 iterations. For the imputation models, we considered all variables included in the statistical models for the analyses (outcomes, potential confounders and total crop area in 500m) and data on prescription of antidepressants, anxiolytics and hypnotic-sedative drugs in 2012, paid work, and urbanization degree, which were predictive of some of the imputed variables but were not included in the statistical models. Collinearity between predictors was measured by variance inflation factors (VIF) using a cut-off of 3. Binary variables were imputed using logistic regression and continuous and categorical variables were imputed using predictive mean matching.

Main analysis

We applied logistic regression, building four models with increasing covariate adjustment for each crop-specific land use buffer and outcome combination:

- *basic model*, adjusted for the area of the considered crop that remained until 500m ('complementary donut'), age and sex,
- *individual confounders model*, consisting of basic model and body mass index (BMI), country of origin, marital status, educational level, living with children, having a chronic condition and the presence (yes/no) of other crops within 500m of the participant's residence (except for when 'all crops' was the exposure),

- *lifestyle confounders model*, extending the individual model with physical activity,[22] alcohol status and smoking status,
- *full model*, adding neighborhood socioeconomic position[23] and NDVI (Normalised Difference Vegetation Index, a measure of green space)[24] within 500m of the participant's residence to lifestyle models (see Table 1 for confounder categories).

Sensitivity and additional analyses

We performed four sensitivity analyses on full models. First, we restricted analysis to participants living in a rural setting (<1000 addresses/km^2 at neighborhood level) to assess potential bias from a semi-urban environment. Second, since changes in address may be related to physical or mental health problems, we excluded participants that changed address during 2009–2012. Third, we restricted the analysis to complete cases to assess the impact of using imputed datasets. Forth, in order to exclude a possible influence of occupational exposure, we linked microdata on employment and self-employment available from CBS to identify and exclude people that worked for at least one year in agriculture in the period 2009–2012. Furthermore, because the epidemiology of psychological distress differs between women and men, we conducted a stratified analysis by sex.[25] Finally, since the elderly were oversampled in the survey we conducted an analysis stratified by age (<65 vs 65+ year olds).
For completeness we also conducted additional analyses using different exposure metrics. We calculated the area of cultivated crops around the residence as "donuts with holes" (<100m, 100–250m and 250–500m) that we used as continuous and as binary ("presence"/"absence" of crop) variables in the analyses. We also computed the average distances to nearest crop (categorized into <50m, 50–100m, 100–250m, 250-500m, ≥500m) in 2009–2012.
Statistical analyses were performed in R version 3.4.1 (2017–06–30).

RESULTS

Table 1 describes the study population's demographic characteristics. We included 216 932 participants (46% men, median age 61); 78 355–78 522 (36.12%-36.20%) participants had low to severe risk of PD and 56 615–56 688 (26.10%-26.13%) had moderate to very poor SPH, depending on the imputed data set. There were 21 148 (9.75%), 43 737 (20.16%), 106 122 (48.92%), and 168 088 (77.48%) people exposed to at least one type of treated-crop in the 50, 100, 250 and 500m buffer, respectively. We observed mostly low (Pearson correlation <0.39) correlations between the buffer areas

of crops, except for the moderate correlations between winter wheat, potatoes for consumption and beets (Pearson correlations ranging from 0.56 to 0.61), which are grown in a rotation scheme.

Table 2 shows the number of unexposed and exposed participants per buffer size and type of crop. For some of the less prevalent crops, such as potatoes for starch and other cereals, the number of exposed people in the smaller buffers was very low, resulting in wide confidence intervals of the estimates. This table also displays the Odds Ratios (OR) per increase in 1 hectare of area of (specific) crop and their 95% Confidence Intervals for the full models. We found no clear evidence of associations between presence of specific crops and SPD and SPH. We observed overall patterns of OR below unity for SPH, with increasing gradient of effect sizes from the smaller to the largest buffers. Nevertheless, none of these associations showed statistically significant results consistently among the four buffers. Solely the association between 'all crops' and SPH showed statistically significant results in the 50, 100 and 250m buffers, with OR ranging from 0.77 [0.63, 0.93] in the 50m buffer to 0.99 [0.98, 1.00] in the 250m buffer. Increasing covariate adjustment in the 50m and 100m buffer models resulted in effect estimates that were, in general, closer to unity and, in rare cases, change in direction of effect (Tables S1 and S2). Furthermore, higher levels of adjustment resulted in loss of statistical significance across all buffer models. Neither sensitivity (Tables S3 and S4) nor stratified (Tables S5 and S6) analyses showed material changes in effect estimates. Results using the donuts and distance exposure metrics, shown in Tables S7 and S8, did not change our interpretation of the findings.

DISCUSSION

We used a cross-sectional national survey to study the association between presence of crops near residences and self-reported psychological distress (SPD) and self-perceived health (SPH). Analyses did not indicate that living close to treated-crops was associated with increased risks of SPD or poorer SPH. In fact, we observed overall negative non-significant associations.

The use of the national Public Health Monitor survey allowed us to include a large population of over 200 000 participants, making it one of the largest studies on SPD and residential pesticide exposure. It also enabled for adjustment of a range of relevant lifestyle aspects, although possibly some bias might have been introduced due to self-reporting and average response rate of 47%. We were also able to adjust for the presence of green space in the living environment, which previously has been associ-

ated with better mental and physical health in the Public Health Monitor 2012.[24,26,27] Furthermore, although we did not have exposure data based on measurements, we were able to estimate proxies of exposure at individual level that represented the specific pesticide mixtures and farming methods used in these crops.

Although pesticide exposure was not measured, it was assessed at individual level based on the area of crops around residences. This has been shown previously to be suitable in estimating pesticide levels in residences located near crops,[28] but entail important assumptions and limitations that could have resulted in exposure misclassification. First, we did not consider participants' time-activity patterns, their presence at the residence during spraying events (which may influence exposure levels) or other relevant locations for exposure, such as the workplace. Of all, workplace is probably the most important alternative source of exposure and our sensitivity analysis excluding agricultural workers did not show substantial differences in effect estimates from the main analysis. Domestic use of pesticides or nutritional exposure were not considered but are unlikely to differ substantially within the short distances from treated agricultural fields considered in this study, abating residual confounding. We also did not account for wind speed and direction, which affect spread of pesticides applied in fields. In the Netherlands, prevailing wind is West to South West, but Eastern winds are generally associated with lower wind speed and therefore also important for stable lower spread of pesticides at short distances. In this study we used symmetric (squared) buffers around residences since the best non-symmetrical buffer is difficult to determine. Finally, since no information was available, we were unable to differentiate between conventional and organic crops, but the latter comprised only 1.8% of the total area of investigated crops the Netherlands in 2015.[29]

Because a growing season only lasts for a limited time per year, living at a specific address only for a short period of time increases uncertainty in the exposure estimates. We therefore included people that moved at most once in the period 2009–2012 to minimize uncertainty around exposure that arises from multiple address changes. This resulted in the exclusion of only a minor proportion of the study population (3.5% of the original number of participants in the Public Health Monitor 2012, Figure 1). Changes in address may be related to the investigated outcomes, introducing another source of bias in the study. Unfortunately, we were unable to determine the reasons for moving addresses and we recognize that people might move to residences both further away or nearer to more rural areas for health reasons. Still, a sensitivity analysis restricted to people that never moved addresses in 2009–2012 and a sensitivity analysis restricted to people living in rural areas (<1000 addresses per km^2) did not show major changes in estimates. Outcome misclassification could have been aggravated by the cross-sectional design of the study since we were unable to establish the temporality of onset of the outcomes, which might have occurred before exposure.

Oversampling of elderly could have left the study vulnerable to selection bias. Nevertheless, our stratified analyses did not indicate substantial differences in risk estimates between people under and above 65 years of age. Similarly, even though women have higher risks of psychological distress, no major differences were found in the OR obtained for women and men in a stratified analysis by sex.

Our findings show no associations between proximity to pesticide-treated crops and SPD. This is in line with results from longitudinal studies that found no association between cumulative exposure from pesticide usage among farmers' families and depression.[8–10] Nevertheless, exposure and outcome misclassification were important limitations in these studies, mainly because information was collected via self-report. In contrast, the same studies reported increased risks of depression when pesticide exposure was deemed high enough to induce poisoning. The link between exposure to poisoning inducing pesticide concentrations and increased risk of depression was also reported in a cross-sectional study among farmers and their wives.[7] Two other studies reported positive associations between depression and pesticide exposure as well. One suggested that residential proximity to organophosphate application sites was associated with progression of depressive symptoms in a cohort of Parkinson's disease patients.[30] The other, an ecologic study, reported higher rates of depression among agricultural workers living in areas with intense pesticide application when compared with city dwellers.[6]

We used SPH as an extension of SPD in this study, since SPH is an important component of mental health. SPH has been suggested to be a mediator of the relationship between physical and mental health and shown to be an important indicator of current health status and predictor of depression and mortality.[31,32] We observed that OR for SPH were more often below unity for study participants living close to treated-crops. One would expect an equal ratio between risk estimates above and below unity if presence of crops had no effect on these outcomes. It is unclear why this 'protective' effect was observed, but it may be an indication of uncontrolled bias. Furthermore, given the limitations described above regarding our exposure proxy, it may be possible that we were unable to detect weak to moderate signals. On the other hand, previous studies have also shown negative associations to environmental exposures among the Dutch rural population, namely lower (non-accidental) mortality rates and respiratory problems.[33,34] We are unable to provide data driven explanations for the (statistically non-significant) negative associations found. In the Netherlands, socioeconomic position distribution is relatively similar across all areas of urbanization degree. Tentative explanations could include exposure misclassification, uncontrolled bias and the fact that rural populations in the Netherlands may have a better quality of life in (better air quality, lower costs, less stress, perhaps

more physical activity). It remains, however, unclear why we see this trend in such a small spatial scale (500m), that is, among the rural to semi-urban population itself.
In conclusion, this study provides no evidence that residential proximity to pesticide treated-crops is associated with psychological distress or poorer perceived health. In fact, we observed an overall indication of lower risks of poorer self-perceived health. Exposure and outcome misclassification remain important limitations in studies assessing these associations and hamper interpretation of results, including this one.

ACKNOWLEDGMENTS

The Dutch Public Health Monitor 2012 (Gezondheidsmonitor) was conducted by 28 Public Health Services (GGD), Statistics Netherlands (CBS) and National Institute for Public Health and the Environment (RIVM). Statistical analyses were conducted within the remote access secured environment of CBS.

TABLES

Table 1: Population characteristics (before imputation).

	Study Population (n =216 932)
OUTCOMES	
Self-perceived health	
Good [n (%)]	158 458 (73.05%)
Moderate to bad [n (%)]	55 535 (25.60%)
NA [n (%)]	2939 (1.35%)
Risk of psychological distress	
None [n (%)]	134 976 (62.22%)
Low to high [n (%)]	75 534 (34.82%)
NA [n (%)]	6422 (2.96%)
INDIVIDUAL COVARIATES	
Sex	
Men [n (%)]	99 926 (46.06%)
Women [n (%)]	117 006 (53.94%)
Age[a]	
19–24 [n (%)]	9666 (4.46%)
25–29 [n (%)]	6423 (2.96%)
30–34 [n (%)]	8291 (3.82%)
35–39 [n (%)]	10 003 (4.61%)
40–44 [n (%)]	14 934 (6.88%)
45–49 [n (%)]	17 045 (7.86%)
50–54 [n (%)]	17 822 (8.22%)
55–59 [n (%)]	18 247 (8.41%)
60–64 [n (%)]	19 313 (8.90%)
65–69 [n (%)]	33 709 (15.54%)
70–74 [n (%)]	23 608 (10.88%)
75–79 [n (%)]	19 441 (8.96%)
80–84 [n (%)]	11 791 (5.44%)
85–89 [n (%)]	5095 (2.35%)
90–94 [n (%)]	1367 (0.63%)
96+ [n (%)]	177 (0.08%)
Marital status	
Married/living together [n (%)]	161 735 (74.56%)
Single [n (%)]	18 887 (8.71%)
Divorced [n (%)]	10 225 (4.71%)
Widowed [n (%)]	22 737 (10.48%)
NA [n (%)]	3348 (1.54%)

Table 1: Population characteristics (before imputation). (cont.)

	Study Population (n =216 932)
Country of origin[b]	
Dutch [n (%)]	197 462 (91.02%)
Non-Dutch, western [n (%)]	3535 (1.63%)
Non-western [n (%)]	15 935 (7.35%)
Education level	
Low [n (%)]	19 386 (8.94%)
Middle 1 [n (%)]	77 343 (35.65%)
Middle 2 [n (%)]	60 481 (27.88%)
High [n (%)]	52 610 (24.25%)
NA [n (%)]	7112 (3.28%)
Physical activity[c]	
Complies with none of exercise norms [n (%)]	66 663 (30.73%)
Complies with at least one of exercise norms [n (%)]	132 689 (61.17%)
NA [n (%)]	17 580 (8.10%)
Chronic disease	
No chronic diseases [n (%)]	61 124 (28.18%)
At least 1 chronic disease [n (%)]	130 973 (60.38%)
NA [n (%)]	24 835 (11.45%)
Alcohol use	
Never [n (%)]	22 440 (10.34%)
Former [n (%)]	12 043 (5.55%)
current [n (%)]	176 613 (81.41%)
NA [n (%)]	5836 (2.69%)
Smoking	
Never [n (%)]	82 553 (38.05%)
Former [n (%)]	84 064 (38.75%)
Current [n (%)]	36 124 (16.65%)
NA [n (%)]	14 191 (6.54%)
Body mass index[d]	
Underweight [n (%)]	2326 (1.07%)
Normal [n (%)]	95 240 (43.90%)
Pre-obesity [n (%)]	2326 (1.07%)
Obesity I [n (%)]	22 127 (10.20%)
Obesity II [n (%)]	4644 (2.14%)
Obesity III [n (%)]	1134 (0.52%)
NA [n (%)]	8600 (3.96%)

	Study Population (n =216 932)
Children	
No children [n (%)]	139 605 (64.35%)
Lives with children <18 years old [n (%)]	45 439 (20.95%)
Lives with children ≥18 years old [n (%)]	16 092 (7.42%)
NA [n (%)]	15 796 (7.28%)
COUNTRY LEVEL COVARIATES	
Greenspace (NDVI) in 500m buffer [mean (sd)][e]	0.58 (0.09)
Neighborhood socioeconomic status score [mean (sd)][f]	0.38 (1.01)

a Age was categorized into 5-year categories for <65 year-olds and into 10-year categories for ≥65 year-olds

b Origin was defined as the country of birth of the mother (or that of the father if information on the mother was unavailable). Countries were grouped into Western and non-Western, except for the Netherlands which constitutes a separate category.

c The Nederlandse Norm Gezond Bewegen (NNGB) and Fitnorm are two Dutch common standards for healthy exercise that take into account the amount of time, frequency and intensity of physical activity. Participants were classified into two categories depending on whether they complied with 0="none" or 1="at least one" of these norms.

d Body Mass Index (BMI) categories were defined according to the World Health Organization nutritional status, where Underweight = <18.5, Normal weight = 18.5–24.9, Pre-obesity = 25.0–29.9, Obesity class I = 30.0–34.9, Obesity class II = 35.0–39.9, and Obesity class III = ≥40.

e The NDVI (Normalised Difference Vegetation Index) describes the amount of green vegetation using reflectance measured by satellites. Here, we used the average NDVI within 500 meters of the participant's residence (values: 0 to 1) as calculated by Klompmaker et al.

f We used socioeconomic position as defined by the SCP (Sociaal en Cultureel Planbureau); it is a social status score taking into account average income, percentage of people with a low income, percentage of people with a low education and percentage of unemployed people in a postal code area

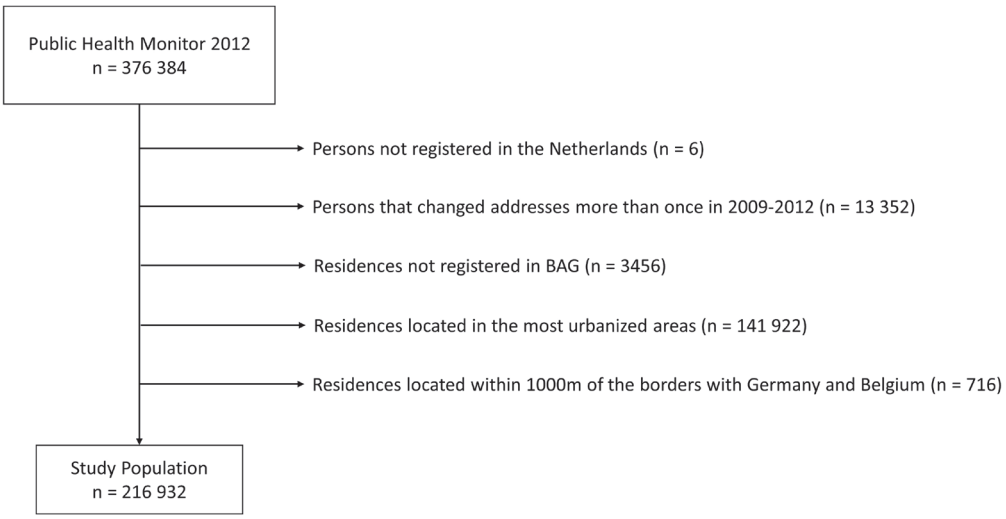
Table 2: Odds Ratios of self-reporting low to severe psychological distress and of self-reporting moderate to very poor health per increase in 1 hectare of area of (specific) treated-crop and their 95% Confidence Intervals (full models).

Crop	Buffer size	Number of participants		Odds Ratio [95% Confidence Interval][1]	
		Unexposed (0 hectares of crop)	Exposed (>0 hectares of crop)	Self-reported psychological distress	Self-perceived health
MAIZE	500m	204959 (94.5%)	11973 (5.5%)	0.79 [0.59, 1.05]	0.68 [0.49, 0.96]
	1000m	190279 (87.7%)	26653 (12.3%)	0.94 [0.89, 1.00]	0.90 [0.84, 0.97]
	1500m	140671 (64.8%)	76261 (35.2%)	0.99 [0.98, 1.00]	0.99 [0.98, 1.00]
	2000m	75816 (34.9%)	141116 (65.1%)	1.00 [1.00, 1.00]	1.00 [1.00, 1.01]
WINTER WHEAT	500m	211278 (97.4%)	5654 (2.6%)	1.06 [0.57, 1.98]	0.53 [0.25, 1.12]
	1000m	204186 (94.1%)	12746 (5.9%)	1.09 [0.96, 1.23]	0.90 [0.77, 1.05]
	1500m	178060 (82.1%)	38872 (17.9%)	1.02 [1.00, 1.04]	0.99 [0.97, 1.02]
	2000m	139146 (64.1%)	77786 (35.9%)	0.99 [0.99, 0.99]	1.00 [0.99, 1.00]
SUMMER BARLEY	500m	215507 (99.3%)	1425 (0.7%)	0.94 [0.17, 5.37]	0.15 [0.02, 1.41]
	1000m	213407 (98.4%)	3525 (1.6%)	0.92 [0.64, 1.33]	0.79 [0.50, 1.23]
	1500m	202898 (93.5%)	14034 (6.5%)	0.97 [0.91, 1.03]	0.99 [0.92, 1.06]
	2000m	176862 (81.5%)	40070 (18.5%)	0.97 [0.96, 0.98]	0.99 [0.97, 1.00]
SUMMER WHEAT	500m	215323 (99.3%)	1609 (0.7%)	0.57 [0.09, 3.54]	0.41 [0.04, 3.79]
	1000m	213107 (98.2%)	3825 (1.8%)	0.85 [0.58, 1.24]	0.70 [0.44, 1.10]
	1500m	202053 (93.1%)	14879 (6.9%)	0.96 [0.91, 1.02]	0.95 [0.88, 1.02]
	2000m	174236 (80.3%)	42696 (19.7%)	0.98 [0.97, 0.99]	0.99 [0.97, 1.00]
OTHER CEREALS	500m	215758 (99.5%)	1174 (0.5%)	0.75 [0.10, 5.37]	0.92 [0.10, 8.78]
	1000m	213659 (98.5%)	3273 (1.5%)	1.02 [0.68, 1.51]	0.90 [0.56, 1.44]
	1500m	202758 (93.5%)	14174 (6.5%)	1.03 [0.97, 1.10]	0.94 [0.87, 1.01]
	2000m	177844 (82.0%)	39088 (18.0%)	1.00 [0.99, 1.02]	1.02 [1.00, 1.03]
POTATOES FOR CONSUMPTION	500m	211606 (97.5%)	5326 (2.5%)	1.16 [0.43, 3.13]	0.82 [0.25, 2.68]
	1000m	204719 (94.4%)	12213 (5.6%)	1.11 [0.91, 1.35]	1.02 [0.81, 1.29]
	1500m	178343 (82.2%)	38589 (17.8%)	1.00 [0.97, 1.03]	1.00 [0.96, 1.04]
	2000m	136573 (63.0%)	80359 (37.0%)	0.99 [0.99, 1.00]	1.00 [0.99, 1.00]
POTATOES FOR STARCH	500m	216525 (99.8%)	407 (0.2%)	0.06 [0.00, 1.04]	1.93 [0.09, 42.32]
	1000m	216107 (99.6%)	825 (0.4%)	0.61 [0.35, 1.06]	0.95 [0.49, 1.83]
	1500m	214709 (99.0%)	2223 (1.0%)	0.98 [0.89, 1.08]	0.99 [0.88, 1.11]
	2000m	211675 (97.6%)	5257 (2.4%)	0.99 [0.98, 1.01]	1.00 [0.99, 1.01]
SEED POTATOES	500m	215616 (99.4%)	1316 (0.6%)	1.03 [0.14, 7.46]	0.39 [0.03, 4.62]
	1000m	214192 (98.7%)	2740 (1.3%)	1.02 [0.70, 1.49]	0.80 [0.50, 1.28]
	1500m	208320 (96.0%)	8612 (4.0%)	1.00 [0.94, 1.06]	0.95 [0.88, 1.02]
	2000m	196123 (90.4%)	20809 (9.6%)	0.99 [0.98, 1.00]	0.98 [0.97, 0.99]

Crop	Buffer size	Number of participants		Odds Ratio [95% Confidence Interval][1]	
		Unexposed (0 hectares of crop)	Exposed (>0 hectares of crop)	Self-reported psychological distress	Self-perceived health
BEETS	500m	212462 (97.9%)	4470 (2.1%)	1.31 [0.42, 4.12]	0.35 [0.09, 1.40]
	1000m	206307 (95.1%)	10625 (4.9%)	1.04 [0.83, 1.31]	0.73 [0.55, 0.96]
	1500m	181942 (83.9%)	34990 (16.1%)	0.98 [0.95, 1.02]	0.95 [0.91, 0.99]
	2000m	141899 (65.4%)	75033 (34.6%)	0.99 [0.98, 0.99]	1.00 [1.00, 1.01]
ORNAMENTAL PLANTS & TREE NURSERIES	500m	215257 (99.2%)	1675 (0.8%)	1.90 [0.90, 4.01]	0.76 [0.32, 1.81]
	1000m	212927 (98.2%)	4005 (1.8%)	1.06 [0.90, 1.25]	0.90 [0.74, 1.08]
	1500m	202567 (93.4%)	14365 (6.6%)	1.00 [0.97, 1.03]	1.00 [0.97, 1.03]
	2000m	177835 (82.0%)	39097 (18.0%)	1.00 [0.99, 1.00]	1.01 [1.00, 1.01]
VEGETABLES	500m	214255 (98.8%)	2677 (1.2%)	0.76 [0.34, 1.72]	0.78 [0.31, 2.00]
	1000m	210653 (97.1%)	6279 (2.9%)	0.95 [0.81, 1.12]	0.87 [0.72, 1.05]
	1500m	194848 (89.8%)	22084 (10.2%)	0.99 [0.97, 1.02]	0.97 [0.94, 1.00]
	2000m	162344 (74.8%)	54588 (25.2%)	1.00 [0.99, 1.00]	1.01 [1.00, 1.01]
FRUIT	500m	214867 (99.0%)	2065 (1.0%)	0.96 [0.55, 1.67]	0.84 [0.45, 1.60]
	1000m	211778 (97.6%)	5154 (2.4%)	1.00 [0.88, 1.13]	0.91 [0.79, 1.06]
	1500m	200303 (92.3%)	16629 (7.7%)	0.99 [0.97, 1.01]	1.00 [0.98, 1.02]
	2000m	180137 (83.0%)	36795 (17.0%)	0.99 [0.99, 1.00]	1.00 [1.00, 1.01]
FLOWER BULBS	500m	215774 (99.5%)	1158 (0.5%)	1.59 [0.49, 5.16]	1.95 [0.52, 7.29]
	1000m	214311 (98.8%)	2621 (1.2%)	1.22 [0.95, 1.57]	1.28 [0.96, 1.71]
	1500m	208613 (96.2%)	8319 (3.8%)	1.01 [0.97, 1.06]	1.04 [0.99, 1.09]
	2000m	196249 (90.5%)	20683 (9.5%)	0.99 [0.99, 1.00]	0.99 [0.99, 1.00]
ALL CROPS	500m	195784 (90.3%)	21148 (9.7%)	0.97 [0.82, 1.15]	0.77 [0.63, 0.93]
	1000m	173195 (79.8%)	43737 (20.2%)	1.00 [0.96, 1.03]	0.93 [0.89, 0.97]
	1500m	110810 (51.1%)	106122 (48.9%)	1.00 [0.99, 1.00]	0.99 [0.98, 1.00]
	2000m	168088 (77.5%)	48844 (22.5%)	1.00 [1.00, 1.00]	1.00 [1.00, 1.00]

[1] Models were adjusted for: adjusted for the area of the considered crop that remained until 500m ('complementary donut'), age, sex, body mass index (BMI), country of origin, marital status, educational level, living with children, having a chronic condition, presence of other crops within 500m of the participant's residence (except for when 'all crops' was the exposure), physical activity, alcohol status, smoking status, neighborhood socioeconomic position and Normalised Difference Vegetation Index (NDVI) within 500m of the participant's residence.
The referent ("unexposed") group in all models was participants with zero hectares of (specific) crop within 500m of their residences.

FIGURES



BAG = Basisregistratie Adressen en Gebouwen, the cadastral dataset containing all addresses in the Netherlands used to compute individual residential exposure proxies.

Figure 1: Flowchart of the study population.

REFERENCES

1. Rehm J, Shield KD. Global Burden of Disease and the Impact of Mental and Addictive Disorders. Curr Psychiatry Rep [Internet]. 2019;21(2):10. Available from: https://doi.org/10.1007/s11920-019-0997-0
2. Vos T, Allen C, Arora M, Barber RM, Bhutta ZA, Brown A, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet [Internet]. 2016 Oct 8;388(10053):1545–602. Available from: https://doi.org/10.1016/S0140-6736(16)31678-6
3. Dickerson AS, Wu AC, Liew Z, Weisskopf M. A Scoping Review of Non-Occupational Exposures to Environmental Pollutants and Adult Depression, Anxiety, and Suicide. Curr Environ Heal Reports [Internet]. 2020;7(3):256–71. Available from: https://doi.org/10.1007/s40572-020-00280-7
4. Ferraro KF, Nuriddin TA. Psychological distress and mortality: are women more vulnerable? J Health Soc Behav [Internet]. 2006 Sep;47(3):227–41. Available from: https://pubmed.ncbi.nlm.nih.gov/17066774
5. Khan N, Kennedy A, Cotton J, Brumby S. A Pest to Mental Health? Exploring the Link between Exposure to Agrichemicals in Farmers and Mental Health. Int J Environ Res Public Health [Internet]. 2019 Apr 12;16(8):1327. Available from: https://pubmed.ncbi.nlm.nih.gov/31013861
6. Meyer A, Koifman S, Koifman RJ, Moreira JC, de Rezende Chrisman J, Abreu-Villaca Y. Mood disorders hospitalizations, suicide attempts, and suicide mortality among agricultural workers and residents in an area with intensive use of pesticides in Brazil. J Toxicol Environ Health A. 2010;73(13–14):866–77.
7. Stallones L, Beseler C. Pesticide poisoning and depressive symptoms among farm residents. Ann Epidemiol. 2002 Aug;12(6):389–94.
8. Beseler CL, Stallones L, Hoppin JA, Alavanja MCR, Blair A, Keefe T, et al. Depression and pesticide exposures among private pesticide applicators enrolled in the Agricultural Health Study. Environ Health Perspect. 2008 Dec;116(12):1713–9.
9. Beard JD, Hoppin JA, Richards M, Alavanja MCR, Blair A, Sandler DP, et al. Pesticide exposure and self-reported incident depression among wives in the Agricultural Health Study. Environ Res. 2013 Oct;126:31–42.
10. Beseler C, Stallones L, Hoppin JA, Alavanja MCR, Blair A, Keefe T, et al. Depression and pesticide exposures in female spouses of licensed pesticide applicators in the agricultural health study cohort. J Occup Environ Med. 2006 Oct;48(10):1005–13.
11. Benyamini Y. Why does self-rated health predict mortality? An update on current knowledge and a research agenda for psychologists. Psychol Health [Internet]. 2011 Nov 1;26(11):1407–13. Available from: https://doi.org/10.1080/08870446.2011.621703
12. Reile R, Stickley A, Leinsalu M. Large variation in predictors of mortality by levels of self-rated health: Results from an 18-year follow-up study. Public Health. 2017 Apr;145:59–66.

13. Tessler R, Mechanic D. Psychological distress and perceived health status. J Health Soc Behav. 1978 Sep;19(3):254–62.
14. Iversen VC, Sam DL, Helvik A-S. Psychological distress and perceived health in inmates in Norwegian prisons. Scand J Public Health [Internet]. 2013 Oct 2;42(2):171–6. Available from: https://doi.org/10.1177/1403494813505728
15. Li J, Dong L, Tian D, Zhao Y, Yang H, Zhi X, et al. Association between pesticide exposure intensity and self-rated health among greenhouse vegetable farmers in Ningxia, China. PLoS One [Internet]. 2018 Dec 21;13(12):e0209566. Available from: https://doi.org/10.1371/journal.pone.0209566
16. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SLT, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med. 2002 Aug;32(6):959–76.
17. Donker T, Comijs H, Cuijpers P, Terluin B, Nolen W, Zitman F, et al. The validity of the Dutch K10 and extended K10 screening scales for depressive and anxiety disorders. Psychiatry Res. 2010 Mar;176(1):45–50.
18. Nationaal Georegister (NGR). Dutch cadastral key-registry of buildings and addresses (BAG) [Internet]. 2016 [cited 2016 Jun 11]. Available from: https://data.overheid.nl/data/dataset/57437-bag-verblijfsobjecten
19. Statistics Netherlands (CBS). StatLine: Landbouw; gewassen, dieren, grondgebruik en arbeid op nationaal niveau 2009–2013 [Internet]. 2013 [cited 2021 Feb 2]. Available from: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/81302ned/table?fromstatweb
20. Garron CA, Davis KC, Ernst WR. Near-field air concentrations of pesticides in potato agriculture in Prince Edward Island. Pest Manag Sci [Internet]. 2009 Jun 1;65(6):688–96. Available from: https://doi.org/10.1002/ps.1746
21. Siebers J, Binner R, Wittich K-P. Investigation on downwind short-range transport of pesticides after application in agricultural crops. Chemosphere. 2003 May;51(5):397–407.
22. Sportdeelname.nl. Definities van de verschillende beweegnormen: NNGB, fitnorm en conbinorm [Internet]. 2015 [cited 2018 Jan 25]. p. 1. Available from: http://www.sportdeelname.nl/wp-content/uploads/sites/2/2015/03/Bijlage-16_definities-NNGB.pdf
23. Knol FA. Van hoog naar laag, van laag naar hoog: de sociaal-ruimtelijke ontwikkeling van wijken tussen 1971–1995 [in Dutch]. The Hague: Sociaal en Cultureel Planbureau; 1998.
24. Klompmaker JO, Hoek G, Bloemsma LD, Gehring U, Strak M, Wijga AH, et al. Green space definition affects associations of green space with overweight and physical activity. Environ Res. 2018 Jan;160:531–40.
25. McHenry J, Carrier N, Hull E, Kabbaj M. Sex differences in anxiety and depression: role of testosterone. Front Neuroendocrinol. 2014 Jan;35(1):42–57.
26. Klompmaker JO, Janssen NAH, Bloemsma LD, Gehring U, Wijga AH, van den Brink C, et al. Residential surrounding green, air pollution, traffic noise and self-perceived general health. Environ Res [Internet]. 2019;179:108751. Available from: https://www.sciencedirect.com/science/article/pii/S0013935119305481
27. Klompmaker JO, Hoek G, Bloemsma LD, Wijga AH, van den Brink C, Brunekreef B, et al. Associations of combined exposures to surrounding green, air pollution and traffic noise on mental health. Environ Int. 2019 Aug;129:525–37.
28. Ward MH, Lubin J, Giglierano J, Colt JS, Wolter C, Bekiroglu N, et al. Proximity to crops and residential to agricultural herbicides in Iowa. Environ Health Perspect. 2006;114(6):893–7.

29. Statistics Netherlands (CBS). StatLine: Activiteiten van biologische landbouwbedrijven; regio. 2021; Available from: https://opendata.cbs.nl/#/CBS/nl/dataset/83922NED/table?ts=1626271082204
30. Paul KC, Sinsheimer JS, Cockburn M, Bronstein JM, Bordelon Y, Ritz B. Organophosphate pesticides and PON1 L55M in Parkinson's disease progression. Environ Int [Internet]. 2017/07/06. 2017 Oct;107:75–81. Available from: https://pubmed.ncbi.nlm.nih.gov/28689109
31. Lorem G, Cook S, Leon DA, Emaus N, Schirmer H. Self-reported health as a predictor of mortality: A cohort study of its relation to other health measurements and observation time. Sci Rep [Internet]. 2020;10(1):4886. Available from: https://doi.org/10.1038/s41598-020-61603-0
32. Segel-Karpas D. Number of Illnesses, Self-perceived Health, and Depressive Symptoms: The Moderating Role of Employment in Older Adulthood and Old Age. Work Aging Retire [Internet]. 2015 Oct 1;1(4):382–92. Available from: https://doi.org/10.1093/workar/wav009
33. Fischer PH, Marra M, Ameling CB, Hoek G, Beelen R, de Hoogh K, et al. Air Pollution and Mortality in Seven Million Adults: The Dutch Environmental Longitudinal Study (DUELS). Environ Health Perspect. 2015 Jul;123(7):697–704.
34. Smit LAM, Hooiveld M, van der Sman-de Beer F, Opstal-van Winden AWJ, Beekhuizen J, Wouters IM, et al. Air pollution from livestock farms, and asthma, allergic rhinitis and COPD among neighbouring residents. Occup Environ Med [Internet]. 2014 Feb 1;71(2):134 LP – 140. Available from: http://oem.bmj.com/content/71/2/134.abstract

SUPPLEMENTAL MATERIAL

Online only: http://dx.doi.org/10.1136/oemed-2021–107544

Contents

CHAPTER 5

EXPLORING ASSOCIATIONS BETWEEN RESIDENTIAL EXPOSURE TO PESTICIDES AND BIRTH OUTCOMES USING THE DUTCH BIRTH REGISTRY

Mariana Simões, Roel Vermeulen, Lützen Portengen, Anke Huss*, Nicole Janssen*

**Shared last author*

Submitted

ABSTRACT

Background: Maternal occupational exposure to pesticides has been linked to adverse birth outcomes but associations with residential pesticide exposures are inconclusive.
Objectives: To explore associations between residential exposure to specific pesticides and birth outcomes using individual level exposure and pregnancy/birth data.
Methods: From all 2009–2013 singleton births in the Dutch birth registry, we selected mothers >16 years old living in non-urban areas, who had complete address history and changed addresses at most once during pregnancy (N=339,947). We estimated amount (kg) of 139 active ingredients (AI) used within buffers of 50, 100, 250 and 500 meters around each mother's home during pregnancy. We used generalized linear models to investigate associations between 12 AIs with evidence of reproductive toxicity and gestational age (GA), birth weight (BW), perinatal mortality, child´s sex, prematurity, low birth weight (LBW), small for gestational age (SGA) and large for gestational age (LGA), adjusting for individual and area-level confounders. For the remainder 127 AIs, we used minimax concave penalty with a stability selection step to identify those that could be related to birth outcomes.
Results: Regression analyses showed that maternal residential exposure to fluroxypyr-meptyl was associated with longer GA, glufosinate-ammonium with higher risk of LBW, linuron with higher BW and higher odds of LGA, thiacloprid with lower odds of perinatal mortality and vinclozolin with longer GA. Variable selection analysis revealed that picoxystrobin was associated with higher odds of LGA. We found no evidence of associations with other AIs. Sensitivity and additional analysis supported these results except for thiacloprid.
Discussion: In this exploratory study, pregnant women residing near crops where fluroxypyr-meptyl, glufosinate-ammonium, linuron, vinclozolin and picoxystrobin were applied had higher risk for certain potentially adverse birth outcomes. Our findings provide leads for confirmatory investigations on these compounds and/or compounds with similar modes of action.
KEYWORDS: pesticides, residential exposure, spatial analysis, general population, birth outcomes, birth registry

INTRODUCTION

Maternal exposure to pesticides has been linked to adverse birth outcomes such as low birth weight, decreased gestational age, being small for the gestational age, prematurity, changed sex ratio and stillborn and infant mortality.[1–3] There is a substantial body of literature supporting these associations among mothers occupationally exposed to pesticides, but the picture is less clear when mothers are residentially exposed. Although research has demonstrated that pesticides are found in residences located near crops[4–6], studies on the typically low-dose non-occupational exposure and birth outcomes present inconsistent results. This discrepancy may be due to the application of different exposure assessment methodology (often based on (single time point) measurement of biomarkers of pesticide exposure, questionnaires or proximity of residencies to agricultural plots) or differences between countries regarding agricultural scenarios and permitted types of pesticides and pesticide application methods. For example, some studies reported associations between exposure to pesticides and decreased birth weight and/or gestational age, these associations were not shown in other studies[7–10] Clear evidence that residential exposure to pesticides presents a health risk for pregnant women and their babies is thus still lacking and, consequently, making it uncertain whether precautionary measures are needed to reduce pesticide exposure for this especially vulnerable population group.

In this paper we explore the possible associations between residential exposure to specific active ingredients during pregnancy and several birth outcomes, namely gestational age, birth weight, perinatal mortality and child´s sex. We used crop maps and farmer's surveys to individually estimate residential exposure and the Dutch birth registry for outcome assessment. We first investigated active ingredients reported to have reproductive and developmental effects (i.e. hypothesis testing) and then further used a variable selection method to identify other relevant active ingredients among 139 active ingredients used during the exposure period assessed (2009–2013) (discovery analyses).

METHODS

Study population

The Perinatale Registratie Nederland (PRN) comprises data on pregnancy and births registered by medical professionals such as midwives, general practitioners, gynecologists and pediatricians/neonatologists. The data are linked to the municipal reg-

istration (GBA – Gemeentelijke Basis Administratie) within Statistics Netherlands (CBS - Centraal Bureau voor de Statistiek). The resulting dataset includes all mother-infant pairs for which mothers were registered at the GBA and infants had a gestational age of 22 weeks or more at birth.
For this study we selected singleton births that occurred before 01–01–2014 for which the day of conception was estimated to be on or after 01–01–2009. We excluded mothers that were aged ≤16 years at the child's birth, mothers with an unknown address sometime during pregnancy, and mothers who changed addresses more than once during pregnancy. Furthermore, mothers who lived within 500m of the borders with Germany or Belgium, and for whom we were unable to compute exposure, were excluded. Finally, we restricted our study to mothers living in non-urban areas of The Netherlands (<1500 addresses per km^2), thus excluding those whose health behaviors, lifestyles and environmental exposures are likely different from those living in rural areas and whose residences are likely not located near any crops (Figure 1).

Birth outcomes

We evaluated the following main outcomes: gestational age, birth weight, perinatal mortality (up to 1 year of age) and child's sex. We further explored transformations of the outcomes gestational age and birth weight as low birth weight, small for gestational age, large for gestational age and prematurity as follows:

- **gestational age**: the number of days between the last menstruation and birth;
- **birth weight**: weight at birth, in grams;
- **low birth weight**: less than 2500 gram at birth (binary variable),
- **small for gestational age** and **large for gestational age**: we constructed birth weight curves based on this study population (singletons, years 2009–2013) using the Lambda Mu and Sigma (LMS) method (library GAMLSS from R) considering gestational age (in days), sex of the baby (male, female), migration background of the mother (Dutch, Western, Non-Western) and parity (primipara, multipara). These curves were used to estimate the 10^{th} and 90^{th} birth weight centiles, which defined the thresholds for small for gestational age and large for gestational age, respectively (binary variables);[11]
- **perinatal mortality**: a binary composite variable including still births and infant mortality within the first year of life;
- **prematurity**: birth occurring before 37 weeks of gestation (binary variable);[12]
- **child's sex:** sex assigned at birth (binary variable).

Exposure assessment

We estimated mothers' residential exposure to pesticides by computing the amount (kg) of specific active ingredients (AIs) used within buffers of 50, 100, 250 and 500m around their residences during pregnancy. For this we used the Basisregistratie Adressen en Gebouwen containing geocoded residences from 2016 (BAG) and the annual polygon land use maps from 2009 to 2013 (Basisregistratie Gewaspercelen, BRP) to compute the area, in hectares (ha), of 12 groups of crops grown in open fields (greenhouses thus excluded) within these buffers: maize, winter wheat, summer barley, summer wheat, potatoes for consumption, potatoes for starch, seed potatoes, beets, ornamental plants and tree nurseries, vegetables, fruit trees and flower bulbs[13,14] These crops account for 86–87% of the total area of arable land (excluding grassland) of the Netherlands in 2009-2013[15]. For computational reasons, we converted the land use maps to raster maps with a resolution of 10m by 10m and used a moving average to obtain squared buffers that correspond roughly to radii of 50m, 100m, 250m and 500m, as previously described[16] We then used data from the 2008 and 2012 CBS's Farmers' Survey, a national survey administered to a sample of farmers roughly every 4 years, to obtain information on the average annual dosage of active ingredients used (amount per hectare, kg/ha) for each crop group. The total amount of active ingredient used in a crop within a buffer around a residence was then estimated by multiplying the dosage used by the estimated crop area. The total annual amount of AI used within a buffer around a residence was then obtained by summing the amounts used for all crops present in the buffer. We used these estimates as proxies for the amount of AIs mothers were exposed to at their residences during pregnancy. When conception and birth occurred in different years, we used the average of estimates for the separate years, weighted by the number of gestation days in each year. A similar approach was used for estimating exposure for mothers that changed address during pregnancy. A more detailed description and example calculations can be found in S1 of the Supplementary Material.

We included 139 AIs that were used by at least 10% of the surveyed farmers in 2008 and 2012 in these calculations. Of these, 12 were classified as reproductive toxicants from Category 1B (presumed human reproductive toxicant based clear evidence of an adverse effects from animals studies) in the European Union Pesticide Database (EUPDB) or as "known" to cause a problem in reproduction or development in the Pesticide Properties Database (PPDB)[17,18] Another 19 were classified as Category 2 (suspected human reproductive toxicant based on some evidence from humans or experimental animals studies) or as "possible" to cause a problem in reproduction or development in the EUPDB and the PPDB, respectively. There is no evidence of reproductive or developmental effects for the remainder AIs (Table 1).

Potential confounders

We included the following individual covariates as potential confounders: gestational age (not for the gestational age and prematurity analyses), sex of the baby (not for the child's sex analyses), parity, mother's migration background, maternal age at delivery, mother's educational level, mother's household income, mother's marital status, mother's job status, and year of birth. We further considered one area level covariate, namely urbanization degree of the neighborhood where the mother lived. Classes of these covariates can be seen in Table 2.

Statistical analysis

Imputation

Some of the covariates had missing data: household income (1.0%), educational level (28.1%) and degree of urbanization (0.002%). We used multiple imputation by chained equations (MICE) to impute missing values, using the *MICE* package in R, and included all outcome variables, all covariates and the total area of crops in the 500m buffer for the imputation models. Additionally, we also considered available variables that were determined to be highly predictive of the variables with missing values: job sector of mother and father's job, proportion of people employed in the neighborhood and proportion of people with low income in the neighborhood. Variables were deemed predictors and included in the imputation models when their correlation to the variable being imputed was >0.20 and the proportion of usable cases was >0.25. Collinearity of the selected variables assessed by the variance inflation factor (VIF). Variables with VIF>5 were considered collinear and the variable with highest VIF was excluded from the imputation model. Due to the large number of observations in the dataset, we used the predictive mean matching to impute all variables. Since the predictive mean matching algorithm only imputes values that are already present in the data, the original classes of categorical variables were maintained[19] We imputed 5 data sets, using 8 iterations, to limit computational overhead. We assessed the imputed datasets by evaluating plots of the mean and standard deviation of the imputed values, per iteration. These showed that convergence was fast and achieved after the second or third iteration, with very little trend (not shown). The kernel density estimates for the marginal distributions showed that the densities of the observed data and the five imputed data sets, per variable, had essentially the same shape.

Regression analyses and exposure variables selection

We investigated associations between the amount of AIs used in the vicinity of pregnant women's residences and birth outcomes using two approaches.
One, from the pool of 139 AIs, we selected those classified as Category 1B or "known to cause problems" (henceforth referred to as "*a priori* selected AIs", n=12, see Table 1), i.e, AIs with clear evidence for toxic effects on reproduction and fetal development. For these AIs we investigated associations with gestational age and birthweight using linear regression models, and with perinatal mortality, child´s sex, prematurity, low birth weight, small for gestational, and large for gestational age using logistic regression models. Each model included all other AIs which correlation to the AI under investigation was below 0.7 (see S2 for correlations between AIs and the footnote from Table 3 for the AIs included in the models). Associations were explored separately for each buffer size. We considered that the narrower buffers (i.e with a radius of 50 and 100m) were more likely to capture direct spray drift of pesticide droplets. This drift has an exponential decrease in concentration with distance and is highest within the first meters of application, although it can be detected up to 100 meter away from a field edge depending on the application technique and meteorological conditions[20,21] The larger two buffers were considered to capture secondary emission processes such as volatilization. Moreover, a Dutch pesticide exposure assessment study reported high contrasts in pesticide concentration in air and house dust between residences located within 250m and beyond 500m from flower bulb crops[6] For these reasons, we considered that relevant residential exposure occurred within 500m and each buffer model was adjusted for the amount of AI used in the area up to 500m, a variable we named "complimentary donut" (for example, for the 50m buffer model we included the amount of AI used in a "donut" of 50 to 500m around the residence as a covariate). Consequently, the referent ("unexposed") group in our analyses consisted of mothers with zero kilograms of (specific) active ingredient within 500m of their residences. Buffers and complementary donuts were used as continuous variables in the analyses. We also explored increasingly adjusted models, starting with a *basic* model adjusted for gestational age (when gestational age or prematurity were not the outcomes), sex of the baby (when child´s sex was not the outcome), parity and the complementary donut, followed by an *intermediate* model consisting of the basic model and further adjusted for household income, mother's education, mother's marital status, mother's origin, mother's age at birth and mother's job status, and a *full* model, which additionally included the degree of urbanization of the residence location and year of birth of the baby. For binary outcomes, analyses were performed when there were at least 10 exposed cases.

Second, from the remainder 127 AIs, we included those with correlations below 0.90 in the 500m buffer to run the variable selection models (S3). Consequently, AIs were selected from a pool of 47 AIs (Table 4). The models were adjusted for all considered covariates (i.e., equivalent to the full model of the regression analyses) and for the AIs for which the previous regression analyses showed evidence of an association with a birth outcome – no penalization was applied to the coefficients for these covariates and they were therefore not subject to variable selection. Variable selection was performed using penalized regression with the minimax concave penalty (MCP) for which the appropriate level of penalization was selected with 10-fold cross-validation. We then applied a stability selection step to estimate selection probabilities for each variable under random resampling of the data. This approach allows for (multiple-testing) error control and yields more consistent selections. For the stability step we used R's *stabs* package, but modified the algorithm to allow use of the ncvreg() function with the MPC penalty. Since this approach has exploratory purposes, stability selection models were run with per-family error rate (PFER) set at the value of 1, a rather lenient approach in regard to controlling for false discoveries. Furthermore, the model sampled from the five different imputed data sets in each subsampling iteration, using stratified sampling so that every dataset was used as often as the others.
Spatial analyses to compute the crop areas within buffers around residences was performed using ArcGIS/ArcPy 10.4 (2016) and STATA/MP 14. Statistical analyses, as well as the imputations and the calculation of the birth weight curves, were performed in R versions 3.4.1 (2017–06–30) and 3.6.3 (2019–12–12), within the remote secure environment of CBS.

Criteria for interpretation

After assessing whether effect sizes of the AIs materially changed with increasing confounder adjustment in the regression analyses, we based the interpretation of the results on a pre-defined set of criteria, similarly to what we used in previous studies from our research group[16,22] We considered that there was evidence of an association between an AI and a birth outcome (a *finding*) when results from the regression analyses met all the following criteria:

- Consistency among buffer models: the associations had the same direction across all buffers
- Trend in strength of the associations (monotonicity): there was a monotonic trend from the smallest to the largest buffer
- Statistical significance: at least one statistically significant result (p-value < 0.05) among the four buffers

We further defined a criterium to interpret the results from variable selection analysis:

- the AI was selected in at least three of the buffers

If the variable selection analyses pointed to an association between an AI and a birth outcome in at least three buffers, we conducted regression analyses and applied the 3 criteria to assess if there was evidence of an association between the selected AI and the birth outcome.

Sensitivity and additional analyses

We performed several sensitivity analyses for the associations that were considered findings according to the abovementioned criteria. First, we excluded gestational age in the models where it was used as a covariate, since it is unclear whether (low) gestational age is a confounder or (proxy for) some intermediate factor on the causal path to other birth outcomes. Second, we excluded all mothers that worked in the agricultural sector as they may experience occupational pesticide exposure besides residential exposure. Third, we excluded all mothers and fathers that worked in the agricultural setting, focusing analysis on residential exposures only. By excluding mothers working in agriculture during pregnancy we excluded sources maternal occupational exposure to pesticides. Since fathers working in the agriculture could be carriers of pesticides into their homes and there may also be an association between paternal occupational pesticide exposure and adverse birth outcomes, excluding both parents further decreased the contribution of occupational and para-occupational sources of exposure to pesticides. Fourth, we restricted analyses to mothers living in the most rural areas (< 1000 addresses per km^2), to assess potential confounding from living in a semi-urban environment. Fifth, we restricted analyses to autochthonous mothers, since having a migration background may correlate both with exposure and different (health) behaviors during pregnancy and delivery. Sixth, we performed a complete case analysis, to assess whether data could be missing not at random[23] Seventh, since exposure was estimated based on annual usage of pesticides and did not take into account the time of year when each AI was applied, changes in address may have contributed to exposure misclassification. To decrease uncertainty around the exposure estimates, we performed an analysis restricted to mothers that did not change addresses during pregnancy.

Additionally, for associations that were considered findings, we performed analyses including an interaction term between sex and the AI and stratified analyses by sex if there was at least some evidence of differential effects from experimental or toxicological studies for the AI under investigation.

RESULTS

We included a total of 339,947 mother-child pairs in our analysis. Table 2 describes the demographic characteristics of the study population. The relatively small number of births in 2009 reflects the fact that we only included pregnancies conceived after 01–01–2009. Exposure to at least one active ingredient occurred 6%, 14%, 40% and 70% of mothers living within 50m, 100m, 250m and 500m of a crop, respectively. The most common active ingredients which mothers were exposed to were florasulam and fluroxypyr-meptyl (4%, 10% and 32% in the in 50m, 100m and 250m) and dimethenamid (62%) in 500m. Interquartile ranges (IQRs) of amount if AI used ranged from 0.0001 kg for florasulam to 2.5 kg for sulphur in 50m and from 0.006 kg for rimsulffuron to 79.2 kg for sulphur in 500m (see S4 for descriptive statistics of each AI). Among exposed mothers, the median number of AIs mothers were exposed to was 22 [minimum 9, maximum 107] in 50m, 25 [9,121] in 100m, 34 [9, 133] in 250m and 50 [9, 139] in 500m. In the 500m buffer, among the 12 *a priori* selected AIs, correlations were high (>0.70) between glufosinate ammonium and propioconazol, between glufosinate ammonium and triadimenol, between thiacloprid and carbetamide, between asulam and thiacloprid, and between asulam and vinclozolin (S2).

Models adjusted for all potential confounders showed differences in effect sizes larger than 10% in more than half of the models where GA and BW were the outcomes, showing mostly smaller effect sizes. For consistency, Table 3 shows the results of the fully adjusted regression analyses for all outcomes expressed as increase in days for gestational age, increase in grams for birth weight and the Odds Ratio (OR) for the other outcomes per 1 kg increase in AI used within the buffer (results from the basic and intermediate models are shown in S5). Among the *a priori* investigated AIs, six findings emerged that met our criteria for being considered a finding: maternal residential exposure to fluroxypyr-meptyl was associated with longer GA, glufosinate-ammonium was associated with higher risk of having LBW babies, linuron was associated with higher BW and higher odds of having LGA babies, thiacloprid was associated with lower risk of perinatal mortality, and vinclozolin was associated with longer GA. Variable selection indicated that picoxystrobin was associated with being LGA and with perinatal mortality (Table 4). Posteriorly, regression analysis showed that exposure to picoxystrobin was associated with higher odds of LGA and complied with the interpretation criteria, rendering this result an additional finding (Table 5). Regression results for the association between picoxystrobin and mortality did not meet or interpretation criteria and the association was not considered a finding.

Results were mostly robust to sensitivity analyses (Table 6). Although some sensitivity analyses showed relevant changes in effect sizes (more than 10% change compared to

the main analysis effect size), the majority of results complied with our interpretation to be considered findings. The exceptions to this were the sensitivity analysis on the association between exposure to vinclozolin and increased GA, in which most sensitivity analysis did not follow the interpretation criteria, and on the association between exposure to thiacloprid and lower odds of perinatal mortality, in which none complied with the interpretation criteria. Given that linuron and vinclozolin have antiandrogenic effects and thiacloprid has estrogenic effects that could result in sex-differential effects,[24–26] we further conducted an analysis including an interaction term for sex these and AIs and analyses stratified by sex. While the interaction term sex:linuron was significant only in the 50m and 100m buffers, the analysis stratified by sex showed that the exposure to linuron was associated to higher BW in girls and that this effect was absent in boys, given non-compliancy with the interpretation criteria (Table 7). Regarding the association between linuron and being LGA, only the interaction term of the 50m buffer was statistically significant and stratified analysis showed that there was evidence of an association in girls but not in boys. For the associations between vinclozolin and GA and between thiacloprid and perinatal mortality, we observed that none of the interaction terms across the buffers were statistically significant, while stratified analyses showed that results did not comply with the interpretation criteria.

DISCUSSION

We used the Dutch national birth registry to build a birth cohort and investigate associations between maternal exposure to 139 active ingredients used in the vicinity of pregnant women's residences and several birth outcomes. After defining a set of three criteria to evaluate the results (same direction of effect, monotonic trend and statistical significance in at least one buffer), we identified seven findings: maternal residential exposure to fluroxypyr-mepty during pregnancy was associated with longer GA, glufosinate-ammonium with higher odds of having LBW, linuron with higher BW and higher odds of being LGA (namely in baby girls), thiacloprid with lower odds of perinatal mortality, vinclozolin with longer GA, and picoxystrobin with higher odds of being LGA. The first six associations arose from an investigation on a set of 12 AIs known to cause reproductive and/or developmental problems. Picoxystrobin was identified to be associated to LGA using a variable selection method that was applied to the remainder 127 AIs. After performing regression analysis and applying the interpretation criteria, we classified this association as an additional finding. We also observed that picoxystrobin was selected as a variable associated to perinatal mortality but regression analyses did not meet the *a priori* interpretation criteria of an effect. Sensitivity

and additional analyses supported these results, with exception of those pertaining the association between thiacloprid and lower risk of perinatal mortality. They also revealed that the effects of linuron on BW and being LGA were more pronounced and statistically significant in girls.

By using the birth registry we were able to include complete information collected and registered by medical professionals on birth outcomes for all singleton births among mothers living in rural areas of the Netherlands. We were thus able to obtain a large study population of nearly 340,000 mother-infant pairs and circumvent selection bias. We further enriched the data set by including covariables from other administrative microdata, namely mother's educational level, household income, marital status and job status. Another advantage of using administrative data is access to the exact address information and residential history, which allowed for exposure assignment at individual level. Indeed, we were able to assess residential exposure for several active ingredients by estimating the amount of active ingredients used in the vicinity of the residences. This objective exposure assessment was based on registry-based annual land use maps and information on pesticide usage from two Farmers' Surveys. The area of crops around residences has been previously shown to be suitable in estimating pesticide levels in residences located near crops,[5,27] and the Farmers' Surveys were conducted in a representative sample of Dutch farmers, covering all the main crops grown in the country. Nevertheless, the computed proxies are not exempt of assumptions and limitations that resulted in exposure misclassification. Firstly, they do not consider all sources of exposure to pesticides. We did not consider time-activity patterns, presence at the residence during spraying events (which may influence exposure levels) and occupational exposures. Of these, we were able to assess the impact of working in agriculture by doing sensitivity analyses excluding mothers and excluding mothers and fathers working in this sector. These analyses showed important differences (>10%) when compared to the main analyses, especially in the smaller buffer sizes but overall remained indicative of the effects of the AIs on birth outcomes even when occupational exposure was excluded. We could not account for usage of pesticides in and around the home or for nutritional exposure to pesticides. However, it is questionable that these constitute strong confounding factors as it is unlikely that they differ greatly across the exposures to specific AIs, especially within the short distances that we assessed. Secondly, we used symmetric (squared) buffers, not considering wind direction and speed and thus disregarding the actual spread from both direct and secondary drift. The most suitable non-symmetrical shape of the buffer is difficult to determine but previous studies have shown that area of crop, even within symmetrical buffers, is among the most important variables to assess residential pesticides levels[5,27] Thirdly, Farmers' Surveys pertain to an average pesticide usage among a sample of farmers and no information on the actual dose used

in each field was available. In this study, we assumed that all fields growing a certain crop used the dosage of AI reported in the survey. Fourthly, because Farmers' Surveys were available only for 2008 and 2012, we assumed that the dosages used in the years between surveys were either the same or an average of the two available years (see S1), but it is likely that important changes in usage occurred. Furthermore, the availability of information was only on the annual usage of AIs and we could not take into account that most AIs have a seasonal application (which further precluded analyses by trimester, when vulnerability to chemical aggressors may change). Finally, the fact that some mothers (9%) changed addresses during pregnancy may have resulted in some degree of change in the type of pesticides they were exposed to by moving to a residence that could be near other types of crops, especially regarding the closest crops. To tackle this, we computed weighted averages of the exposure for the different addresses, and additionally performed a sensitivity analysis restricted to mothers that did not change addresses. These analyses showed mostly attenuated estimates (closer to the null) in the two smaller buffers, but no major differences in the larger buffers. This may be related to the abovementioned differences in type of crop that resulted in a "dilution" of the exposure in the smaller buffers after computing the weighted average. Of note, none of these factors contributing to exposure misclassification would be considered non-differential, which means that they may have led to underestimation of effects from the regression analyses.

Since we used administrative databases, we were unable to account for individual level lifestyle factors such as smoking, alcohol use, drug use or pre-pregnancy Body Mass Index (BMI). For most of these factors one would not expect that they are associated to exposure to specific AIs and it is also unlikely that they differ substantially within the short distances investigated. Major differences in lifestyle are seen between urban and rural populations and to reduce this potential residual confounding we conducted a sensitivity analysis restricted to mothers living in the most rural areas (<1000 addresses per km^2). These analyses showed no material differences from the estimates of the main analysis that would change the overall interpretation of results. Furthermore, a similar study on residential proximity to crops and depression and perceived health on a sample of the Dutch population, several of these lifestyle factors were taken into account in the models, but results showed no major changes in the estimates compared to simpler models[16].

The study entails a high number of tests, since several outcomes, AIs and buffer sizes were assessed. To address the resulting elevated number of false positives from the regression analyses, we applied a set of criteria for interpretation. Although we might consequently have overlooked a signal, the application of these criteria was not so stringent as to preclude identification of the most important associations. Among the six AIs that showed evidence of an association with a birth outcome according to our

criteria, three pertain to pesticides that have been reported to have endocrine disruptive effects. Linuron, which we found to be associated to higher BW and higher odds of being LGA, is a herbicide that inhibits root growth (Table 8). Vinclozolin, which was associated with longer GA is a fungicide that prevents spore germination and mycelial growth. Both linuron and vinclozolin are endocrine disruptors with antiandrogenic effects, as shown in toxicologic/experimental research in both in vitro and in vivo studies[28,29] While linuron was reported to be hazardous to male reproductive development, namely by hampering gonadal organ development, vinclozolin was reported to inhibit sex differentiation in males[24] Our additional analyses showed that the effects observed for linuron were mainly seen in girls. It is unclear how exposure to linuron could have resulted in larger baby girls but the balance in testosterone/estrogen levels is likely important to fetal growth. Vinclozolin is among one of the least used (S4). Together with the loss of power when restricting analysis to certain groups, this could at least partially contribute to why sensitivity and additional analyses did not fully corroborate the association between exposure to vinclozolin and increased GA found in the main analysis. Thiacloprid is an insecticide which was also reported to have endocrine-disrupting effects, namely estrogenic activity[26] Toxicological studies on rats and rabbits observed reduced maternal and fetal body weight, altered sex hormones during pregnancy leading to dystocia and delayed sexual maturation[30] Although these effects may be relevant to humans, the reasons for observed lower risks of perinatal mortality in mothers exposed to this AI in the two larger buffers remain unclear, especially given that toxicological studies also point to reduced pup viability. This could have been a chance finding in this study, since sensitivity and additional analyses did not support this result. Toxicological and experimental literature on the potential effects of fluroxypyr-meptyl and glufosinate-ammonium is scarcer, but they are listed in the PPDB as having reproductive and developmental effects. Fluroxypyr-meptyl is a formulation of the herbicide fluroxypyr and is listed in the PPDB as "known to cause a problem", but EFSA's peer-review on its parent compound reports no evidence of reproductive or developmental effects[31] We observed higher risk of LBW in mothers exposed to the herbicide glufosinate-ammonium, but toxicological studies in rats and rabbits showed that this AI induced lower fertility, abortions, fetus death and premature deliveries[32] Of note is that glufosinate-ammonium was highly correlated to several other AIs (Table 8) including propiconazole, an AI also listed in the EUPDB as a reproductive toxicant. Our results showed that propiconazole was indeed associated with increased odds of LBW, except in the 250m buffer, but none of the OR were statistically significant. A reason for this discrepancy could be the lower number of exposed cases, compared to glufosinate-ammonium. However, we cannot completely rule out propiconazole as potentially being associated with LBW.

In this exploratory study we have used a variable selection method to uncover other potential links to adverse birth outcomes. From many possibilities, we chose the MPC method since it allows for a relatively fast computation on a large data set. MPC works in a similar way to the LASSO penalty but applies less shrinkage to the nonzero coefficients, reducing bias comparatively to LASSO. We set one interpretation criterium (the AI had to be selected in at least three buffers) to reduce the risk of a Type 1 error (false positive findings) due to multiple testing and analyzing a large data set, but we were still mostly interested in identifying potential pesticide candidates for further investigation. Indeed, in this study we considered it more important to avoid a high false negative rate than to wrongly select an AI and we were already hampered by the expected weak associations due to the small effect sizes usually seen in environmental epidemiological studies and the discussed sources of exposure misclassification. We observed that the associations between picoxystrobin and LGA and perinatal mortality complied with our *a priori* interpretation criteria. These results were further investigated with regression analyses. While these later identified the association between picoxystrobin and increased risk of being LGA as an additional finding (but not the association with perinatal mortality), to the best of our knowledge, studies in rats and rabbits have not provided evidence of fertility, reproductive or developmental effects up to date[33] Correlations between picoxystrobin and other AIs were low (<0.5), and therefore no other candidates for the observed effects were considered.

In general, our findings point to higher BW, higher risk of LGA and longer GA. Although lower birth weight is usually of more concern for newborns health and survival, studies have indicated that higher BW and being LGA are associated with increased risk of neonatal complications and with increased risk of obesity and cardiovascular diseases later in life[34,35] Except for fluroxypyr-meptyl, none of the findings' AIs are nowadays approved in the European Union (EU), but exposure via contaminated foods from countries in which it is still used is possible. Additionally, these AIs were highly correlated to other AIs (Table 8) and/or share the same modes of action with AIs that are currently in use in the EU and other countries worldwide. Therefore, future research on the effects of maternal exposure to pesticides can consider the findings of this study as leads.

In conclusion, we observed associations between residential exposure to five AIs (fluroxypyr-meptyl, glufosinate-ammonium, linuron, vinclozolin and picoxystrobin) and potentially adverse birth outcomes. The underlying mechanism driving these effects are unclear, but the findings warrant more research into the effects of (non-occupational) exposure to these pesticides on human health, especially in the vulnerable population of pregnant women and their babies. AIs that were correlated or that share the same modes of action with the identified in this study may also be considered as leads for further research.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Health, Welfare and Sport (VWS), in the context of the Policy Advisory on Plant Protection Products. We would like to thank Perined for allowing the use of their data on pregnancy and birth outcomes and Statistics Netherlands (CBS) for providing remote access to a secured environment for statistical analyses. We would also like to thank Rob Vijftigschild (CBS) for his assistance regarding the pesticide usage data.

TABLES

Table 1: Active ingredients included in the study (n=139). The first 12 active ingredients have evidence of adverse reproductive and/or developmental effects in humans and were explored using regression analyses.

Reproduction and/or development effects	Active ingredients			
EUPDB Category 1B (H360D, H360F, H360Df, H360FD, H360Fd) Presumed human reproductive toxicant (evidence from animal studies) **PPDB** Yes, known to cause a problem in reproduction and/or development	Asulam	Glufosinate-ammonium		
	Carbetamide	Linuron		
	Cyproconazole	Propiconazole		
	Epoxiconazole	Thiacloprid		
	Fluroxypyr-meptyl	Triadimenol		
	Glufosinate	Vinclozolin		
EUPDB Category 2 (H361d, H361f, H361fd) Suspected human reproductive toxicant (some evidence from human and/or animal studies) **PPDB** Possibly capable of causing a problem in reproduction and/or development, status not identified	Abamectin	Ioxynil octanoate	Tembotrione	
	Amitrole	Mancozeb	Tepraloxydim	
	Benthiavalicarb isopropyl	Maneb	Triflusulfuron-methyl	
	Cycloxydim	Penconazole		
	Cymoxanil	Spirodiclofen		
	Fenpropimorph	Spirotetramat		
	Fluazifop-p-butyl	Sulcotrione		
	Fluazinam	Tebuconazole		
EUPDB Not classified as reproductive toxicant **PPDB** Known not to cause a problem	2,4-D	Dimethoate	Iodosulfuron-methyl-sodium	Prochloraz
	Acetamiprid	Dimethomorph	Iprodione	Procymidone
	Aclonifen	Diquat	Isoproturon	Propamocarb
	Azoxystrobin	Dithianon	Kresoxim-methyl	Propyzamide
	Bentazone	Dodine	lambda-Cyhalothrin	Prosulfocarb
	Bifenazate	Emamectin benzoate	Mandipropamid	Prothioconazole
	Bifenox	Esfenvalerate	MCPA	Pymetrozine
	Bitertanol	Ethofumesate	Mecoprop-P	Pyraclostrobin
	Bixafen	Fenamidone	Mepanipyrim	Pyridate
	Boscalid	Fenhexamid	Mesosulfuron-methyl	Pyrimethanil
	Bupirimate	Fenoxycarb	Mesotrione	Pyroxsulam
	Captan	Fenpropidin	Metalaxyl-M	Quinoclamine
	Chlorantraniliprole	Flonicamid	Metamitron	Quizalofop-P-ethyl
	Chloridazon	Florasulam	Metazachlor	Rimsulfuron
	Chlorothalonil	Fludioxonil	Methiocarb	S-Metolachlor
	Chlorpropham	Fluopicolide	Methoxyfenozide	Spinosad
	Clomazone	Fluoxastrobin	Metiram	Sulphur
	Clopyralid	Fluroxypyr	Metoxuron	Tebufenpyrad
	Copper oxychloride	Folpet	Metribuzin	Teflubenzuron

→

Table 1: Active ingredients included in the study (n=139). The first 12 active ingredients have evidence of adverse reproductive and/or developmental effects in humans and were explored using regression analyses. (cont.)

Reproduction and/or development effects	Active ingredients			
EUPDB Not classified as reproductive toxicant	Cyazofamid	Foramsulfuron	Metsulfuron-methyl	Terbuthylazine
	Cydia pomonella granulovirus	Fosetyl	Nicosulfuron	Thiamethoxam
PPDB Known not to cause a problem	Cyprodinil	Fosetyl-aluminium	Pencycuron	Thiophanate-methyl
	Deltamethrin	Glyphosate	Pendimethalin	Thiram
	Desmedipham	Haloxyfop-p-methyl	Phenmedipham	Tolclofos-methyl
	Dicamba	Imazalil	Picoxystrobin	Topramezone
	Difenoconazole	Imidacloprid	Pirimicarb	Tri-allate
	Dimethenamid-P	Indoxacarb	Pirimiphos-methyl	Trifloxystrobin

EUPDB = European Union Pesticide Database
PPDB = Pesticide Properties Database
Classification (Reg. 1272/2008):
H360D = May damage the unborn child.
H360Df = May damage the unborn child. Suspected of damaging fertility.
H360FD = May damage fertility. May damage the unborn child.
H361f = Suspected of damaging fertility.
H361d = Suspected of damaging the unborn child.
H361fd = Suspected of damaging fertility. Suspected of damaging the unborn child.

Table 2: Demographic characteristics of the study population (before imputation).

	Study Population (n =339,947)
PREGNANCY OUTCOMES	
Gestational age (days) [mean±sd]	276.2±14.1
Birth weight (g) [mean±sd]	3,460.5±589.3
Perinatal mortality [n (%)]	2,926 (0.9)
Infant sex (boys) [n (%)]	174,373 (51.3)
Premature babies (<37 weeks) [n (%)]	20,792 (6.1)
Low birth weight [n (%)]	15,738 (4.6)
Small for gestational age [n (%)]	32,951 (9.7)
Large for gestational age [n (%)]	35,141 (10.3)
INDIVIDUAL COVARIATES	
Parity	
1 [n (%)]	147,231 (43.3)
≥2 [n (%)]	192,716 (56.7)

	Study Population (n =339,947)
Migration background	
Dutch [n (%)]	288,446 (84.9)
Non-Dutch, western [n (%)]	27,597 (8.1)
Non-western [n (%)]	23,904 (7.0)
Maternal age at delivery	
17–19 [n (%)]	3,158 (0.9)
20–29 [n (%)]	138,841 (40.8)
30–34 [n (%)]	129,847 (38.2)
35–40 [n (%)]	57,877 (17.0)
≥40 [n (%)]	10,224 (3.0)
Maternal education level	
Low [n (%)]	7,164 (2.1)
Medium [n (%)]	125,446 (36.9)
High [n (%)]	111,824 (32.9)
Missing [n (%)]	95,513 (28.1)
Household income (quintiles)	
1st quintile [n (%)]	49,741 (14.6)
2nd quintile [n (%)]	64,929 (19.1)
3rd quintile [n (%)]	91,518 (26.9)
4th quintile [n (%)]	80,799 (23.8)
5th quintile [n (%)]	49,444 (14.5)
Missing [n (%)]	3,516 (1.0)
Marital status	
Married/living together [n (%)]	210,795 (62.0)
Single/divorced/widowed [n (%)]	129,,152 (38.0)
Mother's job status	
Employed [n (%)]	277,110 (81.5)
Unemployed [n (%)]	62,837 (18.5)
COUNTRY/NEIGHBORHOOD LEVEL COVARIATES	
Year of birth	
2009 [n (%)]	19,764 (5.8)
2010 [n (%)]	83,687 (24.6)
2011 [n (%)]	82,032 (24.1)
2012 [n (%)]	79,072 (23.3)
2013 [n (%)]	75,392 (22.2)
Urbanization degree	
1000–1500 addresses per km² [n (%)]	129,313 (38.0)
500–1000 addresses per km² [n (%)]	113,692 (33.4)
≤500 addresses per km² [n (%)]	96,934 (28.5)
Missing [n (%)]	8 (0.0)

n = number of observations, sd = standard deviation

Table 3: Results of the regression analysis on the 12 *a priori* selected active ingredients.

Active ingredient (AI)	Outcome	Buffer size			
		50m	100m	250m	500m
Asulamaa	Gestational age	β = 0.74 (-1.97, 3.46)	β = -0.00 (-0.57, 0.57)	β = 2.2e-03 (-0.09, 0.09)	β = 0.03 (0.01, 0.04)
	Birth weight	β = -94.30 (-176.06, -12.53)	β = -25.20 (-42.41, -7.99)	β = -2.01 (-4.74, 0.71)	β = 0.85 (0.45, 1.24)
	Perinatal mortality	OR = 0.82 (0.01, 52.50)	OR = 0.91 (0.41, 2.03)	OR = 0.95 (0.84, 1.07)	OR = 0.98 (0.96, 1.00)
	Child's sex	OR = 0.74 (0.50, 1.09)	OR = 0.95 (0.87, 1.03)	OR = 1.00 (0.99, 1.01)	OR = 1.00 (1.00, 1.00)
	Prematurity	OR = 0.67 (0.24, 1.86)	OR = 0.96 (0.79, 1.17)	OR = 0.99 (0.96, 1.02)	OR = 1.00 (0.99, 1.00)
	Low birth weight	OR = 1.51 (0.25, 9.04)	OR = 1.15 (0.80, 1.66)	OR = 1.01 (0.95, 1.06)	OR = 0.99 (0.99, 1.00)
	Large for gestational age	OR = 0.57 (0.31, 1.06)	OR = 0.90 (0.80, 1.02)	OR = 0.99 (0.98, 1.01)	OR = 1.00 (1.00, 1.01)
	Small for gestational age	OR = 0.93 (0.41, 2.11)	OR = 1.01 (0.86, 1.19)	OR = 1.00 (0.97, 1.03)	OR = 0.99 (0.99, 1.00)
Carbetamideb	Gestational age	β = 0.68 (-6.15, 7.50)	β = -0.21 (-1.52, 1.09)	β = -0.05 (-0.25, 0.15)	β = 0.02 (-0.02, 0.06)
	Birth weight	β = 40.52 (-164.79, 245.83)	β = -6.48 (-45.71, 32.75)	β = -3.65 (-9.55, 2.25)	β = -1.55 (-2.68, -0.41)
	Perinatal mortality	OR = 0.04 (9.8e-06, 199.06)	OR = 0.53 (0.11, 2.43)	OR = 0.82 (0.65, 1.02)	OR = 0.98 (0.94, 1.02)
	Child's sex	OR = 1.15 (0.44, 3.05)	OR = 1.10 (0.91, 1.32)	OR = 1.01 (0.99, 1.04)	OR = 1.00 (0.99, 1.00)
	Prematurity	OR = 1.19 (0.14, 10.06)	OR = 1.20 (0.81, 1.77)	OR = 1.04 (0.98, 1.10)	OR = 1.00 (0.99, 1.01)
	Low birth weight	OR = 1.45 (0.04, 50.60)	OR = 0.96 (0.49, 1.88)	OR = 0.94 (0.85, 1.05)	OR = 0.99 (0.97, 1.01)
	Large for gestational age	OR = 0.97 (0.22, 4.33)	OR = 0.94 (0.70, 1.26)	OR = 0.98 (0.93, 1.02)	OR = 0.99 (0.98, 1.00)
	Small for gestational age	OR = 1.81 (0.35, 9.27)	OR = 1.20 (0.88, 1.64)	OR = 1.02 (0.97, 1.07)	OR = 1.00 (0.99, 1.01)
Cyproconazoleb	Gestational age	β = -22.97 (-280.15, 234.21)	β = 17.86 (-30.49, 66.21)	β = 0.66 (-6.56, 7.87)	β = -1.84 (-3.57, -0.11)
	Birth weight	β = 4.3e+03 (-3447.78, 1.2e+04)	β = 1.5e+03 (29.63, 2.9e+03)	β = 394.08 (176.91, 611.26)	β = -46.49 (-98.66, 5.68)
	Perinatal mortality	OR = 2.0e-81 (7.4e-206, 5.4e+43)	OR = 2.0e-14 (2.1e-36, 2.0e+08)	OR = 0.05 (3.5e-05, 82.77)	OR = 0.27 (0.05, 1.49)
	Child's sex	OR = 191.80 (2.5e-14, 1.5e+18)	OR = 0.45 (4.6e-04, 437.60)	OR = 1.03 (0.37, 2.88)	OR = 1.00 (0.78, 1.27)
	Prematurity	OR = 6.5e+19 (1.2e-13, 3.4e+52)	OR = 5.42 (2.6e-06, 1.1e+07)	OR = 0.79 (0.09, 7.17)	OR = 1.05 (0.62, 1.75)
	Low birth weight	OR = 5.1e+05 (4.8e-51, 5.3e+61)	OR = 7.7e+03 (1.5e-07, 3.9e+14)	OR = 0.08 (1.7e-03, 3.82)	OR = 0.87 (0.37, 2.04)
	Large for gestational age	OR = 8.0e+20 (1.7e-04, 3.8e+45)	OR = 173.54 (2.8e-03, 1.1e+07)	OR = 7.03 (1.38, 35.85)	OR = 0.69 (0.46, 1.05)
	Small for gestational age	OR = 4.1e-10 (2.9e-39, 5.7e+19)	OR = 2.6e-06 (6.3e-12, 1.06)	OR = 0.08 (0.01, 0.51)	OR = 0.82 (0.54, 1.24)

Active ingredient (AI)	Outcome	Buffer size			
		50m	100m	250m	500m
Epoxiconazoleb	Gestational age	β = 36.13 (-6.75, 79.01)	β = 2.75 (-5.55, 11.04)	β = 0.72 (-0.54, 1.98)	β = 0.05 (-0.16, 0.26)
	Birth weight	β = 76.02 (-1214.41, 1.4e+03)	β = -47.86 (-297.55, 201.82)	β = -13.85 (-51.81, 24.10)	β = -5.61 (-11.91, 0.69)
	Perinatal mortality	OR = 3.9e+07 (2.0e-09, 7.7e+23)	OR = 39.10 (0.02, 9.3e+04)	OR = 1.10 (0.32, 3.80)	OR = 1.09 (0.89, 1.33)
	Child's sex	OR = 0.45 (1.0e-03, 200.74)	OR = 1.11 (0.34, 3.60)	OR = 0.99 (0.83, 1.18)	OR = 1.00 (0.97, 1.03)
	Prematurity	OR = 1.8e-04 (1.5e-10, 226.44)	OR = 0.67 (0.05, 8.90)	OR = 0.84 (0.57, 1.24)	OR = 1.03 (0.97, 1.09)
	Low birth weight	OR = 0.43 (2.0e-11, 9.3e+09)	OR = 0.24 (2.5e-03, 22.40)	OR = 0.63 (0.32, 1.25)	OR = 0.95 (0.85, 1.06)
	Large for gestational age	OR = 0.04 (1.7e-06, 975.37)	OR = 0.72 (0.11, 4.94)	OR = 1.04 (0.78, 1.39)	OR = 1.00 (0.95, 1.05)
	Small for gestational age	OR = 0.02 (4.2e-07, 716.07)	OR = 0.74 (0.09, 5.80)	OR = 1.02 (0.75, 1.38)	OR = 1.06 (1.01, 1.11)
Fluroxypyr_ meptylb	Gestational age	β = 9.48 (-10.99, 29.95)	β = 4.96 (0.89, 9.03)	β = 0.87 (0.22, 1.52)	β = 0.06 (-0.06, 0.19)
	Birth weight	β = 445.85 (-170.30, 1.1e+03)	β = 142.72 (20.09, 265.36)	β = 36.19 (16.59, 55.79)	β = -1.48 (-5.28, 2.31)
	Perinatal mortality	OR = 2.39 (4.4e-09, 1.3e+09)	OR = 0.20 (2.9e-03, 13.45)	OR = 0.91 (0.47, 1.76)	OR = 1.03 (0.91, 1.16)
	Child's sex	OR = 0.17 (9.1e-03, 3.11)	OR = 0.66 (0.37, 1.18)	OR = 0.98 (0.89, 1.07)	OR = 1.00 (0.99, 1.02)
	Prematurity	OR = 0.23 (3.6e-04, 143.16)	OR = 0.45 (0.12, 1.61)	OR = 0.84 (0.68, 1.02)	OR = 0.96 (0.93, 1.00)
	Low birth weight	OR = 0.07 (1.7e-06, 2.8e+03)	OR = 0.36 (0.04, 2.94)	OR = 0.89 (0.63, 1.24)	OR = 1.01 (0.95, 1.08)
	Large for gestational age	OR = 1.76 (0.02, 200.58)	OR = 1.37 (0.54, 3.49)	OR = 1.10 (0.95, 1.28)	OR = 0.98 (0.95, 1.01)
	Small for gestational age	OR = 1.60 (0.01, 254.46)	OR = 0.56 (0.20, 1.56)	OR = 0.88 (0.75, 1.03)	OR = 1.00 (0.97, 1.03)
Glufosinateb	Gestational age	β = -2.92 (-36.39, 30.54)	β = 2.47 (-4.83, 9.77)	β = 0.08 (-1.12, 1.27)	β = 0.31 (0.09, 0.53)
	Birth weight	β = -360.66 (-1368.05, 646.73)	β = -16.49 (-236.35, 203.38)	β = -9.25 (-45.17, 26.66)	β = -2.97 (-9.51, 3.58)
	Perinatal mortality	OR = 1.8e+13 (3.9e+03, 8.2e+22)	OR = 1.1e+03 (3.99, 3.2e+05)	OR = 1.20 (0.34, 4.16)	OR = 0.93 (0.73, 1.19)
	Child's sex	OR = 1.64 (0.01, 192.22)	OR = 0.75 (0.27, 2.13)	OR = 0.82 (0.69, 0.97)	OR = 0.99 (0.96, 1.02)
	Prematurity	OR = 1.62 (9.8e-06, 2.7e+05)	OR = 0.25 (0.02, 3.56)	OR = 0.83 (0.56, 1.21)	OR = 0.90 (0.84, 0.97)
	Low birth weight	OR = 2.4e-13 (4.9e-25, 0.12)	OR = 1.6e-03 (1.2e-05, 0.22)	OR = 0.87 (0.47, 1.61)	OR = 1.00 (0.90, 1.12)
	Large for gestational age	OR = 2.6e-03 (6.8e-07, 9.75)	OR = 0.41 (0.07, 2.30)	OR = 0.96 (0.73, 1.27)	OR = 0.96 (0.91, 1.01)
	Small for gestational age	OR = 2.19 (5.0e-04, 9.6e+03)	OR = 1.10 (0.18, 6.64)	OR = 1.12 (0.84, 1.48)	OR = 1.02 (0.96, 1.07)

→

Table 3: Results of the regression analysis on the 12 a priori selected active ingredients. (cont.)

Active ingredient (AI)	Outcome	Buffer size			
		50m	100m	250m	500m
Glufosinate_ ammoniumc	Gestational age	β = 11.58 (-5.61, 28.77)	β = 1.52 (-2.24, 5.28)	β = -0.16 (-0.79, 0.47)	β = 0.23 (0.12, 0.34)
	Birth weight	β = 165.48 (-351.80, 682.77)	β = 38.72 (-74.44, 151.88)	β = -12.70 (-31.66, 6.26)	β = -5.57 (-8.94, -2.19)
	Perinatal mortality	OR = 0.44 (4.5e-08, 4.3e+06)	OR = 0.90 (0.02, 46.39)	OR = 0.91 (0.48, 1.72)	OR = 0.91 (0.80, 1.04)
	Child's sex	OR = 1.04 (0.09, 12.04)	OR = 0.88 (0.52, 1.51)	OR = 1.06 (0.97, 1.16)	OR = 1.00 (0.98, 1.01)
	Prematurity	OR = 5.8e-03 (8.5e-06, 4.02)	OR = 0.47 (0.13, 1.75)	OR = 1.01 (0.83, 1.23)	OR = 0.98 (0.94, 1.01)
	Low birth weight	OR = 4.1e+04 (8.54, 1.9e+08)	OR = 18.83 (3.19, 111.33)	OR = 1.16 (0.85, 1.59)	OR = 1.05 (1.00, 1.11)
	Large for gestational age	OR = 6.91 (0.17, 287.36)	OR = 1.48 (0.64, 3.42)	OR = 0.95 (0.82, 1.10)	OR = 0.97 (0.95, 1.00)
	Small for gestational age	OR = 0.03 (2.5e-04, 2.84)	OR = 0.48 (0.18, 1.29)	OR = 1.06 (0.91, 1.24)	OR = 1.01 (0.99, 1.04)
Linuronb	Gestational age	β = 4.97 (-2.72, 12.67)	β = 0.99 (-0.50, 2.48)	β = 0.16 (-0.06, 0.39)	β = -0.09 (-0.13, -0.05)
	Birth weight	β = 256.85 (25.27, 488.43)	β = 66.05 (21.12, 110.98)	β = 13.24 (6.43, 20.04)	β = 2.54 (1.25, 3.84)
	Perinatal mortality	OR = 1.61 (1.1e-03, 2.4e+03)	OR = 1.06 (0.27, 4.19)	OR = 1.03 (0.84, 1.27)	OR = 1.00 (0.96, 1.04)
	Child's sex	OR = 0.66 (0.22, 1.99)	OR = 0.90 (0.73, 1.12)	OR = 1.00 (0.96, 1.03)	OR = 1.00 (1.00, 1.01)
	Prematurity	OR = 0.41 (0.04, 4.53)	OR = 0.85 (0.54, 1.34)	OR = 0.98 (0.91, 1.04)	OR = 1.01 (1.00, 1.02)
	Low birth weight	OR = 0.04 (5.1e-04, 3.02)	OR = 0.60 (0.27, 1.35)	OR = 0.95 (0.85, 1.07)	OR = 1.00 (0.98, 1.02)
	Large for gestational age	OR = 4.19 (0.80, 22.13)	OR = 1.43 (1.03, 1.98)	OR = 1.07 (1.02, 1.13)	OR = 1.01 (1.00, 1.02)
	Small for gestational age	OR = 0.95 (0.14, 6.46)	OR = 0.94 (0.64, 1.36)	OR = 0.95 (0.90, 1.00)	OR = 0.99 (0.98, 1.01)
Propiconazoleb	Gestational age	β = 4.60 (-15.38, 24.59)	β = 1.59 (-2.81, 5.98)	β = 0.11 (-0.63, 0.84)	β = 0.06 (-0.08, 0.19)
	Birth weight	β = 207.34 (-394.19, 808.87)	β = 31.87 (-100.39, 164.12)	β = -17.94 (-40.18, 4.29)	β = -5.61 (-9.72, -1.50)
	Perinatal mortality	OR = 22.65 (2.4e-08, 2.1e+10)	OR = 17.32 (0.24, 1.2e+03)	OR = 1.86 (0.89, 3.89)	OR = 0.90 (0.77, 1.06)
	Child's sex	OR = 3.79 (0.22, 65.29)	OR = 1.29 (0.69, 2.42)	OR = 1.12 (1.01, 1.25)	OR = 0.99 (0.97, 1.01)
	Prematurity	OR = 0.12 (7.0e-05, 222.08)	OR = 0.51 (0.11, 2.26)	OR = 0.97 (0.77, 1.21)	OR = 1.00 (0.96, 1.04)
	Low birth weight	OR = 1.3e+04 (0.15, 1.1e+09)	OR = 3.26 (0.37, 28.38)	OR = 0.90 (0.63, 1.30)	OR = 1.03 (0.97, 1.10)
	Large for gestational age	OR = 11.76 (0.13, 1.0e+03)	OR = 1.31 (0.48, 3.57)	OR = 0.90 (0.76, 1.08)	OR = 0.95 (0.92, 0.98)
	Small for gestational age	OR = 0.18 (8.7e-04, 35.30)	OR = 0.61 (0.20, 1.87)	OR = 1.00 (0.84, 1.19)	OR = 1.00 (0.97, 1.03)

Active ingredient (AI)	Outcome	Buffer size			
		50m	100m	250m	500m
Thiaclopridd	Gestational age	β = 3.91 (-19.16, 26.98)	β = -0.07 (-4.61, 4.47)	β = -0.09 (-0.78, 0.60)	β = 0.15 (0.04, 0.27)
	Birth weight	β = -93.41 (-788.02, 601.20)	β = -39.31 (-176.07, 97.45)	β = -10.28 (-31.03, 10.48)	β = 1.07 (-2.31, 4.45)
	Perinatal mortality	OR = 1.9e-07 (5.1e-21, 7.1e+06)	OR = 0.01 (2.6e-05, 3.86)	OR = 0.39 (0.17, 0.88)	OR = 0.89 (0.78, 1.02)
	Child's sex	OR = 0.17 (6.2e-03, 4.51)	OR = 0.82 (0.43, 1.56)	OR = 1.01 (0.91, 1.11)	OR = 1.00 (0.98, 1.01)
	Prematurity	OR = 0.43 (2.4e-04, 759.17)	OR = 1.48 (0.37, 5.97)	OR = 1.11 (0.90, 1.37)	OR = 0.98 (0.94, 1.01)
	Low birth weight	OR = 0.12 (2.0e-07, 7.5e+04)	OR = 1.14 (0.10, 13.36)	OR = 0.95 (0.66, 1.35)	OR = 1.00 (0.94, 1.05)
	Large for gestational age	OR = 0.13 (7.3e-04, 22.34)	OR = 0.75 (0.27, 2.08)	OR = 0.95 (0.81, 1.11)	OR = 0.99 (0.97, 1.02)
	Small for gestational age	OR = 0.66 (1.7e-03, 254.20)	OR = 1.10 (0.34, 3.50)	OR = 0.96 (0.80, 1.14)	OR = 0.97 (0.94, 0.99)
Triadimenolb	Gestational age	β = 15.92 (-44.33, 76.16)	β = -2.48 (-15.61, 10.66)	β = -1.05 (-3.16, 1.06)	β = 0.37 (-0.01, 0.75)
	Birth weight	β = 103.29 (-1710.01, 1.9e+03)	β = 39.86 (-355.48, 435.20)	β = 17.97 (-45.55, 81.49)	β = -9.01 (-20.41, 2.38)
	Perinatal mortality	OR = 4.9e-10 (2.9e-32, 8.4e+12)	OR = 1.8e-04 (2.2e-09, 15.25)	OR = 0.21 (0.03, 1.67)	OR = 0.87 (0.59, 1.28)
	Child's sex	OR = 0.02 (3.2e-06, 93.12)	OR = 0.35 (0.05, 2.26)	OR = 0.79 (0.59, 1.07)	OR = 1.02 (0.97, 1.08)
	Prematurity	OR = 6.4e-06 (3.8e-15, 1.1e+04)	OR = 0.65 (9.3e-03, 45.30)	OR = 1.24 (0.65, 2.35)	OR = 0.99 (0.88, 1.11)
	Low birth weight	OR = 0.05 (6.0e-20, 4.8e+16)	OR = 190.83 (0.17, 2.1e+05)	OR = 2.45 (0.88, 6.79)	OR = 1.08 (0.89, 1.30)
	Large for gestational age	OR = 0.61 (3.4e-07, 1.1e+06)	OR = 1.47 (0.07, 30.47)	OR = 1.14 (0.70, 1.85)	OR = 1.01 (0.93, 1.10)
	Small for gestational age	OR = 1.3e-03 (3.8e-10, 4.2e+03)	OR = 0.65 (0.03, 14.89)	OR = 1.37 (0.84, 2.23)	OR = 1.11 (1.02, 1.21)
Vinclozolinb	Gestational age	β = 486.69 (-560.90, 1.5e+03)	β = 38.81 (-184.54, 262.16)	β = 7.05 (-28.80, 42.90)	β = 8.40 (3.12, 13.67)
	Birth weight	β = -2.7e+04 (-5.9e+04, 4.3e+03)	β = -8.0e+03 (-1.5Ee+04, -1.3e+03)	β = -484.91 (-1.6e+03, 594.32)	β = 257.78 (99.19, 416.37)
	Perinatal mortality	OR = 8.1e+178 (2.8e-287, Inf)	OR = 4.9e+38 (4.2e-70, 5.7e+146)	OR = 5.2e-04 (1.9e-26, 1.4e+19)	OR = 6.1e-05 (3.9e-09, 0.95)
	Child's sex	OR = 6.9e-16 (1.2e-80, 4.1e+49)	OR = 4.03 (6.5e-14, 2.5e+14)	OR = 30.32 (0.18, 5.0e+03)	OR = 1.49 (0.70, 3.15)
	Prematurity	OR = 3.8e-88 (2.8e-292, 5.1e+116)	OR = 1.7e-14 (6.1e-52, 4.9e+23)	OR = 1.4e-03 (4.6e-09, 404.35)	OR = 0.15 (0.02, 0.96)
	Low birth weight	OR = 3.0e+185 (7.8e-87, Inf)	OR = 5.8e+27 (3.0e-37, 1.1e+92)	OR = 3.3e-04 (2.7e-15, 4.1e+07)	OR = 0.02 (8.4e-04, 0.64)
	Large for gestational age	OR = 4.3e-65 (2.8e-164, 6.6e+34)	OR = 7.5e-15 (1.2e-35, 4.6e+06)	OR = 0.09 (5.5e-05, 158.84)	OR = 3.49 (1.23, 9.90)
	Small for gestational age	OR = 1.9e-05 (1.7e-137, 2.0e+127)	OR = 1.9e+03 (2.8e-24, 1.3e+30)	OR = 9.34 (5.8e-04, 1.5e+05)	OR = 0.21 (0.05, 0.96)

The referent ("unexposed") group in all models was mothers with zero kilogram of specific AI within 500m of their residences.

AIs included in the models (assuring that the correlations between the AIs included in the models was <0.70):
a asulam, carbetamide, cyproconazole, epoxiconazole, glufosinate, linuron, fluroxypyr-meptyl, propiconazole and triadimenol
b carbetamide, cyproconazole, epoxiconazole, glufosinate, linuron, fluroxypyr-meptyl, propiconazole, triadimenol and vinclozolin
c carbetamide, cyproconazole, epoxiconazole, glufosinate, linuron, fluroxypyr_meptyl, glufosinate-ammonium and vinclozolin
d cyproconazole, epoxiconazole, glufosinate, linuron, fluroxypyr-meptyl, propiconazole, thiacloprid, triadimenol and vinclozolin

All models were further adjusted for the amount of AI used in the area up to 500m ("complimentary donut"), gestational age (when gestational age or prematurity were not the outcomes), sex of the baby (when child's sex was not the outcome), parity, household income, mother's education, mother's marital status, mother's origin, mother's age at birth, mother's job status, urbanization degree and year of birth.
NA = not computed because there were <10 exposed cases.

Table 4: Results of the variable selection models.

Pool of AIs	Buffer	Gestational age[a]	Birth weight[b]	Child's sex	Perinatal mortality[c]
azoxystrobin, boscalid, clopyralid, cymoxanil, dimethoate, diquat, phenmedipham, fenpropimorph, flonicamid, fluroxypyr, fosetyl, haloxyfop_p-methyl, imidacloprid, copper oxychloride, kresoxim_methyl, mancozeb, mcpa, mecoprop_p, metamitron, metiram, metribuzin, metsulfuron_methyl, pirimicarb, prosulfocarb, pyraclostrobin, quinoclamine, rimsulfuron, spinosad, sulcotrione, tebufenpyrad, teflubenzuron, thiophanate_methyl, tolclofos_methyl, topramezone, trifloxystrobin, triflusulfuron_methyl, aclonifen, emamectin benzoate, fenamidone, fenpropidin, penconazole, pyroxsulam, tembotrione, thiamethoxam, thiram, sulphur	50m	diquat emamectin benzoate	diquat		diquat picoxystrobin
	100m		copper oxychloride		picoxystrobin
	250m		fenpropidin	diquat tebufenpyrad	picoxystrobin
	500m	rimsulfuron, sulcotrione	tolclofos-methyl		sulcotrione

Pool of AIs	Buffer	Prematurity	Low birth weight[d]	Large for gestational age[b]	Small for gestational age
azoxystrobin, boscalid, clopyralid, cymoxanil, dimethoate, diquat, phenmedipham, fenpropimorph, flonicamid, fluroxypyr, fosetyl, haloxyfop_p-methyl, imidacloprid, copper oxychloride, kresoxim_methyl, mancozeb, mcpa, mecoprop_p, metamitron, metiram, metribuzin, metsulfuron_methyl, pirimicarb, prosulfocarb, pyraclostrobin, quinoclamine, rimsulfuron, spinosad, sulcotrione, tebufenpyrad, teflubenzuron, thiophanate_methyl, tolclofos_methyl, topramezone, trifloxystrobin, triflusulfuron_methyl, aclonifen, emamectin benzoate, fenamidone, fenpropidin, penconazole, pyroxsulam, tembotrione, thiamethoxam, thiram, sulphur	50m		diquat metiram picoxystrobin	picoxystrobin	diquat
	100m		tebufenpyrad	picoxystrobin	copper oxychloride
	250m	triflusulfuron-methyl pyroxsulam		picoxystrobin, fenpropidin	triflusulfuron-methyl
	500m	rimsulfuron	sulcotrione		prosulfocarb

All models were adjusted for gestational age (when gestational age or prematurity were not the outcomes), sex of the baby (when child's sex was not the outcome), parity, household income, mother's education, mother's marital status, mother's origin, mother's age at birth, mother's job status, urbanization degree and year of birth. Models were further adjusted for the amount of AIs used in the area up to 500m ("complimentary donut"). Finally, models were adjusted for AIs (and their complementary donuts) for which evidence for an association was determined in the previous (linear regression) analyses:
a fluroxypyr-meptyl and vinclozolin,
b linuron,
c thiachloprid and
d glufosinate-ammonium

Table 5: Results of the regression analyses on the associations between picoxystrobin and the birth outcomes large for gestational age and perinatal mortality.

Active ingredient (AI)	Buffer size	Large for gestational age	Perinatal mortality
Picoxystrobin	50m	OR = 1.4e+54 (3.0e-05, 6.7e+112)	OR = 1.3e+30 (6.1e-171, 3.0e+230)
	100m	OR = 2.7e+19 (1.8e+08, 4.2e+30)	OR = 6.7e+05 (5.5e-38, 8.1e+48)
	250m	OR = 325.52 (3.62, 2.9e+04)	OR = 1.0e+06 (0.12, 8.1e+12)
	500m	OR = 2.03 (0.74, 5.55)	OR = 4.57 (0.07, 292.54)

Table 6: Results of the sensitivity analyses performed on the associations that were considered findings.

Analysis		50m	100m	250m	500m
Fluroxypyr-meptyl ~ Gestational age	Main analysis	β = 9.48 (-10.99, 29.95)	β = 4.96 (0.89, 9.03)	β = 0.87 (0.22, 1.52)	β = 0.06 (-0.06, 0.19)
	(2) Excl. mothers in Agric.	β = 11.45 (-9.31, 32.22)	β = 4.36 (0.24, 8.49)	β = 0.59 (-0.07, 1.24)	β = 0.04 (-0.09, 0.16)
	(3) Excl. mothers & fathers in Agric.	β = 9.51 (-13.12, 32.14)	β = 3.34 (-1.14, 7.82)	β = 0.35 (-0.35, 1.05)	β = 2.5e-03 (-0.13, 0.14)
	(4) Rural areas	β = 7.91 (-12.99, 28.82)	β = 4.75 (0.56, 8.93)	β = 0.96 (0.28, 1.64)	β = 0.10 (-0.04, 0.23)
	(5) Autochthonous mothers	β = 16.73 (-4.40, 37.86)	β = 5.68 (1.47, 9.90)	β = 0.81 (0.14, 1.49)	β = 0.06 (-0.07, 0.20)
	(6) Complete case analysis	β = 6.99 (-17.62, 31.60)	β = 3.41 (-1.44, 8.25)	β = 0.50 (-0.27, 1.27)	β = 2.1e-03 (-0.15, 0.15)
	(7) No address change	β = 8.30 (-13.06, 29.66)	β = 4.10 (-0.14, 8.33)	β = 0.88 (0.21, 1.56)	β = 0.08 (-0.05, 0.21)
Glufosinate-ammonium ~ Low birth weight	Main analysis	OR=4.1e+04 (8.54, 1.9e+08)	OR = 18.83 (3.19, 111.33)	OR = 1.16 (0.85, 1.59)	OR = 1.05 (1.00, 1.11)
	(1) Without GA	OR=3.21 (4.2e-03, 2.47e+03)	OR = 2.26 (0.61, 8.46)	OR = 1.12 (0.89, 1.38)	OR = 1.01 (0.97, 1.05)
	(2) Excl. mothers in Agric.	OR=4.7e+04 (5.70, 4.0e+08)	OR = 20.88 (3.24, 134.65)	OR = 1.14 (0.82, 1.59)	OR = 1.06 (1.00, 1.12)
	(3) Excl. mothers & fathers in Agric.	OR=2.1e+06 (85.82, 5.2e+10)	OR = 34.82 (4.27, 284.15)	OR = 1.14 (0.79, 1.63)	OR = 1.05 (0.98, 1.11)
	(4) Rural areas	OR=1.3e+04 (2.04, 8.1e+07)	OR = 16.39 (2.61, 103.16)	OR = 1.12 (0.81, 1.56)	OR = 1.06 (1.00, 1.12)
	(5) Autochthonous mothers	OR=1.9e+04 (2.42, 1.6e+08)	OR = 16.50 (2.56, 106.49)	OR = 1.16 (0.84, 1.62)	OR = 1.06 (1.00, 1.12)
	(6) Complete case analysis	OR=5.0e+05 (4.17, 1.4e+10)	OR = 45.20 (4.44, 377.76)	OR = 1.20 (0.80, 1.75)	OR = 1.01 (0.94, 1.08)
	(7) No address change	OR=2.1e+04 (3.33, 1.3e+08)	OR = 13.31 (1.97, 89.84)	OR = 1.07 (0.77, 1.49)	OR = 1.05 (0.99, 1.11)
Linuron ~ Birth weight	Main analysis	β = 256.85 (25.27, 488.43)	β = 66.05 (21.12, 110.98)	β = 13.24 (6.43, 20.04)	β = 2.54 (1.25, 3.84)
	(1) Without GA	β = 393.21 (79.97, 706.45)	β = 93.16 (32.39, 153.95)	β = 17.70 (8.50, 26.91)	β = 0.10 (-1.65, 1.85)
	(2) Excl. mothers in Agric.	β = 249.22 (6.51, 491.92)	β = 58.23 (11.19, 105.28)	β = 11.25 (4.18, 18.32)	β = 2.80 (1.45, 4.14)
	(3) Excl. mothers & fathers in Agric.	β = 196.67 (-70.64, 463.98)	β = 55.58 (4.09, 107.07)	β = 10.99 (3.40, 18.58)	β = 2.67 (1.25, 4.09)
	(4) Rural areas	β = 241.37 (6.84, 475.90)	β = 67.66 (21.90, 113.42)	β = 13.87 (6.85, 20.89)	β = 2.72 (1.37, 4.07)
	(5) Autochthonous mothers	β = 280.80 (40.87, 520.74)	β = 69.05 (22.31, 115.79)	β = 12.94 (5.82, 20.06)	β = 2.75 (1.39, 4.11)
	(6) Complete case analysis	β = 317.46 (45.82, 589.11)	β = 62.01 (9.14, 114.89)	β = 14.51 (6.58, 22.43)	β = 2.74 (1.25, 4.24)
	(7) No address change	β = 185.05 (-55.77, 425.87)	β = 56.02 (9.42, 102.62)	β = 12.12 (5.07, 19.16)	β = 2.45 (1.10, 3.79)

→

Table 6: Results of the sensitivity analyses performed on the associations that were considered findings. (cont.)

Analysis		50m	100m	250m	500m
Linuron ~ Large for gestational age	Main analysis	OR = 4.19 (0.80, 22.13)	OR = 1.43 (1.03, 1.98)	OR = 1.07 (1.02, 1.13)	OR = 1.01 (1.00, 1.02)
	(1) Without GA	OR = 4.23 (0. 08, 22.32)	OR = 1.43 (1.03, 1.98)	OR = 1.07 (1.02, 1.13)	OR = 1.01 (1.00, 1.02)
	(2) Excl. mothers in Agric.	OR = 2.94 (0.50, 17.17)	OR = 1.36 (0.96, 1.92)	OR = 1.06 (1.01, 1.12)	OR = 1.02 (1.01, 1.03)
	(3) Excl. mothers & fathers in Agric.	OR = 1.58 (0.21, 11.66)	OR = 1.25 (0.85, 1.84)	OR = 1.05 (0.99, 1.12)	OR = 1.02 (1.01, 1.03)
	(4) Rural areas	OR = 4.36 (0.82, 23.33)	OR = 1.45 (1.04, 2.02)	OR = 1.07 (1.02, 1.13)	OR = 1.02 (1.01, 1.03)
	(5) Autochthonous mothers	OR = 4.34 (0.77, 24.35)	OR = 1.48 (1.05, 2.07)	OR = 1.08 (1.02, 1.14)	OR = 1.01 (1.00, 1.03)
	(6) Complete case analysis	OR = 3.77 (0.50, 25.90)	OR = 1.32 (0.89, 1.94)	OR = 1.08 (1.02, 1.15)	OR = 1.02 (1.00, 1.03)
	(7) No address change	OR = 3.02 (0.53, 17.25)	OR = 1.35 (0.96, 1.90)	OR = 1.07 (1.01, 1.13)	OR = 1.01 (1.00, 1.02)
Picox-ystrobin ~ Large for gestational age	Main analysis	OR = 1.4e+54 (3.0e-05, 6.7e+112)	OR = 2.7e+19 (1.8e+08, 4.2e+30)	OR = 325.52 (3.62, 2.9e+04)	OR = 2.03 (0.74, 5.55)
	(1) Without GA	OR = 4.1e+53 (9.2e-06, 1.9e+112)	OR = 2.0e+19 (1.3e+08, 3.0e+30)	OR = 3.1e02 (3.55, 2.8e+04)	OR = 2.02 (0.74, 5.53)
	(2) Excl. mothers in Agric.	OR = 5.3e+65 (7.2e+06, 4.0e+124)	OR = 5.4e+20 (1.9e+09, 1.5e+32)	OR = 478.86 (4.60, 5.0e+04)	OR = 2.20 (0.77, 6.29)
	(3) Excl. mothers & fathers in Agric.	OR = 8.8e+61 (1.1e-04, 7.1e+127)	OR = 5.1e+21 (1.4e+09, 1.9e+34)	OR = 427.58 (2.98, 6.1e+04)	OR = 1.58 (0.51, 4.85)
	(4) Rural areas	OR = 5.6e+58 (0.08, 4.0e+118)	OR = 4.0e+18 (1.0e+07, 1.6e+30)	OR = 159.04 (1.41, 1.8e+04)	OR = 1.83 (0.64, 5.22)
	(5) Autochthonous mothers	OR = 5.4e+50 (2.1e-11, 1.4e+112)	OR = 1.3e+19 (3.1e+07, 5.6e+30)	OR = 914.48 (8.35, 1.0e+05)	OR = 2.08 (0.71, 6.06)
	(6) Complete case analysis	OR = 1.2e+109 (1.4e+35, 5.6e+177)	OR = 4.2e+31 (3.4e+17, 1.4e+45)	OR = 7.0e+03 (26.14, 1.5e+06)	OR = 1.62 (0.45, 5.58)
	(7) No address change	OR = 7.6e+56 (9.8e-04, 5.9e+116)	OR = 5.7e+19 (2.0e+08, 1.6e+31)	OR = 404.15 (4.04, 4.0e+04)	OR = 2.07 (0.73, 5.85)
Thiacloprid ~ Perinatal mortality	Main analysis	OR = 1.9e-07 (5.1e-21, 7.1e+06)	OR = 0.01 (2.6e-05, 3.86)	OR = 0.39 (0.17, 0.88)	OR = 0.89 (0.78, 1.02)
	(1) Without GA	OR = 1.5e-12 (1.4e-36, 1.5e12)	OR = 2.2e-05 (9.5e-11, 4.90)	OR = 0.42 (0.13, 1.35)	OR = 0.89 (0.77, 1.03)
	(2) Excl. mothers in Agric.	OR = 1.6e-08 (4.0e-32, 5.92e15)	OR = 1.3e-05 (4.8e-12, 35.99)	OR = 0.24 (0.05, 1.19)	OR = 0.89 (0.72, 1.10)
	(3) Excl. mothers & fathers in Agric.	OR = 2.02e-05 (3.9e-29, 1.1e19)	OR = 1.5e-04 (20e-11, 1.1e03)	OR = 0.32 (0.06, 1.80)	OR = 0.90 (0.73, 1.12)
	(4) Rural areas	OR = 1.5e-08 (1.7e-31, 1.3e15)	OR = 7.2e-06 (7.0e-12, 7.46)	OR = 0.23 (0.05, 1.05)	OR = 0.91 (0.73, 1.12)
	(5) Autochthonous mothers	OR = 1.6e-04 (9.7e-26, 2.6e17)	OR = 1.2e-05 (4.6e-12, 31.59)	OR = 0.27 (0.06, 1.26)	OR = 0.89 (0.71, 1.09)
	(6) Complete case analysis	OR = 1.3e-08 (4.5e-40, 1.0e+10)	OR = 1.9e-04 (6.8e-12, 20.43)	OR = 0.31 (0.05, 1.32)	OR = 0.90 (0.69, 1.10)
	(7) No address change	OR=5.6e-07 (3.6e-28, 8.2e14)	OR=5.6e-04 (6.1e-09, 51.40)	OR = 0.41 (0.11, 1.51)	OR = 0.92 (0.76, 1.11)
Vinclozolin ~ Gestational age	Main analysis	β = 486.69 (-560.90, 1.5e+03)	β = 38.81 (-184.54, 262.16)	β = 7.05 (-28.80, 42.90)	β = 8.40 (3.12, 13.67)
	(2) Excl. mothers in Agric.	β = 435.12 (-645.44, 1.5e+03)	β = 50.75 (-177.84, 279.33)	β = 9.01 (-27.36, 45.37)	β = 7.15 (1.72, 12.58)
	(3) Excl. mothers & fathers in Agric.	β = 496.50 (-715.14, 1.7e+03)	β = 51.89 (-203.85, 307.64)	β = 5.87 (-34.23, 45.98)	β = 8.63 (2.70, 14.57)
	(4) Rural areas	β = 457.71 (-653.97, 1.6e+03)	β = 7.80 (-231.16, 246.77)	β = -7.98 (-47.98, 32.02)	β = 8.21 (2.14, 14.28)
	(5) Autochthonous mothers	β = 448.03 (-624.99, 1.5e+03)	β = 42.62 (-186.51, 271.75)	β = 10.17 (-27.11, 47.46)	β = 8.38 (2.82, 13.94)
	(6) Complete case analysis	β = 98.17 (-1254.22, 1.5e+03)	β = -59.18 (-338.74, 220.39)	β = 4.19 (-39.82, 48.20)	β = 9.62 (3.06, 16.17)
	(7) No address change	β = 595.72 (-473.14, 1.7e+03)	β = 50.02 (-180.01, 280.06)	β = 6.30 (-30.91, 43.51)	β = 6.40 (0.90, 11.91)

(1) Analysis with no adjustment for gestational age (GA) (N=339947)
(2) Analysis excluding mothers that worked in agriculture during pregnancy (N=337025)
(3) Analysis excluding mothers and fathers that worked in agriculture during pregnancy (N=327654)
(4) Analysis restricted to the most rural areas of the Netherlands (<1000 addresses per km2) (N=210632)
(5) Analysis restricted to autochthonous mothers (N=288446)
(6) Complete case analysis (N=)
(7) Analysis restricted to mothers who did not change address during pregnancy (N=309527)

Table 7: Results of the additional analysis performed for the findings pertaining AI with reported sex-differential effects.

Finding	Buffer Size	Model with interaction term between sex and the AI			Stratified analysis by sex	
		Sex	AI	Sex: AI	AI (boys)	AI (girls)
Linuron ~ Birth weight	50m	-142.52 (-145.38, -139.65)	61.68 (-218.36, 341.71)	402.32 (77.58, 727.06)	-50.99 (-379.35, 277.38)	575.22 (248.55, 901.89)
	100m	-142.64 (-145.52, -139.76)	32.51 (-20.87, 85.9)	69.40 (9.72, 129.07)	14.18 (-49.42, 77.79)	118.81 (55.32, 182.30)
	250m	-142.67 (-145.62, -139.72)	10.67 (2.93, 18.4)	5.17 (-2.23, 12.57)	10.88 (1.16, 20.61)	15.40 (5.89, 24.92)
	500m	-142.97 (-146.04, -139.89)	2.00 (0.48, 3.51)	1.13 (-0.50, 2.75)	2.64 (0.81, 4.48)	2.43 (0.61, 4.25)
Linuron ~ Large for gestational age	50m	1.00 (0.98, 1.03)	1.26 (0.15, 10.21)	10.58 (1.01, 110.61)	0.58 (0.05, 7.10)	26.57 (2.83, 249.74)
	100m	1.00 (0.98, 1.03)	1.20 (0.80, 1.78)	1.43 (0.92, 2.20)	1.14 (0.72, 1.83)	1.78 (1.13, 2.81)
	250m	1.00 (0.98, 1.03)	1.06 (1.00, 1.13)	1.01 (0.96, 1.07)	1.08 (1.00, 1.16)	1.06 (0.99, 1.14)
	500m	1.00 (0.98, 1.03)	1.01 (1.00, 1.03)	1.00 (0.99, 1.01)	1.01 (1.00, 1.03)	1.02 (1.00, 1.03)
Thiacloprid ~ Perinatal mortality	50m	0.97 (0.88, 1.07)	0.00 (0.00, 1.20E+06)	4.56E+13 (0.00, 7.24E+35)	0.00 (0.00, 4.97E+06)	0.02 (0.00, 3.93258E+14)
	100m	0.97 (0.88, 1.07)	0.00 (0.00, 3.38)	34.19 (0.01, 131001.18)	0.00 (0.00, 3.33)	0.64 (0.00, 1895.86)
	250m	0.97 (0.89, 1.07)	0.38 (0.15, 0.97)	1.01 (0.38, 2.69)	0.37 (0.12, 1.11)	0.43 (0.12, 1.50)
	500m	0.98 (0.89, 1.08)	0.91 (0.78, 1.06)	0.95 (0.78, 1.15)	0.87 (0.72, 1.04)	0.92 (0.76, 1.12)
Vinclozolin ~ Gestational age	50m	0.58 (0.48, 0.67)	290.79 (-1032.48, 1614.06)	436.18 (-1367.06, 2239.42)	304.43 (-1150.26, 1759.12)	728.18 (-786.37, 2242.74)
	100m	0.58 (0.48, 0.67)	-20.5 (-298.45, 257.46)	127.29 (-228.75, 483.32)	-19.31 (-332.91, 294.29)	109.35 (-208.87, 427.57)
	250m	0.58 (0.48, 0.67)	3.76 (-38.65, 46.17)	6.64 (-39.11, 52.39)	6.04 (-44.82, 56.91)	8.03 (-42.43, 58.49)
	500m	0.58 (0.48, 0.67)	7.33 (-0.03, 14.69)	2.11 (-8.04, 12.27)	6.66 (-0.94, 14.27)	10.18 (2.87, 17.48)

Table 8: Additional information on the AIs for which evidence of associations with birth outcomes were identified.

Active ingredient	Observed association	Pesticide type	Chemical group	Mode of action
fluroxypyr-meptyl	longer GA	Herbicide	pyridine compound	Root Growth Inhibitor
glufosinate-ammonium	increased odds of LBW	Herbicide	phosphinic acid	Glutamine Synthesis Inhibitor
linuron	higher BW and increased odds of LGA	Herbicide	urea	Photosynthesis Inhibitor (Photosystem II)
vinclozolin	longer GA	Fungicide	dicarboximides	Osmotic signal transduction, prevents spore germination and mycelial growth
picoxystrobin	higher odds of LGA	Fungicide	methoxy-acrylates	Mitochondrial respiration inhibition

Active ingredient	Reproductive/ /Developmental effects	Toxicologic/ /experimental studies	Current approval in EU	Correlation >0.80 with other AIs
fluroxypyr-meptyl	Category1B	classified on PPDB as "known to cause a problem", but EFSA's peer-review on its parent compound (fluroypyr) reports no evidence of reproductive or developmental effects	approved	florasulam
glufosinate-ammonium	Category1B	lower fertility, abortions, fetus death and premature deliveries	not approved	fenamidone, fosetyl-al, acetamiprid, propiconazole, penconazole, quinoclamine
linuron	Category1B[1]	endocrine disruptor, with antiandrogenic effects - impaired in male reproductive development	not approved	metribuzin
vinclozolin	Category1B	endocrine disruptor, with antiandrogenic effects - inhibited male sex differentiation	not approved	tolclofos-methyl haloxyfop-p-methyl
picoxystrobin	none	–	not approved	none

[1] Category1B (EU Pesticides Database) = Presumed human reproductive toxicant (evidence from animal studies).

FIGURES

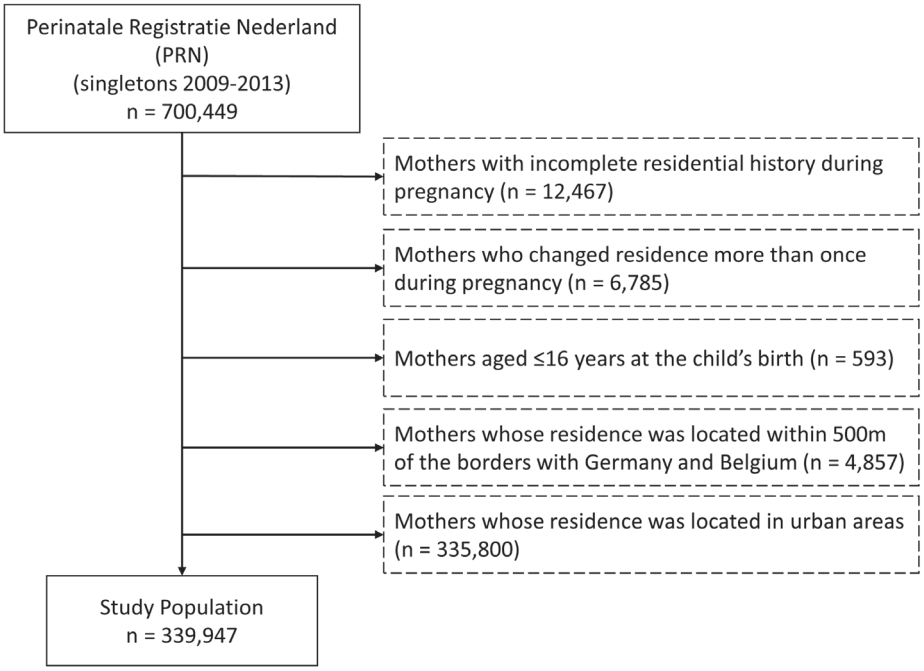


Figure 1: Flow chart of the study population.

REFERENCES

1. Mostafalou S, Abdollahi M. Pesticides: an update of human exposure and toxicity. Arch Toxicol [Internet]. 2017;91(2):549–99. Available from: https://doi.org/10.1007/s00204–016–1849-x
2. Stillerman KP, Mattison DR, Giudice LC, Woodruff TJ. Environmental exposures and adverse pregnancy outcomes: a review of the science. Reprod Sci. 2008 Sep;15(7):631–50.
3. Hu P, Vinturache A, Li H, Tian Y, Yuan L, Cai C, et al. Urinary Organophosphate Metabolite Concentrations and Pregnancy Outcomes among Women Conceiving through in Vitro Fertilization in Shanghai, China. Environ Health Perspect. 2020 Sep;128(9):97007.
4. Coronado GD, Holte S, Vigoren E, Griffith WC, Faustman E, Thompson B. Organophosphate Pesticide Exposure and Residential Proximity to Nearby Fields: Evidence for the Drift Pathway. J Occup Env Med. 2011;53(8):884–91.
5. Ward MH, Lubin J, Giglierano J, Colt JS, Wolter C, Bekiroglu N, et al. Proximity to crops and residential to agricultural herbicides in Iowa. Environ Health Perspect. 2006;114(6):893–7.
6. Gooijer YM, Hoftijser GW, Lageschaar LCC, Figueiredo DM, Huss A, Krop EJM, et al. Research on exposure of residents to pesticides in the Netherlands [Internet]. 2019. Available from: https://www.bestrijdingsmiddelen-omwonenden.nl/english
7. Khoshhali M, Davoodi S, Ebrahimpour K, Shoshtari-Yeganeh B, Kelishadi R. The association between maternal exposure to organophosphate pesticides and neonatal anthropometric measures: A systematic review and meta-analysis. J Res Med Sci Off J Isfahan Univ Med Sci. 2020;25:79.
8. Kamai EM, McElrath TF, Ferguson KK. Fetal growth in environmental epidemiology: mechanisms, limitations, and a review of associations with biomarkers of non-persistent chemical exposures during pregnancy. Environ Health [Internet]. 2019 May 8;18(1):43. Available from: https://pubmed.ncbi.nlm.nih.gov/31068204
9. Windham G, Fenster L. Environmental contaminants and pregnancy outcomes. Fertil Steril [Internet]. 2008;89(2, Supplement):e111–6. Available from: https://www.sciencedirect.com/science/article/pii/S0015028207043208
10. Fucic A, Duca RC, Galea KS, Maric T, Garcia K, Bloom MS, et al. Reproductive Health Risks Associated with Occupational and Environmental Exposure to Pesticides. Int J Environ Res Public Health. 2021 Jun;18(12).
11. Overpeck MD, Hediger ML, Zhang J, Trumble AC, Klebanoff MA. Birth weight for gestational age of Mexican American infants born in the United States. Obstet Gynecol. 1999 Jun;93(6):943–7.
12. ACOG Committee Opinion No 579: Definition of term pregnancy. Obstet Gynecol. 2013 Nov;122(5):1139–40.
13. Nationaal Georegister (NGR). Dutch cadastral key-registry of buildings and addresses (BAG) [Internet]. 2016 [cited 2016 Jun 11]. Available from: https://data.overheid.nl/data/dataset/57437-bag-verblijfsobjecten
14. Dutch Ministry of Interior and Kingdom Relations. Dataportaal van de Nederlandse overheid—BRP Gewaspercelen [Internet]. 2016 [cited 2016 Jun 24]. Available from: https://data.overheid.nl/data/dataset/basisregistratie-gewaspercelen-brp

15. Statistics Netherlands (CBS). StatLine: Landbouw; gewassen, dieren, grondgebruik en arbeid op nationaal niveau 2009–2013 [Internet]. 2013 [cited 2021 Feb 2]. Available from: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/81302ned/table?fromstatweb
16. Simoes M, Huss A, Janssen N, Vermeulen R. Self-reported psychological distress and self-perceived health in residents living near pesticide-treated agricultural land: a cross-sectional study in The Netherlands. Occup Environ Med. 2022 Feb;79(2):127–33.
17. The European Food Safety Authority (EFSA). EU Pesticides Database [Internet]. 2021 [cited 2021 Feb 3]. Available from: https://ec.europa.eu/food/plant/pesticides/eu-pesticides-db_en
18. Lewis KA, Tzilivakis J, Warner DJ, Green A. An international database for pesticide risk assessments and management. Hum Ecol Risk Assess An Int J. 2016 May;22(4):1050–64.
19. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. J Stat Softw [Internet]. 2011 Dec 12;45(3 SE-Articles):1–67. Available from: https://www.jstatsoft.org/index.php/jss/article/view/v045i03
20. Wolters A, Linnemann V, van de Zande JC, Vereecken H. Field experiment on spray drift: Deposition and airborne drift during application to a winter wheat crop. Sci Total Environ [Internet]. 2008 Nov 1 [cited 2018 Jul 26];405(1–3):269–77. Available from: https://www.sciencedirect.com/science/article/pii/S004896970800702X
21. Rautmann D, Streloke M, Winkler R. New basic drift values in the authorization procedure for plant protection products. Workshop on Risk Assessment and Risk Mitigation Measures (WORMM), vol. 13. 2001. 3–141 p.
22. Simões M, Huss A, Brouwer M, Krop E, Janssen N, Vermeulen R. Residential proximity to crops and agricultural pesticide use and cause-specific mortality: A prospective census-based cohort study in the Netherlands. Sci Total Environ. 2022 Apr;817:152932.
23. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009 Jun;338:b2393.
24. Gray LEJ, Ostby JS, Kelce WR. Developmental effects of an environmental antiandrogen: the fungicide vinclozolin alters sex differentiation of the male rat. Toxicol Appl Pharmacol. 1994 Nov;129(1):46–52.
25. Lambright C, Ostby J, Bobseine K, Wilson V, Hotchkiss AK, Mann PC, et al. Cellular and Molecular Mechanisms of Action of Linuron: An Antiandrogenic Herbicide that Produces Reproductive Malformations in Male Rats. Toxicol Sci [Internet]. 2000 Aug 1;56(2):389–99. Available from: https://doi.org/10.1093/toxsci/56.2.389
26. Zhang C, Schilirò T, Gea M, Bianchi S, Spinello A, Magistrato A, et al. Molecular Basis for Endocrine Disruption by Pesticides Targeting Aromatase and Estrogen Receptor. Int J Environ Res Public Health [Internet]. 2020 Aug 5;17(16):5664. Available from: https://pubmed.ncbi.nlm.nih.gov/32764486
27. Figueiredo DM, Duyzer J, Huss A, Krop EJM, Gerritsen-Ebben MG, Gooijer Y, et al. Spatio-temporal variation of outdoor and indoor pesticide air concentrations in homes near

agricultural fields. Atmos Environ [Internet]. 2021;262:118612. Available from: https://www.sciencedirect.com/science/article/pii/S1352231021004349

28. Ding H, Zheng W, Han H, Hu X, Hu B, Wang F, et al. Reproductive toxicity of linuron following gestational exposure in rats and underlying mechanisms. Toxicol Lett [Internet]. 2016/12/19. 2017 Jan 15;266:49–55. Available from: https://pubmed.ncbi.nlm.nih.gov/28007637
29. European Food Safety Authority (EFSA). Peer review of the pesticide risk assessment of the active substance linuron. EFSA J [Internet]. 2016 Jul 1;14(7):e04518. Available from: https://doi.org/10.2903/j.efsa.2016.4518
30. European Food Safety Authority (EFSA), Abdourahime H, Anastassiadou M, Arena M, Auteri D, Barmaz S, et al. Peer review of the pesticide risk assessment of the active substance thiacloprid. EFSA J [Internet]. 2019 Mar 1;17(3):e05595. Available from: https://doi.org/10.2903/j.efsa.2019.5595
31. European Food Safety Authority (EFSA). Conclusion on the peer review of the pesticide risk assessment of the active substance fluroxypyr (evaluated variant fluroxypyr-meptyl). EFSA J [Internet]. 2011 Mar 1;9(3):2091. Available from: https://doi.org/10.2903/j.efsa.2011.2091
32. European Food Safety Authority (EFSA). Conclusion regarding the peer review of the pesticide risk assessment of the active substance glufosinate. EFSA J [Internet]. 2005 Apr 1;3(4):27r. Available from: https://doi.org/10.2903/j.efsa.2005.27r
33. The European Food Safety Authority (EFSA). Peer review of the pesticide risk assessment of the active substance picoxystrobin. EFSA J [Internet]. 2016 Jun 1;14(6):e04515. Available from: https://doi.org/10.2903/j.efsa.2016.4515
34. Palatianou ME, Simos Y V, Andronikou SK, Kiortsis DN. Long-term metabolic effects of high birth weight: a critical review of the literature. Horm Metab Res = Horm und Stoffwechselforsch = Horm Metab. 2014 Dec;46(13):911–20.
35. Scifres CM. Short- and Long-Term Outcomes Associated with Large for Gestational Age Birth Weight. Obstet Gynecol Clin North Am. 2021 Jun;48(2):325–37.

SUPPLEMENTAL MATERIAL

Contents

S1 Exposure assessment

Proxies for residential exposure to pesticides were computed based on (1) the types of crop cultivation present in the vicinity of residences and (2) the amount of active ingredients (AIs) used in those crops.
To determine which crops were located near residences and compute the area of crop within buffers around the residences we used the annual Basisregistratie Gewaspercelen (BRP) polygon maps from 2009 to 2013. We selected crops based on their contribution to the total agricultural surface area in The Netherlands, and grouped them into 12 groups (Table S1.1). By overlaying a point layer pertaining the coordinates of all residences in the Netherlands (the Basisregistraties Adressen en Gebouwen, BAG), we were able to compute buffers of 50m, 100m, 250m and 500m around the residences and extract the area of crop intersecting those buffers.

Table S1.1: Crops selected for the study and their correspondence to the crops featured in the Basisregistratie Gewaspercelen (BRP) land use maps and the Farmers' Surveys crops

Study Crops	Percentage of total agricultural surface[1]	BRP crops (codes) [in Dutch]	Farmers' Surveys crops [in Dutch]
Maize	39.4	Maïs, snij- (259), Maïs, korrel- (316), Maïs, corncob mix (317), Maïs, suiker- (814), Maïs, energie- (2032)	Snijmais
Cereals	18.2		
Winter wheat	9.5	Tarwe, winter- (233)	Wintertarwe
Summer barley	4.4	Gerst, zomer- (236)	Zomergerst
Summer wheat	2.3	Tarwe, zomer- (234)	Zomertarwe
Other grains2	1.9	Triticale (314), Gerst, winter- (235), Haver (238), Rogge (geen snijrogge) (237), Overige granen (661), Boekweit (659), Gierst (660) Graansorgho (658), Overige granen (2652)	
Potatoes	12.2		
Potatoes for consumption	5.7	Aardappelen, consumptie- op zand-/veengrond (1910), consumptie- op kleigrond (1909), cons. op kleigrond (vroeg, loofver voor 15-07) (1911), cons. op zand/veen (vroeg, loofver voor 15-07) (1912), consumptie op zand/veengrond (3792), consumptie op klei/lössgrond (2951)	Consumptie-aardappelen
Potatoes for starch	3.5	Aardappelen, zetmeel- (1934), zetmeel- TBM pootgoed (1935), zetmeel geleverd aan buitenland (859), zetmeel (3732)	Zetmeelaardappelen
Seed potatoes	2.8	Aardappelen, poot op klei, uitgroeiteelt (loofver na 15-08) (1926), poot- op kleigrond (1928), poot- op zand-/veengrond (1929,)poot op klei/lössgrond (3730), poot op zand/veengrond (3731)	Pootaardappelen
Potatoes, other[2]	0.1	Aardappelen, bestrijdingsmaatregel AM (2025)	
Beets	6.1	Bieten, suiker- (256), voeder- (257), voeder- (inclusief aardperen) (2651), Aardperen (1949)	Suikerbieten
Ornamental plants and tree nurseries	5.0	Boomkwekerij en vaste planten (229), Woudbomen met korte omlooptijd (2297), Boomkwekerijgewassen en vaste planten, pot- en containerveld (294), Boomkwekerijgewassen en vaste planten, open grond (3806)	Bloemkwek. gewassen open grond, Bos- en haagplantsoen, Laan- en parkbomen, Rozenstruiken open grond, Sierconiferen, Vaste planten open grond, Vruchtbomen
Vegetables (open field)	4.8	Groenten open grond (inclusief groentezaden) (672)	Aardbeien open grond (productie), Asperges, Bloemkool, Waspeen en bospeen, Broccoli, Prei, Schorseneren, Sluitkool, Spruitkool, Stambonen, Winterpeen, Witlofwortel
Fruit (mainly apples and pears)	3.5	Fruit (212)	Appelen, Peren
Flower bulbs	2.6	Bloembollen en – knollen (176)	Gladiolen, Hyacinten (bollen), Irissen, Lelies (bollen), Narcissen (bollen), Tulpen (bollen)

[1] Average percentage of surface area of total agricultural surface in 2009–2014, BRP. Excluding grassland, nature and fields classified as other.

[2] Not used as an individual crop in this study

Because analyses in this study were performed within CBS´s secure environment and all identifiers (personal numbers, addresses) were masked, we did not know a priori, which addresses corresponded to the mothers in our study population. We therefore computed the areas of crops for all addresses in the Netherlands (~9 million) and in order to decrease computational time in calculating the exact area of crop within a certain buffer around each of these addresses, we converted polygon BRP datasets to raster datasets (10 × 10m resolution) and determined surface area of crops in squares around buildings using a moving average. The resulting buffers were, in fact, squared in shape and had a larger area than a round buffer. The correlation between the calculated crop area of our square buffers, as compared to circular buffers was high (0.88–1.00, depending on the type of crop. For example, for the 50-meter buffer, the area of cultivated crop using the BRP data is calculated for a block consisting of five 10 by 10 meter raster cells above, below, left and right of the centre cell (Figure S1.1). The surface areas of the squared buffers (50, 100, 250 and 500m) around residences covered areas of 1.21 ha, 4.41 ha, 26.01 ha and 102.01 ha, respectively.

Information on AI's dosage used in each crop type was obtained from the Farmers´ Surveys (FS) from 2008 and 2012. The crops included in these surveys do not seamlessly correspond to those in the BRP (Table S1.1). For example, information on dosage used in several types of flower bulbs was available in the FS but such specificity on the type of flower bulbs was not available in the BRP maps from 2009 to 2013. In the case of potatoes, information on dosage was available for potatoes for consumption in general with no distinction between those grown in clay or in sandy soils. In such instances, we either calculated the average dosage used in the crops that fall within the BRP crop group or assumed that the same dosage was used in the several BRP crops that make up a FS crop. The FS are conducted roughly every four years and data on pesticide usage is not available for the years between surveys. To compute the amount of AIs used each year around a residence, we assigned the dosage of a FS to its consecutive years and calculated the mean dosage for the year that was not consecutive to either FS (Table S1.2). To compute the amount of a specific AI used in buffers around residences in a certain year, we multiplied the dosage (kg/ha) reported to be used in a certain crop by the area of that crop that was present within the buffers. By summing all the amounts used in the crops surrounding the residence, we obtained the total amount of AI used around that residence. Then, by averaging these amounts over the study period (2009–2013), we obtained the the final exposure proxy for residential pesticide exposure: the average amount of AI used in 2009–2013 within buffers of 50m, 100m, 250m and 500m around residences.

Figure S1.2 shows a step-by-step example of these calculations for the amount of AI1 used within a 100m buffer around residence A.

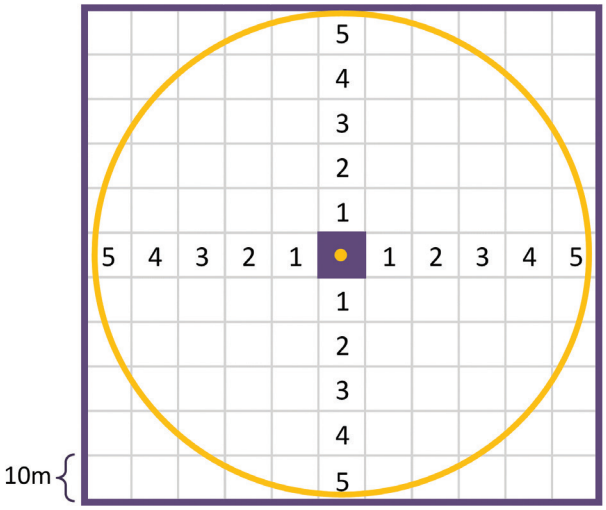
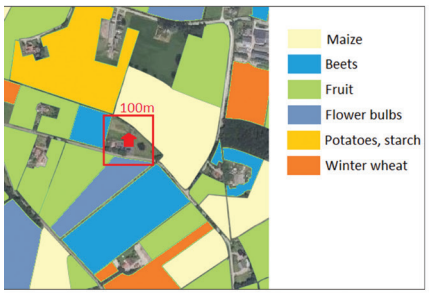


Figure S1.1: For a 50-meter buffer, the area around a residence using the moving average method (purple) results in a square shaped area around a raster cell. This area is therefore larger than an area calculated using the Buffer_analysis method in ArcGIS (orange).

Table S1.2: Correspondence between annual crop maps (BRP) and quadrennial Farmers' Surveys.

Available BRP years	Available Farmers´ Survey years	Farmers´ Survey used for correspondence to BRP
–	2008	–
2009	–	2008
2010	–	mean 2008–2012
2011	–	2012
2012	2012	2012
2013	–	2012

Figure S1.2: Example of the calculation of the average amount of AI1 used in 100m around residence A in 2009–2013.



Residence A	100m buffer (ha)					
	Maize	Beets	Fruit	Flower bulbs	Potatoes, starch	Winter wheat
2009	0.225	0.033	0.025	0.045	0	0
2010	0.225	0	0.025	0.055	0.034	0
2011	0.225	0	0.025	0.077	0	0.035
2012	0.225	0.033	0.025	0.045	0	0
2013	0.225	0	0.025	0.032	0.033	0

Study year	Farmers' survey year	FS information used	Dosage AI_1 (kg/ha)					
			Maize	Beets	Fruit	Flower bulbs	Potatoes, starch	Winter wheat
	2008		0.88	1.84	1.57	1.64	0.95	1.10
2009	-	**2008**	0.88	1.84	1.57	1.64	0.95	1.10
2010	-	**mean 2008-2012**	0.44	1.33	1.49	2.50	1.57	1.83
2011	-	**2012**	0	0.82	1.40	3.36	0.62	0.73
2012	**2012**	**2012**	0	0.82	1.40	3.36	0.62	0.73
2013	-	**2012**	0	0.82	1.40	3.36	0.62	0.73

$$Amount_{AI_1}(kg) = Area_{Maize} * Dosage_{AI_1 in\ Maize} + Area_{Beets} * Dosage_{AI_1 in\ Beets} + Area_{Fruit} * Dosage_{AI_1 in\ Fruit} + Area_{Flower\ bulbs} * Dosage_{AI_1 in\ Flower\ bulbs} + Area_{Potatoes\ starch} * Dosage_{AI_1 in\ Potatoes\ starch} + Area_{Winter\ wheat} * Dosage_{AI_1 in\ Winter\ wheat}$$

where, *Area* = area, in hectares, of specific crop group within a buffer (50m, 100m, 250m or 500m)
Dosage = reported amount of AI used per hectare in a specific crop group

Amount AI_1 (kg) in residence A		
2009	0.37	Average amount used 2009-2013
2010	0.33	
2011	0.32	**0.28**
2012	0.21	
2013	0.16	

S2 Correlation matrices of the 12 active ingredients with clear evidence of toxic effects on reproduction and fetal development.

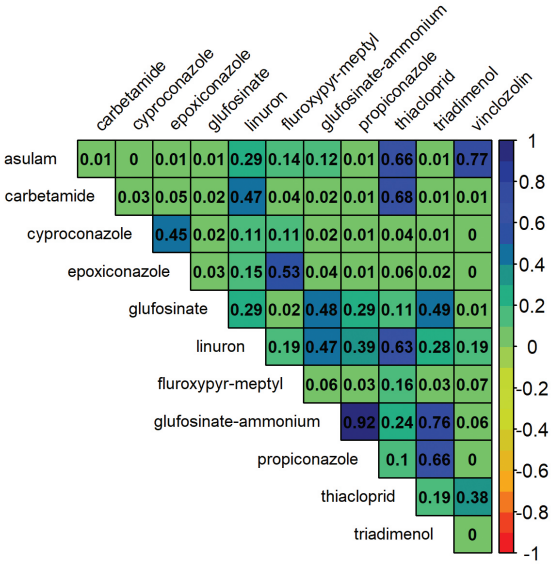


Figure S2.1: Spearman correlations between the amount of 12 AIs with clear evidence of reproductive and/or developmental effects used in 50m around mothers' residences.

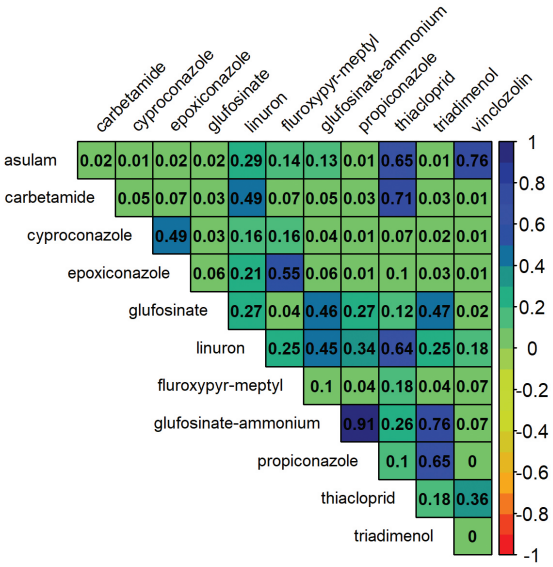


Figure S2.2: Spearman correlations between the amount of 12 AIs with clear evidence of reproductive and/or developmental effects used in 100m around mothers' residences.

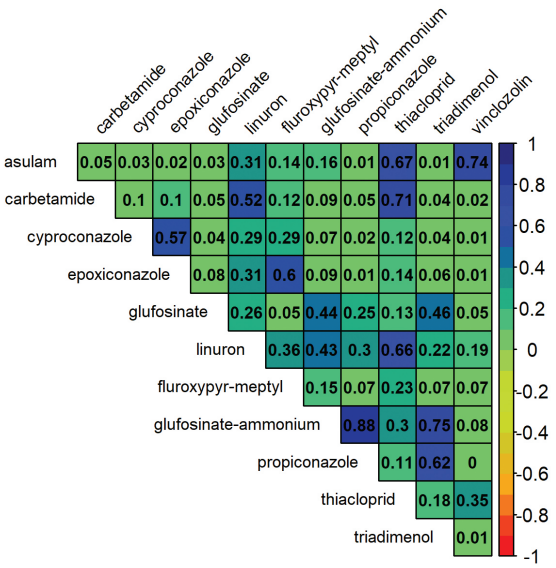


Figure S2.3: Spearman correlations between the amount of 12 AIs with clear evidence of reproductive and/or developmental effects used in 250m around mothers' residences.

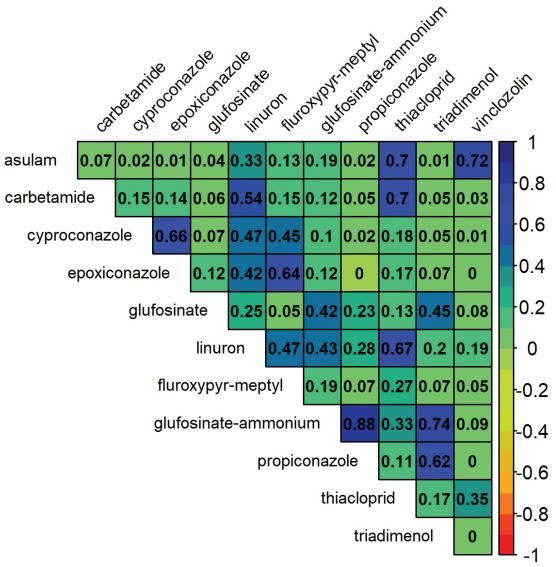


Figure S2.4: Spearman correlations between the amount of 12 AIs with clear evidence of reproductive and/or developmental effects used in 500m around mothers' residences.

S3 Correlation matrices of the all active ingredients included in the study

Legend of Figures S3.1-S3.4

Code	Name	Code	Name	Code	Name	Code	Name
1	2,4-dm	**36**	dithianonm	**71**	glufosinate-ammoniumm	**106**	rimsulfuronm
2	abamectinm	**37**	dodinem	**72**	mandipropamidm	**107**	s-metolachlorm
3	acetamipridm	**38**	epoxiconazolem	**73**	manebm	**108**	spinosadm
4	amitrolm	**39**	esfenvaleratem	**74**	mcpam	**109**	spirodiclofenm
5	asulamm	**40**	ethofumesatem	**75**	mecoprop-pm	**110**	sulcotrionm
6	azoxystrobinm	**41**	fenhexamidm	**76**	mepanipyrimm	**111**	tebuconazolem
7	bentazonm	**42**	phenmediphamm	**77**	mesotrionem	**112**	tebufenpyradm
8	benthiavalicarb isopropylm	**43**	fenoxycarbm	**78**	metalaxyl-mm	**113**	teflubenzuronm
9	bifenazatem	**44**	fenpropimorphm	**79**	metamitronm	**114**	tepraloxydinm
10	bifenoxm	**45**	flonicamidm	**80**	metazachlorm	**115**	terbuthylazinem
11	bitertanolm	**46**	fluazifop-p-butylm	**81**	methiocarb m	**116**	thiacloprid m
12	boscalidm	**47**	fluazinamm	**82**	methoxyfenozidem	**117**	thiophanate-methylm
13	bupirimatem	**48**	fludioxonilm	**83**	metiramm	**118**	tolclofos-methylm
14	captanm	**49**	fluopicolidem	**84**	metoxuronm	**119**	topramezonem
15	carbetamidem	**50**	fluoxastrobinm	**85**	metribuzinm	**120**	triallatem
16	chlorprophamm	**51**	fluroxypyrm	**86**	metsulfuron-methylm	**121**	triadimenolm
17	chlorothalonilm	**52**	folpetm	**87**	nicosulfuronm	**122**	trifloxystrobinm
18	chloridazonm	**53**	foramsulfuronm	**88**	pencycuronm	**123**	triflusulfuron-methylm
19	clomazonem	**54**	fosetylm	**89**	pendimethalinm	**124**	vinclozolinm
20	clopyralidm	**55**	glufosinatem	**90**	picoxystrobinm	**125**	aclonifenm
21	cyazofamidm	**56**	glyphosatem	**91**	pirimicarbm	**126**	bixafenm
22	cycloxydimm	**57**	haloxyfop-p-methylm	**92**	pirimiphos-methylm	**127**	chlorantraniliprolem
23	florasulamm	**58**	imidaclopridm	**93**	prochlorazm	**128**	emamectin benzoatem
24	cydia pomonella gv granulosevirusm	**59**	indoxacarbm	**94**	procymidonem	**129**	fenamidonem
25	cymoxanilm	**60**	iodosulfuron-methyl-sodiumm	**95**	propamocarbm	**130**	fenpropidinm
26	cyproconazolem	**61**	ioxynil octanoatem	**96**	propiconazolem	**131**	imazalilm
27	cyprodinilm	**62**	iprodionm	**97**	propyzamidem	**132**	mesosulfuron-methylm
28	deltamethrinm	**63**	isoproturonm	**98**	prosulfocarbm	**133**	penconazolem
29	desmediphamm	**64**	copper oxychloridem	**99**	prothioconazolem	**134**	pyroxsulamm
30	dicambam	**65**	kresoxim-methylm	**100**	pymetrozinem	**135**	spirotetramatm
31	difenoconazolem	**66**	lambda-cyhalothrinm	**101**	pyraclostrobinm	**136**	tembotrionem
32	dimethenamid-pm	**67**	linuronm	**102**	pyridatem	**137**	thiamethoxamm
33	dimethoatem	**68**	mancozebm	**103**	pyrimethanilm	**138**	thiramm
34	dimethomorphm	**69**	fluroxypyr-meptylm	**104**	quinoclaminem	**139**	sulphurm
35	diquatm	**70**	fosetyl-alm	**105**	quizalofop-p-ethylm		

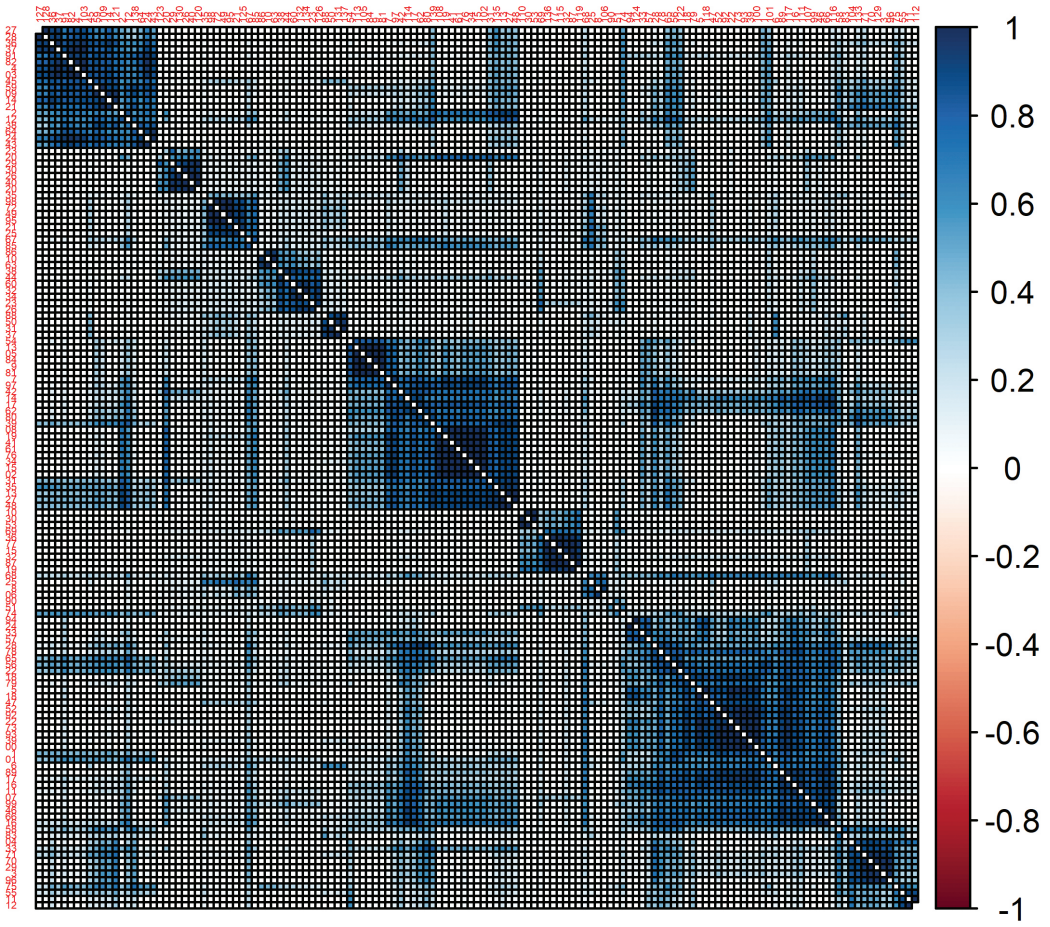


Figure S3.1: Spearman correlations between the 139 used in 50m around mothers' residences.

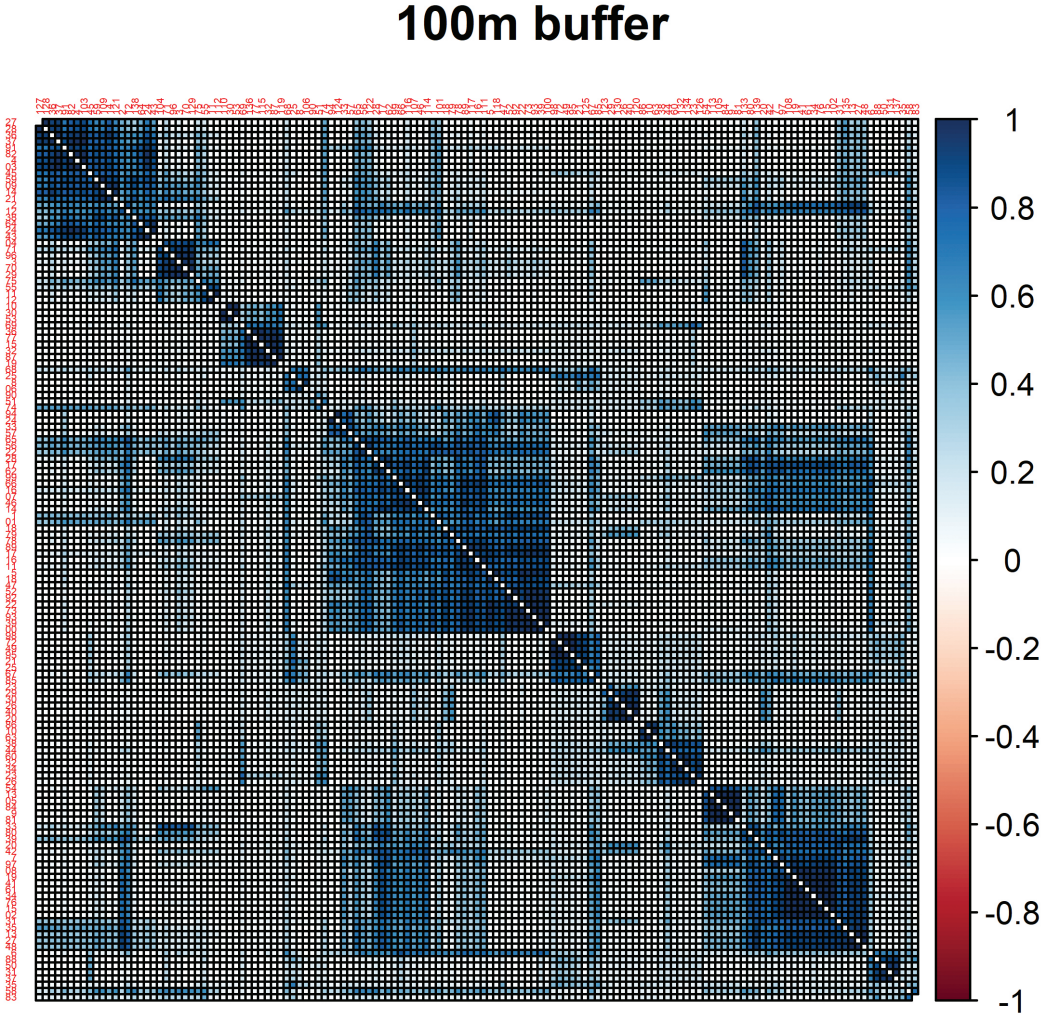


Figure S3.2: Spearman correlations between the 139 used in 5010 around mothers' residences.

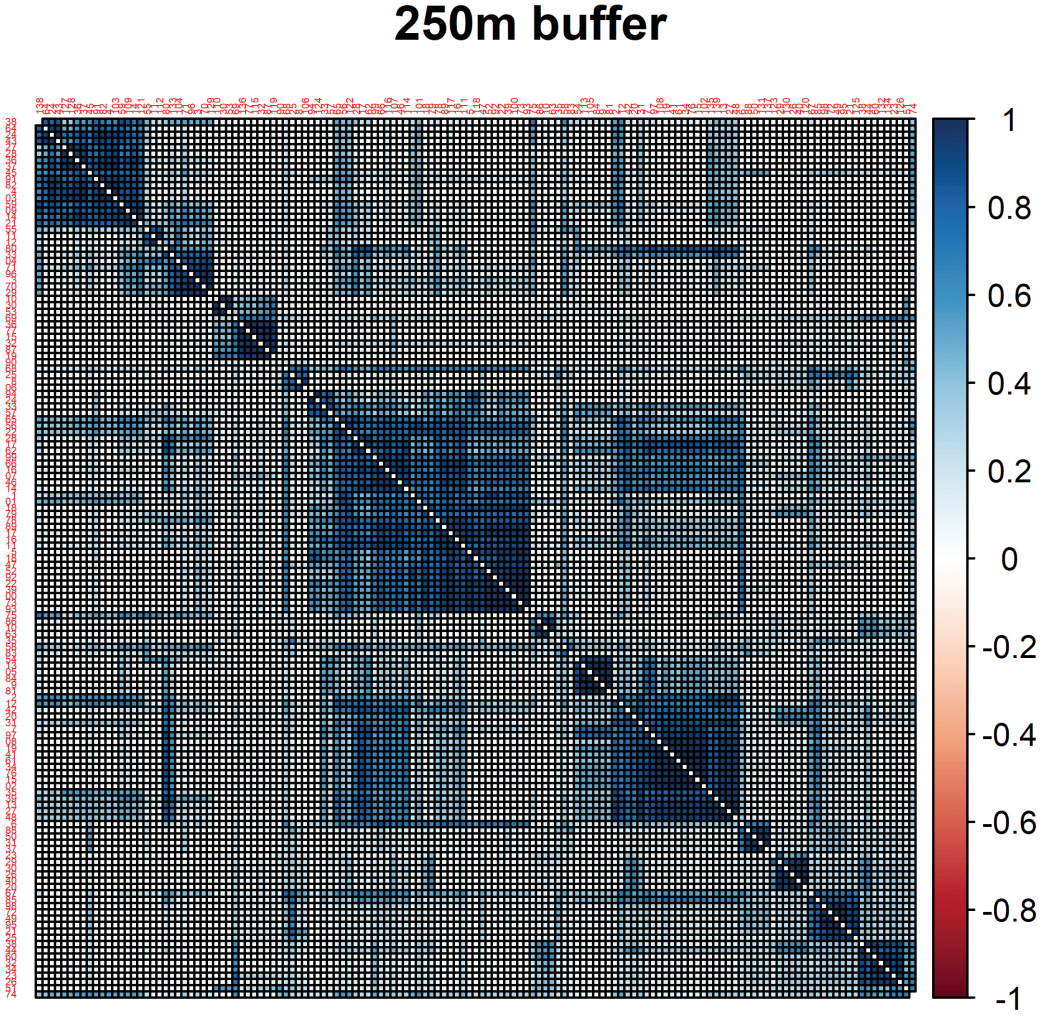


Figure S3.3: Spearman correlations between the 139 used in 250m around mothers' residences.

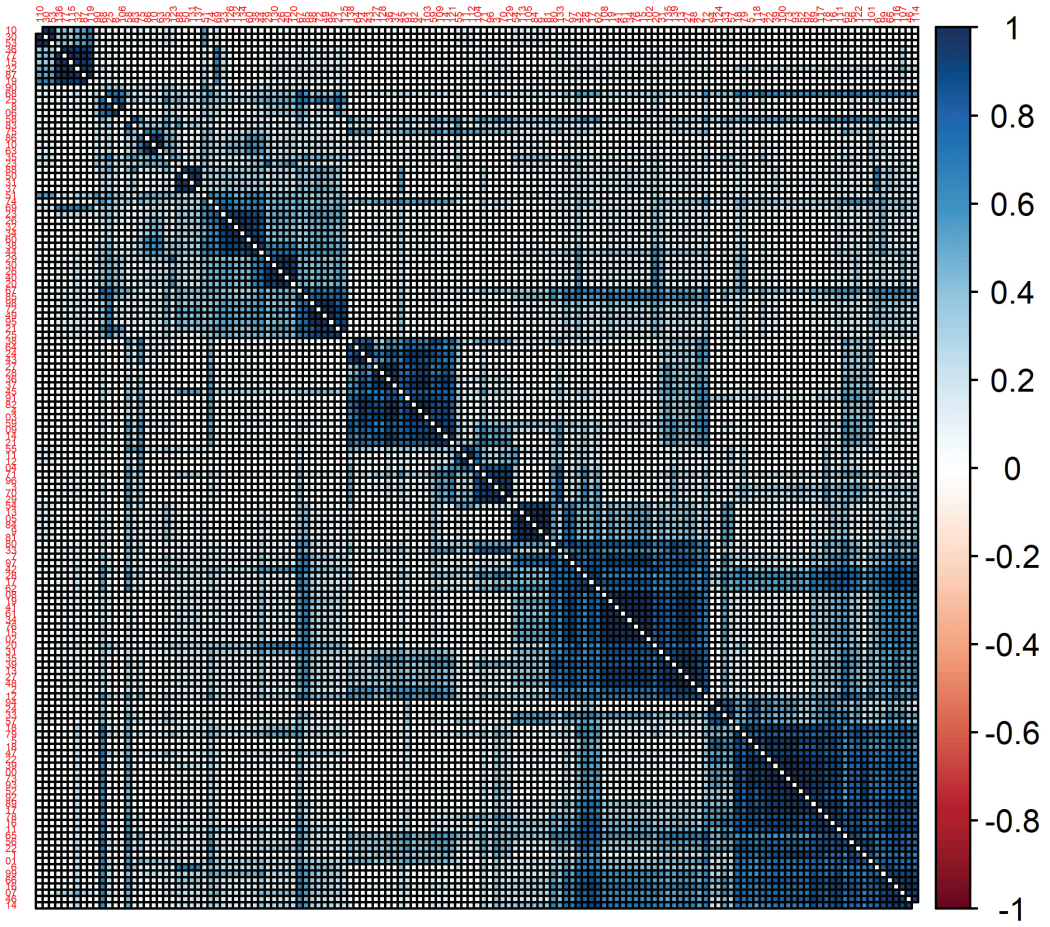


Figure S3.4: Spearman correlations between the 139 used in 500m around mothers' residences.

S4 Descritpive statistics of the exposure variables (amount of Active ingredient used in buffers around mothers' residences)

Active ingredient	50m buffer							
	Not exposed		Exposed					
	n	%	n	%	median	iqr	p1	p90
2,4-d	336621	99.0%	3326	1.0%	0.0034	0.0090	0.0003	0.0224
abamectin	337129	99.2%	2818	0.8%	0.0008	0.0019	0.0001	0.0038
acetamiprid	332517	97.8%	7430	2.2%	0.0006	0.0023	0.0000	0.0121
amitrol	338555	99.6%	1392	0.4%	0.0591	0.1236	0.0071	0.2358
asulam	339098	99.8%	849	0.2%	0.1640	0.3476	0.0179	0.7093
azoxystrobin	333256	98.0%	6691	2.0%	0.0044	0.0140	0.0004	0.0429
bentazon	334738	98.5%	5209	1.5%	0.0042	0.0134	0.0002	0.0401
benthiavalicarb isopropyl	337310	99.2%	2637	0.8%	0.0020	0.0053	0.0002	0.0133
bifenazate	339175	99.8%	772	0.2%	0.0020	0.0045	0.0002	0.0093
bifenox	338404	99.5%	1543	0.5%	0.0030	0.0065	0.0003	0.0127
bitertanol	339304	99.8%	643	0.2%	0.0023	0.0052	0.0003	0.0102
boscalid	330625	97.3%	9322	2.7%	0.0063	0.0153	0.0006	0.0371
bupirimate	335643	98.7%	4304	1.3%	0.0268	0.0618	0.0028	0.1459
captan	334992	98.5%	4955	1.5%	0.1759	1.1430	0.0011	3.1157
carbetamide	338260	99.5%	1687	0.5%	0.0479	0.1054	0.0057	0.2014
chlorpropham	336213	98.9%	3734	1.1%	0.0267	0.0624	0.0030	0.1491
chlorothalonil	334802	98.5%	5145	1.5%	0.0213	0.0730	0.0013	0.1751
chloridazon	336762	99.1%	3185	0.9%	0.0242	0.0632	0.0024	0.1429
clomazone	336342	98.9%	3605	1.1%	0.0007	0.0022	0.0001	0.0068
clopyralid	336239	98.9%	3708	1.1%	0.0007	0.0017	0.0001	0.0035
cyazofamid	334952	98.5%	4995	1.5%	0.0083	0.0201	0.0008	0.0450
cycloxydim	337431	99.3%	2516	0.7%	0.0009	0.0038	0.0001	0.0131
florasulam	325590	95.8%	14357	4.2%	0.0000	0.0001	0.0000	0.0003
cydia pomonella gv granulosevirus	338555	99.6%	1392	0.4%	0.0003	0.0006	0.0000	0.0012
cymoxanil	334952	98.5%	4995	1.5%	0.0159	0.0363	0.0016	0.0722
cyproconazole	337556	99.3%	2391	0.7%	0.0009	0.0021	0.0001	0.0045
cyprodinil	335738	98.8%	4209	1.2%	0.0126	0.0342	0.0012	0.0785
deltamethrin	330751	97.3%	9196	2.7%	0.0002	0.0010	0.0000	0.0035
desmedipham	337556	99.3%	2391	0.7%	0.0016	0.0038	0.0002	0.0077
dicamba	334925	98.5%	5022	1.5%	0.0007	0.0015	0.0001	0.0028
difenoconazole	334733	98.5%	5214	1.5%	0.0025	0.0062	0.0002	0.0132
dimethenamid-p	326460	96.0%	13487	4.0%	0.0326	0.0714	0.0035	0.1408
dimethoate	335619	98.7%	4328	1.3%	0.0047	0.0181	0.0003	0.0581
dimethomorph	336644	99.0%	3303	1.0%	0.0126	0.0578	0.0006	0.1705

Table S4.1: Number of exposed and unexposed mothers to each of the 139 active ingredients included in the study and measures of center and spread for the exposed mothers in 50m and 100m around mothers' residences.

100m buffer							
Not exposed		Exposed					
n	%	n	%	median	iqr	p1	p90
331613	97.5%	8334	2.5%	0.0083	0.0233	0.0006	0.0639
332651	97.9%	7296	2.1%	0.0021	0.0053	0.0002	0.0121
321945	94.7%	18002	5.3%	0.0017	0.0069	0.0001	0.0331
336339	98.9%	3608	1.1%	0.1427	0.3157	0.0118	0.6816
337913	99.4%	2034	0.6%	0.4305	1.0532	0.0301	2.2279
323703	95.2%	16244	4.8%	0.0119	0.0420	0.0007	0.1372
327220	96.3%	12727	3.7%	0.0120	0.0431	0.0006	0.1312
333856	98.2%	6091	1.8%	0.0069	0.0191	0.0004	0.0506
338000	99.4%	1947	0.6%	0.0054	0.0144	0.0004	0.0289
335953	98.8%	3994	1.2%	0.0082	0.0209	0.0005	0.0440
338237	99.5%	1710	0.5%	0.0054	0.0124	0.0005	0.0290
317469	93.4%	22478	6.6%	0.0180	0.0478	0.0013	0.1167
329006	96.8%	10941	3.2%	0.0684	0.1837	0.0056	0.4587
327645	96.4%	12302	3.6%	0.4467	2.9527	0.0028	8.3749
335570	98.7%	4377	1.3%	0.1343	0.3295	0.0101	0.6798
330579	97.2%	9368	2.8%	0.0661	0.1834	0.0053	0.4772
327056	96.2%	12891	3.8%	0.0539	0.2079	0.0029	0.5380
331856	97.6%	8091	2.4%	0.0685	0.1887	0.0047	0.4625
330818	97.3%	9129	2.7%	0.0020	0.0070	0.0001	0.0214
330371	97.2%	9576	2.8%	0.0021	0.0054	0.0002	0.0118
327952	96.5%	11995	3.5%	0.0264	0.0679	0.0018	0.1586
333744	98.2%	6203	1.8%	0.0025	0.0106	0.0001	0.0400
304873	89.7%	35074	10.3%	0.0001	0.0004	0.0000	0.0010
336339	98.9%	3608	1.1%	0.0006	0.0015	0.0001	0.0036
327952	96.5%	11995	3.5%	0.0479	0.1209	0.0032	0.2639
333670	98.2%	6277	1.8%	0.0028	0.0070	0.0002	0.0152
329312	96.9%	10635	3.1%	0.0320	0.0993	0.0022	0.2473
317573	93.4%	22374	6.6%	0.0005	0.0029	0.0000	0.0105
333670	98.2%	6277	1.8%	0.0050	0.0121	0.0003	0.0264
327371	96.3%	12576	3.7%	0.0019	0.0046	0.0001	0.0093
326714	96.1%	13233	3.9%	0.0069	0.0183	0.0005	0.0429
306936	90.3%	33011	9.7%	0.0956	0.2314	0.0066	0.4789
329135	96.8%	10812	3.2%	0.0127	0.0521	0.0007	0.1803
331936	97.6%	8011	2.4%	0.0401	0.1935	0.0017	0.5517

Table S4.1: Number of exposed and unexposed mothers to each of the 139 active ingredients included in the study and measures of center and spread for the exposed mothers in 50m and 100m around mothers' residences. (cont.)

Active ingredient	50m buffer							
	Not exposed		Exposed					
	n	%	n	%	median	iqr	p1	p90
diquat	335686	98.7%	4261	1.3%	0.0068	0.0224	0.0005	0.0660
dithianon	337227	99.2%	2720	0.8%	0.0373	0.0890	0.0043	0.2214
dodine	338555	99.6%	1392	0.4%	0.0428	0.0843	0.0046	0.1632
epoxiconazole	333365	98.1%	6582	1.9%	0.0035	0.0091	0.0003	0.0193
esfenvalerate	331202	97.4%	8745	2.6%	0.0001	0.0004	0.0000	0.0025
ethofumesate	337556	99.3%	2391	0.7%	0.0311	0.0707	0.0036	0.1501
fenhexamid	337065	99.2%	2882	0.8%	0.0196	0.1142	0.0011	0.3026
phenmedipham	334067	98.3%	5880	1.7%	0.0251	0.0559	0.0029	0.1145
fenoxycarb	338555	99.6%	1392	0.4%	0.0053	0.0132	0.0006	0.0340
fenpropimorph	332120	97.7%	7827	2.3%	0.0072	0.0175	0.0007	0.0363
flonicamid	334317	98.3%	5630	1.7%	0.0012	0.0031	0.0001	0.0082
fluazifop-p-butyl	334929	98.5%	5018	1.5%	0.0021	0.0074	0.0001	0.0256
fluazinam	334218	98.3%	5729	1.7%	0.0139	0.0411	0.0010	0.1415
fludioxonil	335738	98.8%	4209	1.2%	0.0081	0.0220	0.0008	0.0519
fluopicolide	335169	98.6%	4778	1.4%	0.0067	0.0187	0.0005	0.0475
fluoxastrobin	337691	99.3%	2256	0.7%	0.0017	0.0049	0.0002	0.0170
fluroxypyr	329782	97.0%	10165	3.0%	0.0015	0.0033	0.0002	0.0070
folpet	337877	99.4%	2070	0.6%	0.0260	0.0734	0.0023	0.2705
foramsulfuron	334925	98.5%	5022	1.5%	0.0005	0.0010	0.0001	0.0020
fosetyl	338558	99.6%	1389	0.4%	0.0138	0.0320	0.0018	0.0636
glufosinate	338152	99.5%	1795	0.5%	0.0076	0.0166	0.0006	0.0376
glyphosate	328952	96.8%	10995	3.2%	0.0213	0.0711	0.0014	0.1947
haloxyfop-p-methyl	337316	99.2%	2631	0.8%	0.0013	0.0034	0.0001	0.0083
imidacloprid	336003	98.8%	3944	1.2%	0.0019	0.0041	0.0002	0.0085
indoxacarb	336594	99.0%	3353	1.0%	0.0012	0.0026	0.0001	0.0058
iodosulfuron-methyl-sodium	331650	97.6%	8297	2.4%	0.0000	0.0001	0.0000	0.0005
ioxynil octanoate	338260	99.5%	1687	0.5%	0.0083	0.0190	0.0010	0.0390
iprodion	336305	98.9%	3642	1.1%	0.0165	0.0362	0.0019	0.0714
isoproturon	338404	99.5%	1543	0.5%	0.0475	0.1022	0.0052	0.2002
copper oxychloride	339225	99.8%	722	0.2%	0.0059	0.0134	0.0007	0.0256
kresoxim-methyl	333040	98.0%	6907	2.0%	0.0036	0.0112	0.0002	0.0243
lambda-cyhalothrin	330396	97.2%	9551	2.8%	0.0003	0.0009	0.0000	0.0042
linuron	331786	97.6%	8161	2.4%	0.0231	0.0499	0.0025	0.0977
mancozeb	330537	97.2%	9410	2.8%	0.2194	0.5146	0.0214	1.2519
fluroxypyr-meptyl	325506	95.8%	14441	4.2%	0.0060	0.0123	0.0007	0.0239
fosetyl-al	337981	99.4%	1966	0.6%	0.0170	0.0428	0.0014	0.1049

100m buffer							
Not exposed		Exposed					
n	%	n	%	median	iqr	p1	p90
329701	97.0%	10246	3.0%	0.0196	0.0752	0.0010	0.2228
332954	97.9%	6993	2.1%	0.0888	0.2348	0.0079	0.5839
336339	98.9%	3608	1.1%	0.1010	0.2177	0.0090	0.4595
323780	95.2%	16167	4.8%	0.0108	0.0299	0.0007	0.0670
319018	93.8%	20929	6.2%	0.0004	0.0013	0.0000	0.0074
333670	98.2%	6277	1.8%	0.0963	0.2367	0.0068	0.5041
332573	97.8%	7374	2.2%	0.0525	0.3527	0.0021	0.9555
325259	95.7%	14688	4.3%	0.0715	0.1725	0.0054	0.3810
336339	98.9%	3608	1.1%	0.0122	0.0326	0.0010	0.0958
320795	94.4%	19152	5.6%	0.0218	0.0584	0.0014	0.1284
325838	95.8%	14109	4.2%	0.0033	0.0093	0.0002	0.0236
327454	96.3%	12493	3.7%	0.0058	0.0240	0.0002	0.0775
326336	96.0%	13611	4.0%	0.0428	0.1379	0.0026	0.4573
329312	96.9%	10635	3.1%	0.0210	0.0646	0.0014	0.1636
328373	96.6%	11574	3.4%	0.0205	0.0604	0.0014	0.1591
334394	98.4%	5553	1.6%	0.0051	0.0149	0.0004	0.0557
314974	92.7%	24973	7.3%	0.0043	0.0110	0.0003	0.0246
334791	98.5%	5156	1.5%	0.0606	0.1959	0.0045	0.7601
327371	96.3%	12576	3.7%	0.0014	0.0033	0.0001	0.0067
336403	99.0%	3544	1.0%	0.0349	0.0909	0.0028	0.1939
335497	98.7%	4450	1.3%	0.0190	0.0472	0.0013	0.1127
313738	92.3%	26209	7.7%	0.0643	0.2218	0.0035	0.6002
333329	98.1%	6618	1.9%	0.0034	0.0097	0.0002	0.0250
330327	97.2%	9620	2.8%	0.0048	0.0121	0.0003	0.0266
331507	97.5%	8440	2.5%	0.0029	0.0071	0.0002	0.0165
319329	93.9%	20618	6.1%	0.0001	0.0004	0.0000	0.0017
335570	98.7%	4377	1.3%	0.0240	0.0597	0.0018	0.1348
330774	97.3%	9173	2.7%	0.0417	0.1037	0.0034	0.2302
335953	98.8%	3994	1.2%	0.1281	0.3277	0.0080	0.6921
338125	99.5%	1822	0.5%	0.0133	0.0340	0.0010	0.0707
322826	95.0%	17121	5.0%	0.0093	0.0310	0.0006	0.0749
317135	93.3%	22812	6.7%	0.0007	0.0029	0.0000	0.0136
320510	94.3%	19437	5.7%	0.0677	0.1607	0.0052	0.3394
317512	93.4%	22435	6.6%	0.6195	1.6710	0.0440	4.2846
304767	89.7%	35180	10.3%	0.0177	0.0408	0.0014	0.0840
335037	98.6%	4910	1.4%	0.0425	0.1180	0.0035	0.2899

Table S4.1: Number of exposed and unexposed mothers to each of the 139 active ingredients included in the study and measures of center and spread for the exposed mothers in 50m and 100m around mothers' residences. (cont.)

Active ingredient	50m buffer							
	Not exposed		Exposed					
	n	%	n	%	median	iqr	p1	p90
glufosinate-ammonium	335857	98.8%	4090	1.2%	0.0078	0.0180	0.0007	0.0457
mandipropamid	334952	98.5%	4995	1.5%	0.0151	0.0371	0.0013	0.0917
maneb	337877	99.4%	2070	0.6%	0.0082	0.3301	0.0004	1.3719
mcpa	331044	97.4%	8903	2.6%	0.0215	0.0504	0.0019	0.1107
mecoprop-p	335879	98.8%	4068	1.2%	0.0036	0.0088	0.0004	0.0193
mepanipyrim	336740	99.1%	3207	0.9%	0.0138	0.0567	0.0008	0.1619
mesotrione	328923	96.8%	11024	3.2%	0.0034	0.0068	0.0004	0.0134
metalaxyl-m	336213	98.9%	3734	1.1%	0.0051	0.0114	0.0005	0.0266
metamitron	335482	98.7%	4465	1.3%	0.0770	0.1794	0.0091	0.3851
metazachlor	336966	99.1%	2981	0.9%	0.0319	0.0660	0.0038	0.1278
methiocarb	339175	99.8%	772	0.2%	0.0106	0.0239	0.0013	0.0498
methoxyfenozide	338555	99.6%	1392	0.4%	0.0048	0.0101	0.0006	0.0197
metiram	337538	99.3%	2409	0.7%	0.0189	0.0464	0.0018	0.1198
metoxuron	339175	99.8%	772	0.2%	0.0117	0.0263	0.0015	0.0547
metribuzin	333659	98.2%	6288	1.8%	0.0100	0.0220	0.0010	0.0462
metsulfuron-methyl	336424	99.0%	3523	1.0%	0.0004	0.0017	0.0000	0.0065
nicosulfuron	328923	96.8%	11024	3.2%	0.0018	0.0035	0.0002	0.0067
pencycuron	334952	98.5%	4995	1.5%	0.0043	0.0111	0.0004	0.0497
pendimethalin	337474	99.3%	2473	0.7%	0.0376	0.0860	0.0039	0.2011
picoxystrobin	339333	99.8%	614	0.2%	0.0006	0.0012	0.0001	0.0026
pirimicarb	335321	98.6%	4626	1.4%	0.0024	0.0082	0.0002	0.0321
pirimiphos-methyl	339098	99.8%	849	0.2%	0.0113	0.0290	0.0012	0.0644
prochloraz	337776	99.4%	2171	0.6%	0.0076	0.0247	0.0008	0.1079
procymidone	339579	99.9%	368	0.1%	0.0065	0.0150	0.0007	0.0334
propamocarb	333988	98.2%	5959	1.8%	0.0329	0.1247	0.0014	0.3495
propiconazole	338590	99.6%	1357	0.4%	0.0175	0.0457	0.0015	0.0922
propyzamide	336966	99.1%	2981	0.9%	0.0196	0.0483	0.0021	0.1185
prosulfocarb	335314	98.6%	4633	1.4%	0.0554	0.1368	0.0044	0.3497
prothioconazole	332788	97.9%	7159	2.1%	0.0073	0.0172	0.0007	0.0415
pymetrozine	337976	99.4%	1971	0.6%	0.0017	0.0051	0.0001	0.0156
pyraclostrobin	330204	97.1%	9743	2.9%	0.0023	0.0059	0.0002	0.0161
pyridate	338260	99.5%	1687	0.5%	0.0148	0.0323	0.0017	0.0615
pyrimethanil	337325	99.2%	2622	0.8%	0.0035	0.0283	0.0002	0.0741
quinoclamine	338590	99.6%	1357	0.4%	0.0168	0.0335	0.0021	0.0666
quizalofop-p-ethyl	339175	99.8%	772	0.2%	0.0009	0.0021	0.0001	0.0043
rimsulfuron	336306	98.9%	3641	1.1%	0.0001	0.0003	0.0000	0.0011

100m buffer							
Not exposed		Exposed					
n	%	n	%	median	iqr	p1	p90
329864	97.0%	10083	3.0%	0.0208	0.0547	0.0015	0.1342
327952	96.5%	11995	3.5%	0.0473	0.1251	0.0034	0.3123
334791	98.5%	5156	1.5%	0.0196	0.7431	0.0009	3.8265
318086	93.6%	21861	6.4%	0.0580	0.1603	0.0038	0.3659
329660	97.0%	10287	3.0%	0.0094	0.0246	0.0007	0.0567
331798	97.6%	8149	2.4%	0.0345	0.1740	0.0016	0.5109
312485	91.9%	27462	8.1%	0.0101	0.0220	0.0008	0.0450
330579	97.2%	9368	2.8%	0.0134	0.0327	0.0011	0.0778
328584	96.7%	11363	3.3%	0.2107	0.5335	0.0149	1.2313
332254	97.7%	7693	2.3%	0.0820	0.1888	0.0073	0.3989
338000	99.4%	1947	0.6%	0.0291	0.0776	0.0021	0.1553
336339	98.9%	3608	1.1%	0.0115	0.0262	0.0009	0.0573
333973	98.2%	5974	1.8%	0.0459	0.1276	0.0033	0.3590
338000	99.4%	1947	0.6%	0.0320	0.0852	0.0023	0.1705
324928	95.6%	15019	4.4%	0.0311	0.0763	0.0022	0.1651
330923	97.3%	9024	2.7%	0.0012	0.0049	0.0001	0.0208
312485	91.9%	27462	8.1%	0.0052	0.0114	0.0004	0.0228
327952	96.5%	11995	3.5%	0.0129	0.0381	0.0008	0.1499
333795	98.2%	6152	1.8%	0.1042	0.2687	0.0075	0.6444
338348	99.5%	1599	0.5%	0.0019	0.0042	0.0001	0.0090
328255	96.6%	11692	3.4%	0.0062	0.0218	0.0004	0.0829
337913	99.4%	2034	0.6%	0.0306	0.0864	0.0019	0.2049
334467	98.4%	5480	1.6%	0.0169	0.0614	0.0014	0.3099
339013	99.7%	934	0.3%	0.0161	0.0456	0.0010	0.0949
325532	95.8%	14415	4.2%	0.1016	0.4148	0.0033	1.1589
336386	99.0%	3561	1.0%	0.0385	0.1119	0.0023	0.2553
332254	97.7%	7693	2.3%	0.0486	0.1417	0.0039	0.3852
328558	96.6%	11389	3.4%	0.1585	0.4561	0.0104	1.1636
322189	94.8%	17758	5.2%	0.0210	0.0569	0.0014	0.1394
335262	98.6%	4685	1.4%	0.0049	0.0154	0.0003	0.0464
316560	93.1%	23387	6.9%	0.0069	0.0192	0.0005	0.0517
335570	98.7%	4377	1.3%	0.0411	0.1016	0.0031	0.2074
333280	98.0%	6667	2.0%	0.0094	0.0718	0.0004	0.2015
336386	99.0%	3561	1.0%	0.0384	0.0832	0.0037	0.1786
338000	99.4%	1947	0.6%	0.0025	0.0067	0.0002	0.0135
331001	97.4%	8946	2.6%	0.0002	0.0009	0.0000	0.0038

Table S4.1: Number of exposed and unexposed mothers to each of the 139 active ingredients included in the study and measures of center and spread for the exposed mothers in 50m and 100m around mothers' residences (cont.)

Active ingredient	50m buffer							
	Not exposed		Exposed					
	n	%	n	%	median	iqr	p1	p90
s-metolachlor	326173	95.9%	13774	4.1%	0.0110	0.0265	0.0011	0.0689
spinosad	338260	99.5%	1687	0.5%	0.0061	0.0170	0.0005	0.0353
spirodiclofen	336594	99.0%	3353	1.0%	0.0010	0.0025	0.0001	0.0051
sulcotrion	334606	98.4%	5341	1.6%	0.0042	0.0095	0.0004	0.0184
tebuconazole	330349	97.2%	9598	2.8%	0.0039	0.0107	0.0003	0.0302
tebufenpyrad	339304	99.8%	643	0.2%	0.0033	0.0076	0.0005	0.0149
teflubenzuron	339175	99.8%	772	0.2%	0.0034	0.0076	0.0004	0.0158
tepraloxydim	334566	98.4%	5381	1.6%	0.0005	0.0013	0.0000	0.0033
terbuthylazine	328923	96.8%	11024	3.2%	0.0210	0.0432	0.0024	0.0876
thiacloprid	330813	97.3%	9134	2.7%	0.0024	0.0070	0.0002	0.0239
thiophanate-methyl	336396	99.0%	3551	1.0%	0.0161	0.0418	0.0013	0.1168
tolclofos-methyl	339098	99.8%	849	0.2%	0.1062	0.2230	0.0107	0.4837
topramezone	328923	96.8%	11024	3.2%	0.0008	0.0015	0.0001	0.0028
triallate	337556	99.3%	2391	0.7%	0.0025	0.0055	0.0003	0.0117
triadimenol	337227	99.2%	2720	0.8%	0.0055	0.0114	0.0007	0.0222
trifloxystrobin	332799	97.9%	7148	2.1%	0.0035	0.0080	0.0004	0.0170
triflusulfuron-methyl	336034	98.8%	3913	1.2%	0.0007	0.0025	0.0001	0.0350
vinclozolin	339579	99.9%	368	0.1%	0.0005	0.0012	0.0001	0.0028
aclonifen	336524	99.0%	3423	1.0%	0.0176	0.0394	0.0017	0.0878
bixafen	335267	98.6%	4680	1.4%	0.0022	0.0050	0.0002	0.0103
chlorantraniliprole	338691	99.6%	1256	0.4%	0.0022	0.0046	0.0003	0.0089
emamectin benzoate	338691	99.6%	1256	0.4%	0.0007	0.0016	0.0001	0.0030
fenamidone	337981	99.4%	1966	0.6%	0.0016	0.0042	0.0001	0.0102
fenpropidin	337903	99.4%	2044	0.6%	0.0063	0.0144	0.0007	0.0302
imazalil	338988	99.7%	959	0.3%	0.0020	0.0043	0.0002	0.0087
mesosulfuron-methyl	336796	99.1%	3151	0.9%	0.0005	0.0011	0.0001	0.0021
penconazole	336088	98.9%	3859	1.1%	0.0011	0.0027	0.0001	0.0062
pyroxsulam	336796	99.1%	3151	0.9%	0.0004	0.0009	0.0000	0.0018
spirotetramat	336088	98.9%	3859	1.1%	0.0031	0.0091	0.0003	0.0209
tembotrione	330009	97.1%	9938	2.9%	0.0013	0.0027	0.0002	0.0052
thiamethoxam	336913	99.1%	3034	0.9%	0.0001	0.0003	0.0000	0.0019
thiram	337156	99.2%	2791	0.8%	0.2931	0.7849	0.0254	1.8486
sulphur	335643	98.7%	4304	1.3%	1.1691	2.5502	0.1236	5.2112
TOTAL	**318896**	**93.8%**	**21051**	**6.2%**	**0.2723**	**1.2106**	**0.0221**	**4.7950**

100m buffer							
Not exposed		Exposed					
n	%	n	%	median	iqr	p1	p90
306420	90.1%	33527	9.9%	0.0324	0.0867	0.0023	0.2310
335570	98.7%	4377	1.3%	0.0172	0.0521	0.0011	0.1197
331507	97.5%	8440	2.5%	0.0025	0.0064	0.0002	0.0146
326617	96.1%	13330	3.9%	0.0117	0.0299	0.0007	0.0615
316340	93.1%	23607	6.9%	0.0105	0.0327	0.0007	0.0917
338237	99.5%	1710	0.5%	0.0080	0.0182	0.0008	0.0425
338000	99.4%	1947	0.6%	0.0093	0.0247	0.0007	0.0494
326512	96.0%	13435	4.0%	0.0013	0.0039	0.0001	0.0103
312485	91.9%	27462	8.1%	0.0614	0.1383	0.0047	0.2906
318132	93.6%	21815	6.4%	0.0069	0.0218	0.0005	0.0757
331071	97.4%	8876	2.6%	0.0403	0.1224	0.0026	0.3578
337913	99.4%	2034	0.6%	0.2702	0.7002	0.0198	1.5049
312485	91.9%	27462	8.1%	0.0022	0.0048	0.0002	0.0095
333670	98.2%	6277	1.8%	0.0076	0.0184	0.0005	0.0394
332954	97.9%	6993	2.1%	0.0130	0.0295	0.0012	0.0641
322212	94.8%	17735	5.2%	0.0094	0.0237	0.0007	0.0540
329930	97.1%	10017	2.9%	0.0021	0.0082	0.0001	0.1165
339013	99.7%	934	0.3%	0.0013	0.0037	0.0001	0.0078
331444	97.5%	8503	2.5%	0.0509	0.1301	0.0040	0.3020
328050	96.5%	11897	3.5%	0.0063	0.0159	0.0004	0.0347
336640	99.0%	3307	1.0%	0.0051	0.0118	0.0005	0.0258
336640	99.0%	3307	1.0%	0.0017	0.0040	0.0002	0.0088
335037	98.6%	4910	1.4%	0.0039	0.0111	0.0003	0.0282
334556	98.4%	5391	1.6%	0.0190	0.0473	0.0014	0.1015
337655	99.3%	2292	0.7%	0.0062	0.0152	0.0004	0.0317
331948	97.6%	7999	2.4%	0.0014	0.0035	0.0001	0.0074
330086	97.1%	9861	2.9%	0.0027	0.0077	0.0002	0.0183
331948	97.6%	7999	2.4%	0.0012	0.0030	0.0001	0.0063
330086	97.1%	9861	2.9%	0.0081	0.0260	0.0006	0.0651
314945	92.6%	25002	7.4%	0.0038	0.0087	0.0003	0.0177
332371	97.8%	7576	2.2%	0.0003	0.0009	0.0000	0.0060
332911	97.9%	7036	2.1%	0.7198	2.0019	0.0514	5.1584
329006	96.8%	10941	3.2%	2.9843	7.3691	0.2359	16.2599
291366	**85.7%**	**48581**	**14.3%**	**0.8462**	**3.9810**	**0.0488**	**15.2486**

Table S4.2: Number of exposed and unexposed mothers to each of the 139 active ingredients included in the study and measures of center and spread for the exposed mothers in 250m and 500m around mothers' residences.

Active ingredient	250m buffer							
	Not exposed		Exposed					
	n	%	n	%	median	iqr	p1	p90
2,4-d	309539	91.1%	30408	8.9%	0.0260	0.0857	0.0014	0.2376
abamectin	311255	91.6%	28692	8.4%	0.0073	0.0193	0.0005	0.0441
acetamiprid	280507	82.5%	59440	17.5%	0.0069	0.0317	0.0003	0.1285
amitrol	326240	96.0%	13707	4.0%	0.4642	1.1693	0.0362	2.5568
asulam	332659	97.9%	7288	2.1%	1.4174	3.5477	0.0733	8.0524
azoxystrobin	283824	83.5%	56123	16.5%	0.0432	0.1638	0.0021	0.5366
bentazon	294733	86.7%	45214	13.3%	0.0512	0.1770	0.0022	0.5324
benthiavalicarb isopropyl	321715	94.6%	18232	5.4%	0.0277	0.0884	0.0014	0.2387
bifenazate	331775	97.6%	8172	2.4%	0.0167	0.0452	0.0011	0.1027
bifenox	325065	95.6%	14882	4.4%	0.0316	0.0773	0.0018	0.1720
bitertanol	333251	98.0%	6696	2.0%	0.0142	0.0358	0.0010	0.0871
boscalid	267012	78.5%	72935	21.5%	0.0792	0.2095	0.0044	0.4993
bupirimate	299152	88.0%	40795	12.0%	0.2353	0.7076	0.0157	1.7837
captan	295480	86.9%	44467	13.1%	1.4634	9.7909	0.0081	31.2215
carbetamide	321487	94.6%	18460	5.4%	0.4150	1.0913	0.0266	2.4449
chlorpropham	305053	89.7%	34894	10.3%	0.2083	0.6551	0.0131	1.7421
chlorothalonil	293497	86.3%	46450	13.7%	0.2003	0.7597	0.0092	2.0722
chloridazon	308251	90.7%	31696	9.3%	0.2421	0.6498	0.0136	1.5783
clomazone	304647	89.6%	35300	10.4%	0.0072	0.0260	0.0003	0.0788
clopyralid	302309	88.9%	37638	11.1%	0.0080	0.0198	0.0005	0.0447
cyazofamid	299991	88.2%	39956	11.8%	0.1131	0.2989	0.0063	0.6814
cycloxydim	316498	93.1%	23449	6.9%	0.0088	0.0364	0.0003	0.1287
florasulam	230062	67.7%	109885	32.3%	0.0006	0.0018	0.0000	0.0048
cydia pomonella gv granulosevirus	326240	96.0%	13707	4.0%	0.0021	0.0056	0.0002	0.0127
cymoxanil	299991	88.2%	39956	11.8%	0.1952	0.4950	0.0109	1.1182
cyproconazole	314189	92.4%	25758	7.6%	0.0106	0.0255	0.0006	0.0546
cyprodinil	300071	88.3%	39876	11.7%	0.1143	0.3787	0.0060	0.9451
deltamethrin	266622	78.4%	73325	21.6%	0.0024	0.0125	0.0001	0.0407
desmedipham	314189	92.4%	25758	7.6%	0.0187	0.0450	0.0011	0.0954
dicamba	297148	87.4%	42799	12.6%	0.0075	0.0175	0.0005	0.0377
difenoconazole	290775	85.5%	49172	14.5%	0.0270	0.0718	0.0016	0.1685
dimethenamid-p	232603	68.4%	107344	31.6%	0.4189	0.9544	0.0252	2.0079
dimethoate	301214	88.6%	38733	11.4%	0.0521	0.2110	0.0023	0.6795
dimethomorph	310882	91.5%	29065	8.5%	0.1865	0.7753	0.0085	2.1774

500m buffer							
Not exposed		Exposed					
n	%	n	%	median	iqr	p1	p90
264106	77.7%	75841	22.3%	0.0613	0.2295	0.0028	0.6344
268247	78.9%	71700	21.1%	0.0210	0.0544	0.0013	0.1250
209956	61.8%	129991	38.2%	0.0246	0.1071	0.0009	0.3655
306688	90.2%	33259	9.8%	1.1868	3.0941	0.0799	7.1968
319724	94.1%	20223	5.9%	3.0189	7.5247	0.1752	19.0223
214355	63.1%	125592	36.9%	0.1398	0.5322	0.0062	1.6676
233945	68.8%	106002	31.2%	0.1757	0.5924	0.0076	1.6795
299819	88.2%	40128	11.8%	0.0721	0.2363	0.0036	0.7362
316237	93.0%	23710	7.0%	0.0426	0.1143	0.0025	0.2607
303463	89.3%	36484	10.7%	0.0912	0.2227	0.0062	0.5239
320644	94.3%	19303	5.7%	0.0247	0.0709	0.0018	0.1929
186430	54.8%	153517	45.2%	0.2654	0.7321	0.0134	1.7022
241508	71.0%	98439	29.0%	0.7135	2.2066	0.0367	5.3696
235436	69.3%	104511	30.7%	4.1444	25.2923	0.0267	84.0937
288178	84.8%	51769	15.2%	1.1316	2.8242	0.0720	6.2981
252429	74.3%	87518	25.7%	0.5804	1.8581	0.0290	4.7459
230930	67.9%	109017	32.1%	0.6504	2.3543	0.0296	6.2614
259853	76.4%	80094	23.6%	0.7602	1.7335	0.0459	4.1286
251966	74.1%	87981	25.9%	0.0234	0.0825	0.0011	0.2307
248027	73.0%	91920	27.0%	0.0264	0.0586	0.0016	0.1283
247232	72.7%	92715	27.3%	0.3649	0.9215	0.0200	2.0716
281447	82.8%	58500	17.2%	0.0296	0.1045	0.0010	0.3168
131269	38.6%	208678	61.4%	0.0026	0.0069	0.0002	0.0184
306688	90.2%	33259	9.8%	0.0054	0.0146	0.0003	0.0356
247232	72.7%	92715	27.3%	0.5884	1.4236	0.0346	3.2826
273416	80.4%	66531	19.6%	0.0339	0.0707	0.0021	0.1452
243600	71.7%	96347	28.3%	0.3600	1.1736	0.0139	2.8097
188441	55.4%	151506	44.6%	0.0104	0.0465	0.0004	0.1337
273416	80.4%	66531	19.6%	0.0592	0.1235	0.0037	0.2543
250392	73.7%	89555	26.3%	0.0255	0.0547	0.0020	0.1183
228778	67.3%	111169	32.7%	0.0957	0.2327	0.0058	0.5260
128841	37.9%	211106	62.1%	1.5794	3.1416	0.1308	6.5941
247983	72.9%	91964	27.1%	0.1772	0.6925	0.0067	1.9854
267506	78.7%	72441	21.3%	0.6549	2.4164	0.0255	6.1562

Table S4.2: Number of exposed and unexposed mothers to each of the 139 active ingredients included in the study and measures of center and spread for the exposed mothers in 250m and 500m around mothers' residences. (cont.)

Active ingredient	250m buffer							
	Not exposed		Exposed					
	n	%	n	%	median	iqr	p1	p90
diquat	306701	90.2%	33246	9.8%	0.0795	0.3399	0.0033	0.9901
dithianon	313853	92.3%	26094	7.7%	0.2707	0.8155	0.0183	2.2926
dodine	326240	96.0%	13707	4.0%	0.3290	0.8142	0.0266	1.8101
epoxiconazole	285996	84.1%	53951	15.9%	0.0476	0.1287	0.0024	0.2920
esfenvalerate	272622	80.2%	67325	19.8%	0.0016	0.0057	0.0001	0.0313
ethofumesate	314189	92.4%	25758	7.6%	0.3609	0.8647	0.0216	1.8027
fenhexamid	310729	91.4%	29218	8.6%	0.1971	1.2199	0.0058	3.5122
phenmedipham	286722	84.3%	53225	15.7%	0.2660	0.6648	0.0159	1.4937
fenoxycarb	326240	96.0%	13707	4.0%	0.0406	0.1168	0.0029	0.3219
fenpropimorph	276138	81.2%	63809	18.8%	0.0942	0.2516	0.0047	0.5609
flonicamid	289810	85.3%	50137	14.7%	0.0117	0.0339	0.0007	0.0934
fluazifop-p-butyl	293922	86.5%	46025	13.5%	0.0220	0.0874	0.0006	0.2948
fluazinam	295125	86.8%	44822	13.2%	0.1924	0.6058	0.0100	1.9038
fludioxonil	300071	88.3%	39876	11.7%	0.0744	0.2469	0.0040	0.6214
fluopicolide	300946	88.5%	39001	11.5%	0.1009	0.2710	0.0057	0.6529
fluoxastrobin	319594	94.0%	20353	6.0%	0.0175	0.0575	0.0009	0.1982
fluroxypyr	259786	76.4%	80161	23.6%	0.0189	0.0473	0.0011	0.1076
folpet	320683	94.3%	19264	5.7%	0.1645	0.6320	0.0090	2.5436
foramsulfuron	297148	87.4%	42799	12.6%	0.0054	0.0126	0.0003	0.0272
fosetyl	326086	95.9%	13861	4.1%	0.1071	0.2916	0.0074	0.6764
glufosinate	323873	95.3%	16074	4.7%	0.0620	0.1651	0.0042	0.3916
glyphosate	257674	75.8%	82273	24.2%	0.3026	0.9830	0.0131	2.6007
haloxyfop-p-methyl	315548	92.8%	24399	7.2%	0.0117	0.0336	0.0007	0.0920
imidacloprid	305779	89.9%	34168	10.1%	0.0160	0.0414	0.0010	0.0989
indoxacarb	308343	90.7%	31604	9.3%	0.0092	0.0245	0.0006	0.0609
iodosulfuron-methyl-sodium	271289	79.8%	68658	20.2%	0.0004	0.0020	0.0000	0.0071
ioxynil octanoate	321487	94.6%	18460	5.4%	0.0744	0.2040	0.0047	0.4723
iprodion	305624	89.9%	34323	10.1%	0.1371	0.3666	0.0089	0.8561
isoproturon	325065	95.6%	14882	4.4%	0.4965	1.2154	0.0282	2.7025
copper oxychloride	333161	98.0%	6786	2.0%	0.0438	0.1169	0.0031	0.2632
kresoxim-methyl	281203	82.7%	58744	17.3%	0.0369	0.1191	0.0018	0.2904
lambda-cyhalothrin	267208	78.6%	72739	21.4%	0.0034	0.0144	0.0001	0.0582
linuron	276692	81.4%	63255	18.6%	0.2715	0.6603	0.0166	1.4673
mancozeb	267966	78.8%	71981	21.2%	2.4232	7.0766	0.1374	18.3878
fluroxypyr-meptyl	230046	67.7%	109901	32.3%	0.0804	0.1784	0.0053	0.3769
fosetyl-al	321568	94.6%	18379	5.4%	0.1345	0.3678	0.0082	0.9262

500m buffer							
Not exposed		Exposed					
n	%	n	%	median	iqr	p1	p90
268732	79.1%	71215	20.9%	0.2959	1.2302	0.0087	3.2705
273677	80.5%	66270	19.5%	0.5514	2.0402	0.0343	6.0155
306688	90.2%	33259	9.8%	0.8368	2.1835	0.0555	5.0749
221620	65.2%	118327	34.8%	0.1532	0.4296	0.0070	1.0010
197324	58.0%	142623	42.0%	0.0054	0.0221	0.0002	0.1210
273416	80.4%	66531	19.6%	1.1487	2.3572	0.0719	4.7821
263365	77.5%	76582	22.5%	0.6697	3.5643	0.0121	9.5547
217753	64.1%	122194	35.9%	0.8757	2.0948	0.0509	4.5501
306688	90.2%	33259	9.8%	0.1029	0.3097	0.0062	0.8662
203621	59.9%	136326	40.1%	0.3214	0.8717	0.0145	1.9526
223494	65.7%	116453	34.3%	0.0317	0.0974	0.0018	0.2735
232980	68.5%	106967	31.5%	0.0785	0.2816	0.0024	0.8746
237912	70.0%	102035	30.0%	0.6762	2.0046	0.0358	5.7720
243600	71.7%	96347	28.3%	0.2360	0.7683	0.0092	1.8473
248689	73.2%	91258	26.8%	0.3516	0.8898	0.0194	1.9730
285646	84.0%	54301	16.0%	0.0372	0.1377	0.0022	0.5156
178896	52.6%	161051	47.4%	0.0718	0.1647	0.0048	0.3672
287422	84.5%	52525	15.5%	0.3462	1.5293	0.0167	5.6460
250392	73.7%	89555	26.3%	0.0184	0.0395	0.0014	0.0854
302924	89.1%	37023	10.9%	0.2631	0.7585	0.0157	1.8269
300090	88.3%	39857	11.7%	0.1478	0.4180	0.0094	1.0497
174402	51.3%	165545	48.7%	1.1434	3.5767	0.0452	8.8230
282391	83.1%	57556	16.9%	0.0346	0.0984	0.0020	0.2652
256611	75.5%	83336	24.5%	0.0389	0.1082	0.0025	0.2783
261368	76.9%	78579	23.1%	0.0219	0.0642	0.0013	0.1661
199523	58.7%	140424	41.3%	0.0017	0.0085	0.0001	0.0250
288178	84.8%	51769	15.2%	0.2029	0.5243	0.0125	1.2195
253391	74.5%	86556	25.5%	0.3684	1.0302	0.0209	2.4244
303463	89.3%	36484	10.7%	1.4334	3.4994	0.0973	8.2334
323391	95.1%	16556	4.9%	0.1037	0.3026	0.0063	0.7267
210462	61.9%	129485	38.1%	0.1213	0.3745	0.0061	0.9026
189300	55.7%	150647	44.3%	0.0131	0.0602	0.0005	0.1996
203802	60.0%	136145	40.0%	0.8818	2.1719	0.0512	4.7475
190371	56.0%	149576	44.0%	8.3090	23.7082	0.3911	59.4721
132340	38.9%	207607	61.1%	0.3198	0.6546	0.0258	1.3411
289578	85.2%	50369	14.8%	0.2640	0.7572	0.0178	2.1761

→

Table S4.2: Number of exposed and unexposed mothers to each of the 139 active ingredients included in the study and measures of center and spread for the exposed mothers in 250m and 500m around mothers' residences. (cont.)

Active ingredient	250m buffer							
	Not exposed		Exposed					
	n	%	n	%	median	iqr	p1	p90
glufosinate-ammonium	304125	89.5%	35822	10.5%	0.0737	0.2025	0.0046	0.4953
mandipropamid	299991	88.2%	39956	11.8%	0.2270	0.5586	0.0136	1.2903
maneb	320683	94.3%	19264	5.7%	0.0505	1.9702	0.0019	12.3480
mcpa	268928	79.1%	71019	20.9%	0.2499	0.6893	0.0132	1.5949
mecoprop-p	303855	89.4%	36092	10.6%	0.0343	0.0891	0.0021	0.2156
mepanipyrim	308479	90.7%	31468	9.3%	0.1268	0.6404	0.0047	1.9035
mesotrione	247871	72.9%	92076	27.1%	0.0412	0.0880	0.0029	0.1849
metalaxyl-m	305053	89.7%	34894	10.3%	0.0408	0.1112	0.0025	0.2722
metamitron	297218	87.4%	42729	12.6%	0.7422	1.9622	0.0441	4.4325
metazachlor	309734	91.1%	30213	8.9%	0.2506	0.6253	0.0175	1.4371
methiocarb	331775	97.6%	8172	2.4%	0.0898	0.2430	0.0060	0.5515
methoxyfenozide	326240	96.0%	13707	4.0%	0.0375	0.0954	0.0030	0.2097
metiram	317661	93.4%	22286	6.6%	0.1417	0.4112	0.0090	1.2428
metoxuron	331775	97.6%	8172	2.4%	0.0986	0.2668	0.0065	0.6055
metribuzin	289774	85.2%	50173	14.8%	0.1358	0.3238	0.0082	0.7055
metsulfuron-methyl	307234	90.4%	32713	9.6%	0.0045	0.0204	0.0002	0.0816
nicosulfuron	247871	72.9%	92076	27.1%	0.0215	0.0453	0.0015	0.0935
pencycuron	299991	88.2%	39956	11.8%	0.0540	0.1620	0.0028	0.6358
pendimethalin	316142	93.0%	23805	7.0%	0.3204	0.9210	0.0188	2.3030
picoxystrobin	333708	98.2%	6239	1.8%	0.0059	0.0148	0.0004	0.0327
pirimicarb	297065	87.4%	42882	12.6%	0.0199	0.0790	0.0010	0.3065
pirimiphos-methyl	332659	97.9%	7288	2.1%	0.1014	0.3020	0.0044	0.7094
prochloraz	319616	94.0%	20331	6.0%	0.0472	0.1961	0.0029	1.0082
procymidone	336606	99.0%	3341	1.0%	0.0531	0.1517	0.0029	0.3512
propamocarb	291541	85.8%	48406	14.2%	0.5038	1.9051	0.0101	4.7872
propiconazole	325989	95.9%	13958	4.1%	0.1106	0.3162	0.0051	0.7900
propyzamide	309734	91.1%	30213	8.9%	0.1578	0.5085	0.0099	1.3483
prosulfocarb	298470	87.8%	41477	12.2%	0.6252	1.8768	0.0304	4.5360
prothioconazole	279497	82.2%	60450	17.8%	0.0888	0.2358	0.0050	0.5901
pymetrozine	322840	95.0%	17107	5.0%	0.0151	0.0545	0.0008	0.1622
pyraclostrobin	265622	78.1%	74325	21.9%	0.0321	0.0878	0.0018	0.2315
pyridate	321487	94.6%	18460	5.4%	0.1275	0.3348	0.0082	0.7502
pyrimethanil	314920	92.6%	25027	7.4%	0.0308	0.2414	0.0009	0.7575
quinoclamine	325989	95.9%	13958	4.1%	0.1055	0.2405	0.0075	0.5721
quizalofop-p-ethyl	331775	97.6%	8172	2.4%	0.0078	0.0211	0.0005	0.0478
rimsulfuron	308902	90.9%	31045	9.1%	0.0007	0.0028	0.0000	0.0146

500m buffer							
Not exposed		Exposed					
n	%	n	%	median	iqr	p1	p90
253719	74.6%	86228	25.4%	0.1927	0.5338	0.0123	1.3493
247232	72.7%	92715	27.3%	0.7715	1.8368	0.0457	3.9328
287422	84.5%	52525	15.5%	0.1060	4.9588	0.0034	26.9584
189563	55.8%	150384	44.2%	0.8461	2.3196	0.0400	5.2937
254438	74.8%	85509	25.2%	0.0859	0.2449	0.0057	0.6373
259300	76.3%	80647	23.7%	0.4148	1.9498	0.0112	5.2854
151114	44.5%	188833	55.5%	0.1423	0.2779	0.0130	0.5883
252429	74.3%	87518	25.7%	0.1044	0.2912	0.0065	0.7285
235033	69.1%	104914	30.9%	2.2727	5.5138	0.1255	12.0110
260991	76.8%	78956	23.2%	0.6472	1.6987	0.0417	3.8965
316237	93.0%	23710	7.0%	0.2289	0.6139	0.0134	1.4006
306688	90.2%	33259	9.8%	0.0966	0.2515	0.0065	0.5888
282208	83.0%	57739	17.0%	0.2955	0.9971	0.0177	3.2207
316237	93.0%	23710	7.0%	0.2513	0.6740	0.0147	1.5377
227769	67.0%	112178	33.0%	0.4555	1.0582	0.0275	2.2514
256575	75.5%	83372	24.5%	0.0109	0.0555	0.0005	0.2315
151114	44.5%	188833	55.5%	0.0744	0.1420	0.0069	0.2971
247232	72.7%	92715	27.3%	0.1580	0.5003	0.0082	2.0125
277199	81.5%	62748	18.5%	0.8360	2.3726	0.0471	5.9775
321272	94.5%	18675	5.5%	0.0129	0.0306	0.0009	0.0690
237973	70.0%	101974	30.0%	0.0506	0.2129	0.0026	0.8319
319724	94.1%	20223	5.9%	0.2294	0.6350	0.0109	1.6627
284498	83.7%	55449	16.3%	0.0967	0.5052	0.0054	2.1850
330879	97.3%	9068	2.7%	0.1180	0.3161	0.0064	0.8254
229960	67.6%	109987	32.4%	1.9234	6.6439	0.0269	15.1029
299914	88.2%	40033	11.8%	0.1990	0.6266	0.0089	1.7200
260991	76.8%	78956	23.2%	0.4460	1.4894	0.0208	3.7363
241991	71.2%	97956	28.8%	2.1570	6.1899	0.0930	13.8983
206876	60.9%	133071	39.1%	0.2991	0.8113	0.0175	1.9647
295618	87.0%	44329	13.0%	0.0402	0.1413	0.0018	0.3963
190096	55.9%	149851	44.1%	0.1202	0.3334	0.0066	0.8054
288178	84.8%	51769	15.2%	0.3472	0.8655	0.0221	1.9291
276503	81.3%	63444	18.7%	0.0572	0.5797	0.0015	1.9549
299914	88.2%	40033	11.8%	0.1841	0.4875	0.0149	1.2944
316237	93.0%	23710	7.0%	0.0199	0.0533	0.0012	0.1215
262954	77.4%	76993	22.6%	0.0016	0.0057	0.0001	0.0350

Table S4.2: Number of exposed and unexposed mothers to each of the 139 active ingredients included in the study and measures of center and spread for the exposed mothers in 250m and 500m around mothers' residences. (cont.)

Active ingredient	250m buffer							
	Not exposed		Exposed					
	n	%	n	%	median	iqr	p1	p90
s-metolachlor	231807	68.2%	108140	31.8%	0.1440	0.3911	0.0083	1.0256
spinosad	321487	94.6%	18460	5.4%	0.0551	0.1743	0.0029	0.4212
spirodiclofen	308343	90.7%	31604	9.3%	0.0081	0.0220	0.0005	0.0539
sulcotrion	294928	86.8%	45019	13.2%	0.0467	0.1154	0.0026	0.2498
tebuconazole	263109	77.4%	76838	22.6%	0.0474	0.1398	0.0026	0.3941
tebufenpyrad	333251	98.0%	6696	2.0%	0.0208	0.0525	0.0015	0.1278
teflubenzuron	331775	97.6%	8172	2.4%	0.0286	0.0773	0.0019	0.1753
tepraloxydim	290640	85.5%	49307	14.5%	0.0048	0.0145	0.0003	0.0393
terbuthylazine	247871	72.9%	92076	27.1%	0.2523	0.5550	0.0169	1.1934
thiacloprid	269612	79.3%	70335	20.7%	0.0310	0.0993	0.0018	0.3296
thiophanate-methyl	307087	90.3%	32860	9.7%	0.1325	0.4344	0.0074	1.2737
tolclofos-methyl	332659	97.9%	7288	2.1%	0.9212	2.3667	0.0485	5.4862
topramezone	247871	72.9%	92076	27.1%	0.0090	0.0189	0.0006	0.0391
triallate	314189	92.4%	25758	7.6%	0.0285	0.0680	0.0017	0.1416
triadimenol	313853	92.3%	26094	7.7%	0.0404	0.1004	0.0031	0.2333
trifloxystrobin	277795	81.7%	62152	18.3%	0.0359	0.0916	0.0022	0.2137
triflusulfuron-methyl	300959	88.5%	38988	11.5%	0.0079	0.0385	0.0004	0.4737
vinclozolin	336606	99.0%	3341	1.0%	0.0044	0.0125	0.0002	0.0289
aclonifen	309977	91.2%	29970	8.8%	0.2239	0.5491	0.0136	1.2034
bixafen	298356	87.8%	41591	12.2%	0.0254	0.0641	0.0015	0.1438
chlorantraniliprole	327443	96.3%	12504	3.7%	0.0175	0.0447	0.0013	0.1022
emamectin benzoate	327443	96.3%	12504	3.7%	0.0060	0.0152	0.0004	0.0348
fenamidone	321568	94.6%	18379	5.4%	0.0126	0.0353	0.0007	0.0893
fenpropidin	317508	93.4%	22439	6.6%	0.0701	0.1703	0.0041	0.3601
imazalil	331788	97.6%	8159	2.4%	0.0226	0.0562	0.0013	0.1232
mesosulfuron-methyl	310418	91.3%	29529	8.7%	0.0053	0.0133	0.0003	0.0297
penconazole	302788	89.1%	37159	10.9%	0.0095	0.0265	0.0006	0.0640
pyroxsulam	310418	91.3%	29529	8.7%	0.0046	0.0114	0.0003	0.0254
spirotetramat	302788	89.1%	37159	10.9%	0.0289	0.1002	0.0015	0.2513
tembotrione	255159	75.1%	84788	24.9%	0.0155	0.0342	0.0010	0.0716
thiamethoxam	311569	91.7%	28378	8.3%	0.0009	0.0034	0.0000	0.0202
thiram	313586	92.2%	26361	7.8%	2.2419	6.5093	0.1330	17.5267
sulphur	299152	88.0%	40795	12.0%	10.0782	26.7393	0.6973	62.3696
TOTAL	**202850**	**59.7%**	**137097**	**40.3%**	**4.4123**	**23.0183**	**0.2150**	**77.1551**

500m buffer							
Not exposed		Exposed					
n	%	n	%	median	iqr	p1	p90
130425	38.4%	209522	61.6%	0.5874	1.5222	0.0395	3.6989
288178	84.8%	51769	15.2%	0.1522	0.4512	0.0076	1.0828
261368	76.9%	78579	23.1%	0.0191	0.0549	0.0012	0.1440
246284	72.4%	93663	27.6%	0.1607	0.3634	0.0101	0.7873
181305	53.3%	158642	46.7%	0.1773	0.5292	0.0090	1.3670
320644	94.3%	19303	5.7%	0.0362	0.1040	0.0027	0.2828
316237	93.0%	23710	7.0%	0.0728	0.1952	0.0043	0.4453
224806	66.1%	115141	33.9%	0.0157	0.0458	0.0008	0.1180
151114	44.5%	188833	55.5%	0.8716	1.7619	0.0775	3.8029
192853	56.7%	147094	43.3%	0.1139	0.3844	0.0053	1.1314
258434	76.0%	81513	24.0%	0.3755	1.2446	0.0174	3.4418
319724	94.1%	20223	5.9%	1.9555	4.9578	0.1150	13.0430
151114	44.5%	188833	55.5%	0.0311	0.0593	0.0029	0.1239
273416	80.4%	66531	19.6%	0.0905	0.1860	0.0057	0.3758
273677	80.5%	66270	19.5%	0.0870	0.2395	0.0062	0.6156
204395	60.1%	135552	39.9%	0.1199	0.2933	0.0072	0.6670
245406	72.2%	94541	27.8%	0.0259	0.1703	0.0012	1.5279
330879	97.3%	9068	2.7%	0.0097	0.0260	0.0005	0.0679
264798	77.9%	75149	22.1%	0.6815	1.6340	0.0407	3.4524
242777	71.4%	97170	28.6%	0.0755	0.1983	0.0047	0.4625
309651	91.1%	30296	8.9%	0.0448	0.1184	0.0029	0.2837
309651	91.1%	30296	8.9%	0.0153	0.0404	0.0010	0.0967
289578	85.2%	50369	14.8%	0.0249	0.0727	0.0017	0.2084
280534	82.5%	59413	17.5%	0.2148	0.4512	0.0134	0.9314
319165	93.9%	20782	6.1%	0.0610	0.1516	0.0035	0.3384
269662	79.3%	70285	20.7%	0.0171	0.0408	0.0011	0.0911
249753	73.5%	90194	26.5%	0.0272	0.0738	0.0017	0.1796
269662	79.3%	70285	20.7%	0.0146	0.0348	0.0010	0.0779
249753	73.5%	90194	26.5%	0.0874	0.2975	0.0033	0.7306
164701	48.4%	175246	51.6%	0.0531	0.1077	0.0044	0.2272
268078	78.9%	71869	21.1%	0.0024	0.0089	0.0001	0.0547
272468	80.2%	67479	19.8%	5.0638	15.2667	0.3128	43.8247
241508	71.0%	98439	29.0%	29.1116	79.1952	1.6727	182.7477
101783	**29.9%**	**238164**	**70.1%**	**22.2086**	**110.5192**	**1.0250**	**312.9242**

S5 Results of regression analyses with increasing level of adjustment to potential confounders

- *Basic model:* adjusted for gestational age (when gestational age or prematurity were not the outcomes), sex of the baby (when child's sex was not the outcome), parity and the complementary donut
- *Intermediate model:* basic model further adjusted for household income, mother's education, mother's marital status, mother's origin, mother's age at birth and mother's job status
- *Full model:* intermediate model further adjusted for degree of urbanization of the residence location and year of birth of the baby

Outcome	Active ingredient (AI)	Buffer size	Basic model		Intermediate model		Full model	
			Risk estimate	p value	Risk estimate	p value	Risk estimate	p value
Gestational age	asulam	50m	β = 0.88 [-1.84, 3.60]	p = 0.525	β = 0.8 [-1.92, 3.51]	p = 0.564	β = 0.74 [-1.97, 3.46]	p = 0.592
		100m	β = 0.04 [-0.53, 0.61]	p = 0.894	β = 0.01 [-0.56, 0.59]	p = 0.960	β = 0 [-0.57, 0.57]	p = 0.997
		250m	β = 0.01 [-0.08, 0.10]	p = 0.848	β = 0 [-0.09, 0.10]	p = 0.919	β = 0 [-0.09, 0.09]	p = 0.962
		500m	β = 0.02 [0.01, 0.04]	p < 0.001	β = 0.02 [0.01, 0.04]	p < 0.001	β = 0.03 [0.01, 0.04]	p < 0.001
	carbetamide	50m	β = 1.18 [-5.66, 8.01]	p = 0.736	β = 0.91 [-5.91, 7.73]	p = 0.793	β = 0.68 [-6.15, 7.50]	p = 0.846
		100m	β = -0.09 [-1.40, 1.21]	p = 0.889	β = -0.15 [-1.45, 1.15]	p = 0.821	β = -0.21 [-1.52, 1.09]	p = 0.749
		250m	β = -0.02 [-0.21, 0.18]	p = 0.860	β = -0.04 [-0.24, 0.16]	p = 0.688	β = -0.05 [-0.25, 0.15]	p = 0.612
		500m	β = 0.03 [-0.01, 0.07]	p = 0.104	β = 0.02 [-0.02, 0.06]	p = 0.329	β = 0.02 [-0.02, 0.06]	p = 0.352
	cyproconazole	50m	β = -40.62 [-298.42, 217.18]	p = 0.757	β = -27.98 [-285.17, 229.20]	p = 0.831	β = -22.97 [-280.15, 234.21]	p = 0.861
		100m	β = 14.29 [-34.17, 62.75]	p = 0.563	β = 16.84 [-31.51, 65.19]	p = 0.495	β = 17.86 [-30.49, 66.21]	p = 0.469
		250m	β = 0.36 [-6.88, 7.59]	p = 0.923	β = 0.6 [-6.61, 7.82]	p = 0.870	β = 0.66 [-6.56, 7.87]	p = 0.859
		500m	β = -1.49 [-3.22, 0.25]	p = 0.093	β = -1.44 [-3.17, 0.29]	p = 0.102	β = -1.84 [-3.57, -0.11]	p = 0.037
	epoxiconazole	50m	β = 39.24 [-3.74, 82.22]	p = 0.074	β = 36.22 [-6.66, 79.09]	p = 0.098	β = 36.13 [-6.75, 79.01]	p = 0.099
		100m	β = 3.57 [-4.74, 11.88]	p = 0.400	β = 2.73 [-5.57, 11.02]	p = 0.519	β = 2.75 [-5.55, 11.04]	p = 0.516
		250m	β = 0.67 [-0.59, 1.94]	p = 0.295	β = 0.66 [-0.60, 1.92]	p = 0.303	β = 0.72 [-0.54, 1.98]	p = 0.265
		500m	β = -0.1 [-0.30, 0.11]	p = 0.363	β = -0.05 [-0.25, 0.16]	p = 0.671	β = 0.05 [-0.16, 0.26]	p = 0.645
	fluroxypyr-meptyl	50m	β = 6.98 [-13.52, 27.48]	p = 0.505	β = 9.65 [-10.80, 30.11]	p = 0.355	β = 9.48 [-10.99, 29.95]	p = 0.364
		100m	β = 4.43 [0.36, 8.51]	p = 0.033	β = 5.03 [0.96, 9.09]	p = 0.015	β = 4.96 [0.89, 9.03]	p = 0.017
		250m	β = 0.96 [0.31, 1.61]	p = 0.004	β = 0.94 [0.29, 1.59]	p = 0.004	β = 0.87 [0.22, 1.52]	p = 0.009
		500m	β = 0.27 [0.15, 0.39]	p < 0.001	β = 0.21 [0.09, 0.33]	p < 0.001	β = 0.06 [-0.06, 0.19]	p = 0.315
	glufosinate	50m	β = -0.55 [-34.09, 32.98]	p = 0.974	β = -0.55 [-34.00, 32.91]	p = 0.974	β = -2.92 [-36.39, 30.54]	p = 0.864
		100m	β = 2.86 [-4.46, 10.18]	p = 0.444	β = 3.07 [-4.23, 10.38]	p = 0.409	β = 2.47 [-4.83, 9.77]	p = 0.507
		250m	β = 0.04 [-1.16, 1.23]	p = 0.951	β = 0.12 [-1.07, 1.32]	p = 0.841	β = 0.08 [-1.12, 1.27]	p = 0.901
		500m	β = 0.23 [0.02, 0.44]	p = 0.035	β = 0.2 [-0.01, 0.42]	p = 0.062	β = 0.31 [0.09, 0.53]	p = 0.005

Outcome	Active ingredient (AI)	Buffer size	Basic model		Intermediate model		Full model	
			Risk estimate	p value	Risk estimate	p value	Risk estimate	p value
Gestational age	glufosinate-ammonium	50m	β = 10.08 [-7.15, 27.31]	p = 0.252	β = 11.01 [-6.18, 28.20]	p = 0.209	β = 11.58 [-5.61, 28.77]	p = 0.187
		100m	β = 1.18 [-2.59, 4.95]	p = 0.541	β = 1.36 [-2.40, 5.11]	p = 0.480	β = 1.52 [-2.24, 5.28]	p = 0.429
		250m	β = -0.15 [-0.78, 0.48]	p = 0.645	β = -0.15 [-0.78, 0.48]	p = 0.648	β = -0.16 [-0.79, 0.47]	p = 0.627
		500m	β = 0.29 [0.18, 0.40]	p < 0.001	β = 0.29 [0.18, 0.40]	p < 0.001	β = 0.23 [0.12, 0.34]	p < 0.001
	linuron	50m	β = 4.5 [-3.21, 12.22]	p = 0.252	β = 4.56 [-3.13, 12.26]	p = 0.245	β = 4.97 [-2.72, 12.67]	p = 0.205
		100m	β = 0.91 [-0.59, 2.40]	p = 0.234	β = 0.89 [-0.60, 2.39]	p = 0.240	β = 0.99 [-0.50, 2.48]	p = 0.194
		250m	β = 0.15 [-0.08, 0.38]	p = 0.193	β = 0.15 [-0.07, 0.38]	p = 0.188	β = 0.16 [-0.06, 0.39]	p = 0.158
		500m	β = -0.09 [-0.14, -0.05]	p < 0.001	β = -0.08 [-0.12, -0.04]	p < 0.001	β = -0.09 [-0.13, -0.05]	p < 0.001
	propiconazole	50m	β = 4.42 [-15.62, 24.46]	p = 0.665	β = 4.93 [-15.06, 24.91]	p = 0.629	β = 4.6 [-15.38, 24.59]	p = 0.652
		100m	β = 1.47 [-2.93, 5.88]	p = 0.513	β = 1.62 [-2.77, 6.01]	p = 0.470	β = 1.59 [-2.81, 5.98]	p = 0.479
		250m	β = 0.09 [-0.66, 0.83]	p = 0.822	β = 0.1 [-0.64, 0.84]	p = 0.787	β = 0.11 [-0.63, 0.84]	p = 0.779
		500m	β = 0.07 [-0.07, 0.20]	p = 0.330	β = 0.06 [-0.07, 0.20]	p = 0.374	β = 0.06 [-0.08, 0.19]	p = 0.425
	thiacloprid	50m	β = 4.3 [-18.83, 27.43]	p = 0.715	β = 3.68 [-19.39, 26.75]	p = 0.755	β = 3.91 [-19.16, 26.98]	p = 0.740
		100m	β = 0.06 [-4.49, 4.62]	p = 0.978	β = -0.09 [-4.63, 4.46]	p = 0.970	β = -0.07 [-4.61, 4.47]	p = 0.976
		250m	β = -0.03 [-0.72, 0.66]	p = 0.936	β = -0.09 [-0.78, 0.60]	p = 0.807	β = -0.09 [-0.78, 0.60]	p = 0.791
		500m	β = 0.19 [0.08, 0.30]	p < 0.001	β = 0.17 [0.05, 0.28]	p = 0.004	β = 0.15 [0.04, 0.27]	p = 0.007
	triadimenol	50m	β = 12.41 [-47.96, 72.79]	p = 0.687	β = 14.54 [-45.71, 74.78]	p = 0.636	β = 15.92 [-44.33, 76.16]	p = 0.605
		100m	β = -2.96 [-16.13, 10.20]	p = 0.659	β = -2.72 [-15.85, 10.42]	p = 0.685	β = -2.48 [-15.61, 10.66]	p = 0.712
		250m	β = -0.94 [-3.06, 1.17]	p = 0.383	β = -0.94 [-3.05, 1.17]	p = 0.381	β = -1.05 [-3.16, 1.06]	p = 0.330
		500m	β = 0.54 [0.16, 0.92]	p = 0.005	β = 0.51 [0.14, 0.89]	p = 0.008	β = 0.37 [-0.01, 0.75]	p = 0.058
	vinclozolin	50m	β = 566.56 [-482.46, 1.6e+03]	p = 0.290	β = 550.08 [-497.37, 1.6e+03]	p = 0.303	β = 486.69 [-560.90, 1.5e+03]	p = 0.363
		100m	β = 60.84 [-162.83, 284.52]	p = 0.594	β = 54.6 [-168.71, 277.91]	p = 0.632	β = 38.81 [-184.54, 262.16]	p = 0.733
		250m	β = 10.33 [-25.59, 46.25]	p = 0.573	β = 9.35 [-26.50, 45.19]	p = 0.609	β = 7.05 [-28.80, 42.90]	p = 0.700
		500m	β = 7.56 [2.30, 12.82]	p = 0.005	β = 7.1 [1.84, 12.35]	p = 0.008	β = 8.4 [3.12, 13.67]	p = 0.002
Birth weight	asulam	50m	β = -86.44 [-168.67, -4.21]	p = 0.039	β = -90.24 [-172.02, -8.47]	p = 0.031	β = -94.3 [-176.06, -12.53]	p = 0.024
		100m	β = -23.11 [-40.43, -5.79]	p = 0.009	β = -24.15 [-41.36, -6.93]	p = 0.006	β = -25.2 [-42.41, -7.99]	p = 0.004
		250m	β = -1.67 [-4.42, 1.07]	p = 0.232	β = -1.86 [-4.59, 0.87]	p = 0.182	β = -2.01 [-4.74, 0.71]	p = 0.148
		500m	β = 0.85 [0.45, 1.25]	p < 0.001	β = 0.8 [0.40, 1.19]	p < 0.001	β = 0.85 [0.45, 1.24]	p < 0.001
	carbetamide	50m	β = 64.54 [-142.09, 271.17]	p = 0.540	β = 54.82 [-150.52, 260.15]	p = 0.601	β = 40.52 [-164.79, 245.83]	p = 0.699
		100m	β = 0.31 [-39.17, 39.80]	p = 0.988	β = -3.1 [-42.33, 36.13]	p = 0.877	β = -6.48 [-45.71, 32.75]	p = 0.746
		250m	β = -2.01 [-7.95, 3.93]	p = 0.508	β = -3.18 [-9.08, 2.73]	p = 0.292	β = -3.65 [-9.55, 2.25]	p = 0.226
		500m	β = -0.86 [-2.00, 0.28]	p = 0.140	β = -1.56 [-2.70, -0.43]	p = 0.007	β = -1.55 [-2.68, -0.41]	p = 0.008
	cyproconazole	50m	β = 4300 [-3500.01, 1.2e+04]	p = 0.280	β = 4400 [-3322.84, 1.2e+04]	p = 0.263	β = 4300 [-3447.78, 1.2e+04]	p = 0.277
		100m	β = 1500 [38.00, 3.0e+03]	p = 0.044	β = 1500 [71.57, 3.0e+03]	p = 0.040	β = 1500 [29.63, 2.9e+03]	p = 0.046
		250m	β = 419.12 [200.52, 637.72]	p < 0.001	β = 405.41 [188.21, 622.61]	p < 0.001	β = 394.08 [176.91, 611.26]	p < 0.001
		500m	β = -29.41 [-81.82, 23.01]	p = 0.271	β = -35.25 [-87.34, 16.83]	p = 0.185	β = -46.49 [-98.66, 5.68]	p = 0.081
	epoxiconazole	50m	β = -37.43 [-1336.53, 1.3e+03]	p = 0.955	β = -94.22 [-1384.67, 1.2e+03]	p = 0.886	β = 76.02 [-1214.41, 1.4e+03]	p = 0.908
		100m	β = -70.76 [-321.98, 180.46]	p = 0.581	β = -92.15 [-341.79, 157.48]	p = 0.469	β = -47.86 [-297.55, 201.82]	p = 0.707
		250m	β = -24.71 [-62.87, 13.45]	p = 0.204	β = -21.95 [-59.89, 15.99]	p = 0.257	β = -13.85 [-51.81, 24.10]	p = 0.474
		500m	β = -10.2 [-16.49, -3.91]	p = 0.001	β = -6.19 [-12.45, 0.06]	p = 0.052	β = -5.61 [-11.91, 0.69]	p = 0.081

→

S5 Results of regression analyses with increasing level of adjustment to potential confounders (cont.)

Outcome	Active ingredient (AI)	Buffer size	Basic model		Intermediate model		Full model	
			Risk estimate	p value	Risk estimate	p value	Risk estimate	p value
Birth weight	fluroxypyr-meptyl	50m	β = 577.42 [-42.28, 1.2e+03]	p = 0.068	β = 559.58 [-56.12, 1.2e+03]	p = 0.075	β = 445.85 [-170.30, 1.1e+03]	p = 0.156
		100m	β = 180.59 [57.41, 303.77]	p = 0.004	β = 174.23 [51.76, 296.69]	p = 0.005	β = 142.72 [20.09, 265.36]	p = 0.023
		250m	β = 52.21 [32.56, 71.87]	p < 0.001	β = 44.26 [24.71, 63.81]	p < 0.001	β = 36.19 [16.59, 55.79]	p < 0.001
		500m	β = 10.63 [7.04, 14.23]	p < 0.001	β = 3.62 [0.04, 7.20]	p = 0.047	β = -1.48 [-5.28, 2.31]	p = 0.443
	glufosinate	50m	β = -478.25 [-1491.86, 535.36]	p = 0.355	β = -422.54 [-1429.95, 584.86]	p = 0.411	β = -360.66 [-1368.05, 646.73]	p = 0.483
		100m	β = -48.31 [-269.55, 172.94]	p = 0.669	β = -27.67 [-247.53, 192.18]	p = 0.805	β = -16.49 [-236.35, 203.38]	p = 0.883
		250m	β = -11.73 [-47.88, 24.42]	p = 0.525	β = -8.94 [-44.86, 26.98]	p = 0.626	β = -9.25 [-45.17, 26.66]	p = 0.614
		500m	β = 4.22 [-2.26, 10.70]	p = 0.202	β = 1.67 [-4.76, 8.11]	p = 0.610	β = -2.97 [-9.51, 3.58]	p = 0.375
	glufosinate-ammonium	50m	β = 192.19 [-328.56, 712.95]	p = 0.469	β = 203.82 [-313.51, 721.14]	p = 0.440	β = 165.48 [-351.80, 682.77]	p = 0.531
		100m	β = 44.57 [-69.36, 158.50]	p = 0.443	β = 46.81 [-66.36, 159.98]	p = 0.418	β = 38.72 [-74.44, 151.88]	p = 0.502
		250m	β = -9.34 [-28.43, 9.74]	p = 0.337	β = -10.73 [-29.69, 8.23]	p = 0.267	β = -12.7 [-31.66, 6.26]	p = 0.189
		500m	β = -4.74 [-8.10, -1.37]	p = 0.006	β = -5.66 [-9.01, -2.32]	p < 0.001	β = -5.57 [-8.94, -2.19]	p = 0.001
	linuron	50m	β = 264.17 [31.06, 497.28]	p = 0.026	β = 242.13 [10.54, 473.72]	p = 0.040	β = 256.85 [25.27, 488.43]	p = 0.030
		100m	β = 67.96 [22.75, 113.17]	p = 0.003	β = 62.72 [17.79, 107.65]	p = 0.006	β = 66.05 [21.12, 110.98]	p = 0.004
		250m	β = 13.8 [6.95, 20.65]	p < 0.001	β = 13.04 [6.23, 19.84]	p < 0.001	β = 13.24 [6.43, 20.04]	p < 0.001
		500m	β = 2.85 [1.56, 4.14]	p < 0.001	β = 3.25 [1.97, 4.54]	p < 0.001	β = 2.54 [1.25, 3.84]	p < 0.001
	propiconazole	50m	β = 223.11 [-382.59, 828.82]	p = 0.470	β = 251.41 [-350.23, 853.06]	p = 0.413	β = 207.34 [-394.19, 808.87]	p = 0.499
		100m	β = 28.24 [-104.94, 161.42]	p = 0.678	β = 39.32 [-92.97, 171.60]	p = 0.560	β = 31.87 [-100.39, 164.12]	p = 0.637
		250m	β = -19.46 [-41.85, 2.92]	p = 0.088	β = -17.88 [-40.13, 4.36]	p = 0.115	β = -17.94 [-40.18, 4.29]	p = 0.114
		500m	β = -7.37 [-11.49, -3.24]	p < 0.001	β = -7.18 [-11.28, -3.08]	p < 0.001	β = -5.61 [-9.72, -1.50]	p = 0.007
	thiacloprid	50m	β = -55.25 [-754.30, 643.80]	p = 0.877	β = -48.45 [-743.19, 646.29]	p = 0.891	β = -93.41 [-788.02, 601.20]	p = 0.792
		100m	β = -24.6 [-162.26, 113.06]	p = 0.726	β = -27.67 [-164.45, 109.11]	p = 0.692	β = -39.31 [-176.07, 97.45]	p = 0.573
		250m	β = -6.41 [-27.29, 14.48]	p = 0.548	β = -8.71 [-29.47, 12.04]	p = 0.411	β = -10.28 [-31.03, 10.48]	p = 0.332
		500m	β = 0.95 [-2.44, 4.35]	p = 0.582	β = -0.07 [-3.44, 3.31]	p = 0.968	β = 1.07 [-2.31, 4.45]	p = 0.535
	triadimenol	50m	β = 167.95 [-1656.95, 2.0e+03]	p = 0.857	β = 98.77 [-1714.93, 1.9e+03]	p = 0.915	β = 103.29 [-1710.01, 1.9e+03]	p = 0.911
		100m	β = 86.62 [-311.36, 484.60]	p = 0.670	β = 50.25 [-345.17, 445.67]	p = 0.803	β = 39.86 [-355.48, 435.20]	p = 0.843
		250m	β = 36.79 [-27.13, 100.71]	p = 0.259	β = 26.49 [-37.03, 90.01]	p = 0.414	β = 17.97 [-45.55, 81.49]	p = 0.579
		500m	β = 4.51 [-6.87, 15.89]	p = 0.437	β = -1.8 [-13.11, 9.52]	p = 0.756	β = -9.01 [-20.41, 2.38]	p = 0.121
	vinclozolin	50m	β = -26151.45 [-57858.18, 5.6e+03]	p = 0.106	β = -27403.38 [-58943.39, 4.1e+03]	p = 0.089	β = -27228.6 [-58766.56, 4.3e+03]	p = 0.091
		100m	β = -7522.43 [-14282.95, -761.91]	p = 0.029	β = -7934.38 [-14657.68, -1211.08]	p = 0.021	β = -7973.91 [-14697.14, -1250.67]	p = 0.020
		250m	β = -419.56 [-1505.15, 666.04]	p = 0.449	β = -473.29 [-1552.62, 606.03]	p = 0.390	β = -484.91 [-1564.13, 594.32]	p = 0.379
		500m	β = 311.82 [152.73, 470.92]	p < 0.001	β = 278.74 [120.63, 436.85]	p < 0.001	β = 257.78 [99.19, 416.37]	p = 0.001

Outcome	Active ingredient (AI)	Buffer size	Basic model		Intermediate model		Full model	
			Risk estimate	p value	Risk estimate	p value	Risk estimate	p value
Perinatal mortality	asulam	50m	NA	NA	NA	NA	NA	NA
		100m	OR = 0.92 [0.43, 2.00]	p = 0.839	OR = 0.94 [0.43, 2.08]	p = 0.888	OR = 0.91 [0.41, 2.03]	p = 0.822
		250m	OR = 0.96 [0.85, 1.08]	p = 0.526	OR = 0.95 [0.84, 1.07]	p = 0.400	OR = 0.95 [0.84, 1.07]	p = 0.366
		500m	OR = 0.98 [0.96, 1.00]	p = 0.065	OR = 0.98 [0.96, 1.00]	p = 0.034	OR = 0.98 [0.96, 1.00]	p = 0.030
	carbetamide	50m	NA	NA	NA	NA	NA	NA
		100m	OR = 0.58 [0.13, 2.63]	p = 0.483	OR = 0.56 [0.12, 2.54]	p = 0.452	OR = 0.53 [0.11, 2.43]	p = 0.412
		250m	OR = 0.84 [0.67, 1.04]	p = 0.112	OR = 0.82 [0.66, 1.02]	p = 0.080	OR = 0.82 [0.65, 1.02]	p = 0.075
		500m	OR = 0.98 [0.94, 1.02]	p = 0.382	OR = 0.98 [0.94, 1.02]	p = 0.239	OR = 0.98 [0.94, 1.02]	p = 0.234
	cyproconazole	50m	OR = 0 [0.00, 2.4e+49]	p = 0.245	OR = 0 [0.00, 1.4e+46]	p = 0.219	OR = 0 [0.00, 5.4e+43]	p = 0.204
		100m	OR = 0 [0.00, 3.7e+09]	p = 0.275	OR = 0 [0.00, 1.2e+09]	p = 0.251	OR = 0 [0.00, 2.0e+08]	p = 0.222
		250m	OR = 0.08 [0.00, 106.72]	p = 0.496	OR = 0.09 [0.00, 129.13]	p = 0.514	OR = 0.05 [0.00, 82.77]	p = 0.436
		500m	OR = 0.24 [0.04, 1.30]	p = 0.098	OR = 0.26 [0.05, 1.41]	p = 0.118	OR = 0.27 [0.05, 1.49]	p = 0.132
	epoxiconazole	50m	OR = 440000 [0.00, 5.7e+21]	p = 0.493	OR = 8e+06 [0.00, 2.1e+23]	p = 0.410	OR = 3.9e+07 [0.00, 7.7e+23]	p = 0.361
		100m	OR = 20.72 [0.01, 4.5e+04]	p = 0.439	OR = 27.75 [0.01, 7.0e+04]	p = 0.405	OR = 39.1 [0.02, 9.3e+04]	p = 0.355
		250m	OR = 1.09 [0.32, 3.69]	p = 0.892	OR = 1.03 [0.30, 3.56]	p = 0.962	OR = 1.1 [0.32, 3.80]	p = 0.877
		500m	OR = 1.16 [0.95, 1.41]	p = 0.148	OR = 1.12 [0.92, 1.37]	p = 0.257	OR = 1.09 [0.89, 1.33]	p = 0.428
	fluroxypyr-meptyl	50m	OR = 358.86 [0.00, 1.1e+11]	p = 0.556	OR = 6.71 [0.00, 4.0e+09]	p = 0.853	OR = 2.39 [0.00, 1.3e+09]	p = 0.932
		100m	OR = 0.76 [0.01, 48.31]	p = 0.895	OR = 0.25 [0.00, 17.57]	p = 0.526	OR = 0.2 [0.00, 13.45]	p = 0.452
		250m	OR = 1.13 [0.58, 2.17]	p = 0.723	OR = 0.96 [0.50, 1.85]	p = 0.900	OR = 0.91 [0.47, 1.76]	p = 0.785
		500m	OR = 1 [0.89, 1.12]	p = 0.963	OR = 1.01 [0.90, 1.13]	p = 0.890	OR = 1.03 [0.91, 1.16]	p = 0.692
	glufosinate	50m	OR = 1e+12 [189.20, 5.8e+21]	p = 0.016	OR = 4e+12 [1.1e+03, 1.5e+22]	p = 0.010	OR = 1.8e+13 [3.9e+03, 8.2e+22]	p = 0.007
		100m	OR = 678.56 [2.72, 1.7e+05]	p = 0.021	OR = 748.18 [2.60, 2.2e+05]	p = 0.022	OR = 1100 [3.99, 3.2e+05]	p = 0.015
		250m	OR = 1.29 [0.38, 4.42]	p = 0.684	OR = 1.17 [0.34, 4.06]	p = 0.800	OR = 1.2 [0.34, 4.16]	p = 0.780
		500m	OR = 1.02 [0.81, 1.28]	p = 0.844	OR = 1.04 [0.82, 1.31]	p = 0.761	OR = 0.93 [0.73, 1.19]	p = 0.568
	glufosinate-ammonium	50m	OR = 0.31 [0.00, 3.5e+06]	p = 0.887	OR = 1.78 [0.00, 1.2e+07]	p = 0.943	OR = 0.44 [0.00, 4.3e+06]	p = 0.920
		100m	OR = 1.38 [0.03, 66.50]	p = 0.871	OR = 1.22 [0.02, 61.38]	p = 0.921	OR = 0.9 [0.02, 46.39]	p = 0.959
		250m	OR = 0.99 [0.52, 1.87]	p = 0.976	OR = 0.93 [0.49, 1.78]	p = 0.837	OR = 0.91 [0.48, 1.72]	p = 0.766
		500m	OR = 0.9 [0.80, 1.02]	p = 0.112	OR = 0.88 [0.78, 1.00]	p = 0.058	OR = 0.91 [0.80, 1.04]	p = 0.158
	linuron	50m	OR = 3.21 [0.00, 4.5e+03]	p = 0.753	OR = 1.37 [0.00, 2.2e+03]	p = 0.934	OR = 1.61 [0.00, 2.4e+03]	p = 0.898
		100m	OR = 1.17 [0.30, 4.62]	p = 0.820	OR = 1.01 [0.26, 4.02]	p = 0.985	OR = 1.06 [0.27, 4.19]	p = 0.930
		250m	OR = 1.04 [0.84, 1.28]	p = 0.719	OR = 1.03 [0.83, 1.26]	p = 0.808	OR = 1.03 [0.84, 1.27]	p = 0.766
		500m	OR = 1 [0.96, 1.04]	p = 0.915	OR = 1 [0.96, 1.04]	p = 0.907	OR = 1 [0.96, 1.04]	p = 0.981
	propiconazole	50m	OR = 5.98 [0.00, 5.5e+09]	p = 0.865	OR = 51.06 [0.00, 3.5e+10]	p = 0.705	OR = 22.65 [0.00, 2.1e+10]	p = 0.767
		100m	OR = 20.61 [0.32, 1.3e+03]	p = 0.156	OR = 20.34 [0.29, 1.4e+03]	p = 0.165	OR = 17.32 [0.24, 1.2e+03]	p = 0.191
		250m	OR = 1.98 [0.95, 4.12]	p = 0.069	OR = 1.9 [0.90, 3.98]	p = 0.090	OR = 1.86 [0.89, 3.89]	p = 0.100
		500m	OR = 0.89 [0.76, 1.05]	p = 0.171	OR = 0.88 [0.75, 1.04]	p = 0.139	OR = 0.9 [0.77, 1.06]	p = 0.225

S5 Results of regression analyses with increasing level of adjustment to potential confounders (cont.)

Outcome	Active ingredient (AI)	Buffer size	Basic model		Intermediate model		Full model	
			Risk estimate	p value	Risk estimate	p value	Risk estimate	p value
Perinatal mortality	thiacloprid	50m	OR = 0 [0.00, 1.2e+07]	p = 0.367	OR = 0 [0.00, 2.0e+07]	p = 0.370	OR = 0 [0.00, 7.1e+06]	p = 0.332
		100m	OR = 0.02 [0.00, 6.00]	p = 0.175	OR = 0.01 [0.00, 5.08]	p = 0.155	OR = 0.01 [0.00, 3.86]	p = 0.130
		250m	OR = 0.45 [0.20, 1.02]	p = 0.055	OR = 0.39 [0.17, 0.90]	p = 0.026	OR = 0.39 [0.17, 0.88]	p = 0.024
		500m	OR = 0.9 [0.79, 1.03]	p = 0.124	OR = 0.88 [0.77, 1.00]	p = 0.050	OR = 0.89 [0.78, 1.02]	p = 0.091
	triadimenol	50m	OR = 0 [0.00, 3.5e+15]	p = 0.582	OR = 0 [0.00, 1.7e+13]	p = 0.433	OR = 0 [0.00, 8.4e+12]	p = 0.412
		100m	OR = 0 [0.00, 48.67]	p = 0.201	OR = 0 [0.00, 21.44]	p = 0.154	OR = 0 [0.00, 15.25]	p = 0.137
		250m	OR = 0.22 [0.03, 1.78]	p = 0.156	OR = 0.21 [0.03, 1.73]	p = 0.149	OR = 0.21 [0.03, 1.67]	p = 0.140
		500m	OR = 0.92 [0.63, 1.33]	p = 0.652	OR = 0.87 [0.60, 1.28]	p = 0.484	OR = 0.87 [0.59, 1.28]	p = 0.490
	vinclozolin	50m	NA	NA	NA	NA	NA	NA
		100m	NA	NA	NA	NA	NA	NA
		250m	OR = 0.01 [0.00, 1.0e+20]	p = 0.847	OR = 0.01 [0.00, 6.2e+19]	p = 0.845	OR = 0 [0.00, 1.4e+19]	p = 0.774
		500m	OR = 0 [0.00, 5.54]	p = 0.110	OR = 0 [0.00, 3.12]	p = 0.085	OR = 0 [0.00, 0.95]	p = 0.049
Child's sex	asulam	50m	OR = 0.74 [0.50, 1.09]	p = 0.122	OR = 0.74 [0.50, 1.09]	p = 0.126	OR = 0.74 [0.50, 1.09]	p = 0.126
		100m	OR = 0.94 [0.87, 1.02]	p = 0.169	OR = 0.95 [0.87, 1.03]	p = 0.176	OR = 0.95 [0.87, 1.03]	p = 0.177
		250m	OR = 1 [0.99, 1.01]	p = 0.931	OR = 1 [0.99, 1.01]	p = 0.916	OR = 1 [0.99, 1.01]	p = 0.913
		500m	OR = 1 [1.00, 1.00]	p = 0.383	OR = 1 [1.00, 1.00]	p = 0.397	OR = 1 [1.00, 1.00]	p = 0.425
	carbetamide	50m	OR = 1.15 [0.43, 3.03]	p = 0.782	OR = 1.15 [0.43, 3.03]	p = 0.781	OR = 1.15 [0.44, 3.05]	p = 0.771
		100m	OR = 1.1 [0.91, 1.32]	p = 0.330	OR = 1.1 [0.91, 1.32]	p = 0.330	OR = 1.1 [0.91, 1.32]	p = 0.322
		250m	OR = 1.01 [0.99, 1.04]	p = 0.315	OR = 1.01 [0.99, 1.04]	p = 0.320	OR = 1.01 [0.99, 1.04]	p = 0.310
		500m	OR = 1 [0.99, 1.00]	p = 0.668	OR = 1 [0.99, 1.00]	p = 0.624	OR = 1 [0.99, 1.00]	p = 0.624
	cyproconazole	50m	OR = 298.89 [0.00, 2.3e+18]	p = 0.760	OR = 215.35 [0.00, 1.7e+18]	p = 0.773	OR = 191.8 [0.00, 1.5e+18]	p = 0.778
		100m	OR = 0.49 [0.00, 470.19]	p = 0.837	OR = 0.46 [0.00, 448.35]	p = 0.826	OR = 0.45 [0.00, 437.60]	p = 0.820
		250m	OR = 1.04 [0.37, 2.90]	p = 0.943	OR = 1.03 [0.37, 2.88]	p = 0.950	OR = 1.03 [0.37, 2.88]	p = 0.950
		500m	OR = 0.99 [0.77, 1.26]	p = 0.918	OR = 0.99 [0.77, 1.26]	p = 0.906	OR = 1 [0.78, 1.27]	p = 0.972
	epoxiconazole	50m	OR = 0.43 [0.00, 194.73]	p = 0.789	OR = 0.44 [0.00, 196.36]	p = 0.791	OR = 0.45 [0.00, 200.74]	p = 0.796
		100m	OR = 1.11 [0.34, 3.60]	p = 0.867	OR = 1.1 [0.34, 3.58]	p = 0.873	OR = 1.11 [0.34, 3.60]	p = 0.868
		250m	OR = 0.99 [0.83, 1.18]	p = 0.894	OR = 0.99 [0.83, 1.18]	p = 0.895	OR = 0.99 [0.83, 1.18]	p = 0.887
		500m	OR = 1 [0.97, 1.03]	p = 0.837	OR = 1 [0.97, 1.03]	p = 0.891	OR = 1 [0.97, 1.03]	p = 0.758
	fluroxypyr-meptyl	50m	OR = 0.18 [0.01, 3.24]	p = 0.242	OR = 0.17 [0.01, 3.09]	p = 0.230	OR = 0.17 [0.01, 3.11]	p = 0.231
		100m	OR = 0.67 [0.37, 1.19]	p = 0.168	OR = 0.66 [0.37, 1.17]	p = 0.156	OR = 0.66 [0.37, 1.18]	p = 0.157
		250m	OR = 0.98 [0.89, 1.07]	p = 0.632	OR = 0.97 [0.89, 1.07]	p = 0.588	OR = 0.98 [0.89, 1.07]	p = 0.616
		500m	OR = 1 [0.98, 1.02]	p = 0.869	OR = 1 [0.98, 1.02]	p = 0.989	OR = 1 [0.99, 1.02]	p = 0.686
	glufosinate	50m	OR = 1.53 [0.01, 178.13]	p = 0.862	OR = 1.52 [0.01, 177.13]	p = 0.864	OR = 1.64 [0.01, 192.22]	p = 0.838
		100m	OR = 0.74 [0.26, 2.10]	p = 0.573	OR = 0.74 [0.26, 2.09]	p = 0.568	OR = 0.75 [0.27, 2.13]	p = 0.594
		250m	OR = 0.82 [0.69, 0.97]	p = 0.019	OR = 0.82 [0.69, 0.97]	p = 0.019	OR = 0.82 [0.69, 0.97]	p = 0.020
		500m	OR = 0.99 [0.96, 1.02]	p = 0.502	OR = 0.99 [0.96, 1.02]	p = 0.493	OR = 0.99 [0.96, 1.02]	p = 0.359

Outcome	Active ingredient (AI)	Buffer size	Basic model		Intermediate model		Full model	
			Risk estimate	p value	Risk estimate	p value	Risk estimate	p value
Child's sex	glufosinate-ammonium	50m	OR = 1.08 [0.09, 12.47]	p = 0.950	OR = 1.07 [0.09, 12.32]	p = 0.958	OR = 1.04 [0.09, 12.04]	p = 0.973
		100m	OR = 0.89 [0.52, 1.52]	p = 0.666	OR = 0.89 [0.52, 1.51]	p = 0.659	OR = 0.88 [0.52, 1.51]	p = 0.643
		250m	OR = 1.06 [0.97, 1.16]	p = 0.210	OR = 1.06 [0.97, 1.16]	p = 0.213	OR = 1.06 [0.97, 1.16]	p = 0.212
		500m	OR = 1 [0.98, 1.01]	p = 0.753	OR = 1 [0.98, 1.01]	p = 0.708	OR = 1 [0.98, 1.01]	p = 0.867
	linuron	50m	OR = 0.68 [0.23, 2.03]	p = 0.491	OR = 0.67 [0.22, 2.00]	p = 0.473	OR = 0.66 [0.22, 1.99]	p = 0.464
		100m	OR = 0.91 [0.73, 1.12]	p = 0.370	OR = 0.9 [0.73, 1.12]	p = 0.350	OR = 0.9 [0.73, 1.12]	p = 0.341
		250m	OR = 1 [0.97, 1.03]	p = 0.844	OR = 1 [0.96, 1.03]	p = 0.814	OR = 1 [0.96, 1.03]	p = 0.807
		500m	OR = 1 [1.00, 1.01]	p = 0.618	OR = 1 [1.00, 1.01]	p = 0.597	OR = 1 [1.00, 1.01]	p = 0.545
	propiconazole	50m	OR = 3.73 [0.22, 64.23]	p = 0.365	OR = 3.77 [0.22, 64.93]	p = 0.361	OR = 3.79 [0.22, 65.29]	p = 0.359
		100m	OR = 1.29 [0.69, 2.42]	p = 0.422	OR = 1.29 [0.69, 2.42]	p = 0.419	OR = 1.29 [0.69, 2.42]	p = 0.419
		250m	OR = 1.12 [1.01, 1.25]	p = 0.030	OR = 1.12 [1.01, 1.25]	p = 0.030	OR = 1.12 [1.01, 1.25]	p = 0.030
		500m	OR = 0.99 [0.97, 1.01]	p = 0.287	OR = 0.99 [0.97, 1.01]	p = 0.276	OR = 0.99 [0.97, 1.01]	p = 0.291
	thiacloprid	50m	OR = 0.17 [0.01, 4.46]	p = 0.284	OR = 0.17 [0.01, 4.57]	p = 0.291	OR = 0.17 [0.01, 4.51]	p = 0.288
		100m	OR = 0.81 [0.43, 1.55]	p = 0.531	OR = 0.82 [0.43, 1.56]	p = 0.543	OR = 0.82 [0.43, 1.56]	p = 0.539
		250m	OR = 1.01 [0.91, 1.11]	p = 0.907	OR = 1.01 [0.91, 1.11]	p = 0.897	OR = 1.01 [0.91, 1.11]	p = 0.896
		500m	OR = 1 [0.98, 1.01]	p = 0.724	OR = 1 [0.98, 1.01]	p = 0.699	OR = 1 [0.98, 1.01]	p = 0.746
	triadimenol	50m	OR = 0.02 [0.00, 105.21]	p = 0.370	OR = 0.02 [0.00, 96.61]	p = 0.359	OR = 0.02 [0.00, 93.12]	p = 0.355
		100m	OR = 0.36 [0.05, 2.31]	p = 0.279	OR = 0.35 [0.05, 2.28]	p = 0.271	OR = 0.35 [0.05, 2.26]	p = 0.268
		250m	OR = 0.79 [0.59, 1.07]	p = 0.134	OR = 0.79 [0.59, 1.07]	p = 0.129	OR = 0.79 [0.59, 1.07]	p = 0.134
		500m	OR = 1.02 [0.97, 1.08]	p = 0.477	OR = 1.02 [0.97, 1.07]	p = 0.513	OR = 1.02 [0.97, 1.08]	p = 0.437
	vinclozolin	50m	OR = 0 [0.00, 4.2e+48]	p = 0.625	OR = 0 [0.00, 7.4e+48]	p = 0.630	OR = 0 [0.00, 4.1e+49]	p = 0.646
		100m	OR = 2.15 [0.00, 1.3e+14]	p = 0.962	OR = 2.65 [0.00, 1.6e+14]	p = 0.952	OR = 4.03 [0.00, 2.5e+14]	p = 0.931
		250m	OR = 28.24 [0.17, 4.7e+03]	p = 0.200	OR = 28.67 [0.17, 4.8e+03]	p = 0.198	OR = 30.32 [0.18, 5.0e+03]	p = 0.191
		500m	OR = 1.56 [0.74, 3.30]	p = 0.241	OR = 1.54 [0.73, 3.26]	p = 0.254	OR = 1.49 [0.70, 3.15]	p = 0.300
Prematurity	asulam	50m	OR = 0.65 [0.23, 1.79]	p = 0.401	OR = 0.66 [0.24, 1.83]	p = 0.424	OR = 0.67 [0.24, 1.86]	p = 0.445
		100m	OR = 0.95 [0.78, 1.16]	p = 0.609	OR = 0.95 [0.78, 1.16]	p = 0.645	OR = 0.96 [0.79, 1.17]	p = 0.677
		250m	OR = 0.99 [0.96, 1.02]	p = 0.611	OR = 0.99 [0.96, 1.02]	p = 0.644	OR = 0.99 [0.96, 1.02]	p = 0.678
		500m	OR = 1 [0.99, 1.00]	p = 0.082	OR = 1 [0.99, 1.00]	p = 0.088	OR = 1 [0.99, 1.00]	p = 0.044
	carbetamide	50m	OR = 1.11 [0.13, 9.37]	p = 0.920	OR = 1.16 [0.14, 9.74]	p = 0.893	OR = 1.19 [0.14, 10.06]	p = 0.874
		100m	OR = 1.18 [0.80, 1.73]	p = 0.410	OR = 1.19 [0.81, 1.75]	p = 0.383	OR = 1.2 [0.81, 1.77]	p = 0.361
		250m	OR = 1.03 [0.97, 1.09]	p = 0.347	OR = 1.03 [0.97, 1.09]	p = 0.273	OR = 1.04 [0.98, 1.10]	p = 0.243
		500m	OR = 0.99 [0.98, 1.00]	p = 0.268	OR = 1 [0.98, 1.01]	p = 0.438	OR = 1 [0.99, 1.01]	p = 0.518
	cyproconazole	50m	OR = 2.1e+22 [0.00, 8.1e+54]	p = 0.179	OR = 1.7e+21 [0.00, 7.6e+53]	p = 0.203	OR = 6.5e+19 [0.00, 3.4e+52]	p = 0.235
		100m	OR = 20.74 [0.00, 4.0e+07]	p = 0.681	OR = 11.81 [0.00, 2.3e+07]	p = 0.739	OR = 5.42 [0.00, 1.1e+07]	p = 0.820
		250m	OR = 0.98 [0.11, 8.76]	p = 0.986	OR = 0.89 [0.10, 7.99]	p = 0.918	OR = 0.79 [0.09, 7.17]	p = 0.838
		500m	OR = 0.96 [0.58, 1.61]	p = 0.883	OR = 0.94 [0.56, 1.58]	p = 0.824	OR = 1.05 [0.62, 1.75]	p = 0.866
	epoxiconazole	50m	OR = 0 [0.00, 75.83]	p = 0.176	OR = 0 [0.00, 89.30]	p = 0.183	OR = 0 [0.00, 226.44]	p = 0.229
		100m	OR = 0.5 [0.04, 6.63]	p = 0.599	OR = 0.54 [0.04, 7.16]	p = 0.637	OR = 0.67 [0.05, 8.90]	p = 0.759
		250m	OR = 0.83 [0.56, 1.22]	p = 0.341	OR = 0.83 [0.56, 1.22]	p = 0.337	OR = 0.84 [0.57, 1.24]	p = 0.384
		500m	OR = 1.07 [1.01, 1.14]	p = 0.028	OR = 1.06 [1.00, 1.13]	p = 0.047	OR = 1.03 [0.97, 1.09]	p = 0.400

→

S5 Results of regression analyses with increasing level of adjustment to potential confounders (cont.)

Outcome	Active ingredient (AI)	Buffer size	Basic model		Intermediate model		Full model	
			Risk estimate	p value	Risk estimate	p value	Risk estimate	p value
Prematurity	fluroxypyr-meptyl	50m	OR = 0.79 [0.00, 484.22]	p = 0.944	OR = 0.45 [0.00, 283.00]	p = 0.809	OR = 0.23 [0.00, 143.16]	p = 0.653
		100m	OR = 0.6 [0.17, 2.15]	p = 0.431	OR = 0.53 [0.15, 1.90]	p = 0.327	OR = 0.45 [0.12, 1.61]	p = 0.217
		250m	OR = 0.85 [0.70, 1.04]	p = 0.118	OR = 0.85 [0.70, 1.04]	p = 0.115	OR = 0.84 [0.68, 1.02]	p = 0.084
		500m	OR = 0.92 [0.89, 0.96]	p < 0.001	OR = 0.93 [0.90, 0.97]	p < 0.001	OR = 0.96 [0.93, 1.00]	p = 0.057
	glufosinate	50m	OR = 0.39 [0.00, 6.4e+04]	p = 0.879	OR = 0.46 [0.00, 7.4e+04]	p = 0.899	OR = 1.62 [0.00, 2.7e+05]	p = 0.937
		100m	OR = 0.19 [0.01, 2.67]	p = 0.220	OR = 0.19 [0.01, 2.64]	p = 0.216	OR = 0.25 [0.02, 3.56]	p = 0.309
		250m	OR = 0.82 [0.56, 1.20]	p = 0.308	OR = 0.81 [0.56, 1.19]	p = 0.287	OR = 0.83 [0.56, 1.21]	p = 0.327
		500m	OR = 0.96 [0.90, 1.02]	p = 0.205	OR = 0.96 [0.90, 1.03]	p = 0.267	OR = 0.9 [0.84, 0.97]	p = 0.004
	glufosinate-ammonium	50m	OR = 0.01 [0.00, 8.02]	p = 0.184	OR = 0.01 [0.00, 6.24]	p = 0.159	OR = 0.01 [0.00, 4.02]	p = 0.123
		100m	OR = 0.55 [0.15, 2.03]	p = 0.372	OR = 0.53 [0.14, 1.95]	p = 0.337	OR = 0.47 [0.13, 1.75]	p = 0.263
		250m	OR = 1.02 [0.83, 1.24]	p = 0.878	OR = 1.01 [0.83, 1.23]	p = 0.892	OR = 1.01 [0.83, 1.23]	p = 0.944
		500m	OR = 0.95 [0.92, 0.99]	p = 0.007	OR = 0.95 [0.92, 0.99]	p = 0.008	OR = 0.98 [0.94, 1.01]	p = 0.161
	linuron	50m	OR = 0.5 [0.05, 5.44]	p = 0.566	OR = 0.46 [0.04, 5.12]	p = 0.531	OR = 0.41 [0.04, 4.53]	p = 0.466
		100m	OR = 0.89 [0.56, 1.40]	p = 0.604	OR = 0.87 [0.55, 1.38]	p = 0.564	OR = 0.85 [0.54, 1.34]	p = 0.480
		250m	OR = 0.98 [0.92, 1.05]	p = 0.638	OR = 0.98 [0.92, 1.05]	p = 0.563	OR = 0.98 [0.91, 1.04]	p = 0.467
		500m	OR = 1.01 [1.00, 1.02]	p = 0.067	OR = 1.01 [1.00, 1.02]	p = 0.148	OR = 1.01 [1.00, 1.02]	p = 0.119
	propiconazole	50m	OR = 0.16 [0.00, 263.94]	p = 0.624	OR = 0.14 [0.00, 252.82]	p = 0.610	OR = 0.12 [0.00, 222.08]	p = 0.586
		100m	OR = 0.53 [0.12, 2.35]	p = 0.404	OR = 0.52 [0.12, 2.34]	p = 0.397	OR = 0.51 [0.11, 2.26]	p = 0.373
		250m	OR = 0.97 [0.77, 1.21]	p = 0.792	OR = 0.97 [0.78, 1.22]	p = 0.794	OR = 0.97 [0.77, 1.21]	p = 0.783
		500m	OR = 0.99 [0.95, 1.03]	p = 0.644	OR = 0.99 [0.95, 1.03]	p = 0.640	OR = 1 [0.96, 1.04]	p = 0.991
	thiacloprid	50m	OR = 0.5 [0.00, 866.96]	p = 0.855	OR = 0.55 [0.00, 962.94]	p = 0.874	OR = 0.43 [0.00, 759.17]	p = 0.825
		100m	OR = 1.52 [0.38, 6.12]	p = 0.555	OR = 1.55 [0.39, 6.26]	p = 0.536	OR = 1.48 [0.37, 5.97]	p = 0.583
		250m	OR = 1.1 [0.90, 1.36]	p = 0.353	OR = 1.12 [0.91, 1.38]	p = 0.302	OR = 1.11 [0.90, 1.37]	p = 0.314
		500m	OR = 0.96 [0.93, 1.00]	p = 0.041	OR = 0.97 [0.93, 1.00]	p = 0.064	OR = 0.98 [0.94, 1.01]	p = 0.207
	triadimenol	50m	OR = 0 [0.00, 4.3e+04]	p = 0.330	OR = 0 [0.00, 1.9e+04]	p = 0.294	OR = 0 [0.00, 1.1e+04]	p = 0.270
		100m	OR = 0.87 [0.01, 59.99]	p = 0.947	OR = 0.75 [0.01, 51.87]	p = 0.893	OR = 0.65 [0.01, 45.30]	p = 0.842
		250m	OR = 1.27 [0.67, 2.40]	p = 0.467	OR = 1.25 [0.66, 2.36]	p = 0.500	OR = 1.24 [0.65, 2.35]	p = 0.508
		500m	OR = 0.96 [0.85, 1.08]	p = 0.476	OR = 0.97 [0.86, 1.08]	p = 0.555	OR = 0.99 [0.88, 1.11]	p = 0.834
	vinclozolin	50m	OR = 0 [0.00, 1.6e+101]	p = 0.325	OR = 0 [0.00, 1.8e+105]	p = 0.341	OR = 0 [0.00, 5.1e+116]	p = 0.401
		100m	OR = 0 [0.00, 1.8e+20]	p = 0.370	OR = 0 [0.00, 1.5e+21]	p = 0.393	OR = 0 [0.00, 4.9e+23]	p = 0.471
		250m	OR = 0 [0.00, 137.73]	p = 0.235	OR = 0 [0.00, 184.98]	p = 0.252	OR = 0 [0.00, 404.35]	p = 0.305
		500m	OR = 0.3 [0.05, 1.77]	p = 0.183	OR = 0.31 [0.05, 1.84]	p = 0.198	OR = 0.15 [0.02, 0.96]	p = 0.045

Outcome	Active ingredient (AI)	Buffer size	Basic model		Intermediate model		Full model	
			Risk estimate	p value	Risk estimate	p value	Risk estimate	p value
Low birth weight	asulam	50m	OR = 1.57 [0.27, 9.29]	p = 0.620	OR = 1.45 [0.24, 8.80]	p = 0.689	OR = 1.51 [0.25, 9.04]	p = 0.655
		100m	OR = 1.16 [0.81, 1.66]	p = 0.406	OR = 1.14 [0.79, 1.65]	p = 0.473	OR = 1.15 [0.80, 1.66]	p = 0.443
		250m	OR = 1.01 [0.96, 1.07]	p = 0.729	OR = 1.01 [0.95, 1.06]	p = 0.820	OR = 1.01 [0.95, 1.06]	p = 0.769
		500m	OR = 0.99 [0.99, 1.00]	p = 0.102	OR = 0.99 [0.99, 1.00]	p = 0.182	OR = 0.99 [0.99, 1.00]	p = 0.141
	carbetamide	50m	OR = 1.52 [0.04, 53.92]	p = 0.818	OR = 1.42 [0.04, 49.77]	p = 0.846	OR = 1.45 [0.04, 50.60]	p = 0.839
		100m	OR = 0.97 [0.50, 1.90]	p = 0.931	OR = 0.95 [0.49, 1.87]	p = 0.891	OR = 0.96 [0.49, 1.88]	p = 0.911
		250m	OR = 0.94 [0.85, 1.04]	p = 0.224	OR = 0.94 [0.85, 1.04]	p = 0.250	OR = 0.94 [0.85, 1.05]	p = 0.271
		500m	OR = 0.99 [0.97, 1.01]	p = 0.172	OR = 0.99 [0.97, 1.01]	p = 0.448	OR = 0.99 [0.97, 1.01]	p = 0.479
	cyproconazole	50m	OR = 3.82 [0.00, 5.0e+56]	p = 0.984	OR = 2200000 [0.00, 2.0e+62]	p = 0.824	OR = 510000 [0.00, 5.3e+61]	p = 0.842
		100m	OR = 1100 [0.00, 5.2e+13]	p = 0.578	OR = 11000 [0.00, 5.5e+14]	p = 0.458	OR = 7700 [0.00, 3.9e+14]	p = 0.477
		250m	OR = 0.07 [0.00, 3.24]	p = 0.175	OR = 0.08 [0.00, 3.90]	p = 0.205	OR = 0.08 [0.00, 3.82]	p = 0.201
		500m	OR = 0.84 [0.36, 1.95]	p = 0.681	OR = 0.82 [0.35, 1.92]	p = 0.648	OR = 0.87 [0.37, 2.04]	p = 0.747
	epoxiconazole	50m	OR = 2.17 [0.00, 3.2e+10]	p = 0.948	OR = 0.68 [0.00, 1.4e+10]	p = 0.975	OR = 0.43 [0.00, 9.3e+09]	p = 0.945
		100m	OR = 0.27 [0.00, 24.89]	p = 0.573	OR = 0.26 [0.00, 24.05]	p = 0.558	OR = 0.24 [0.00, 22.40]	p = 0.536
		250m	OR = 0.67 [0.34, 1.31]	p = 0.237	OR = 0.65 [0.33, 1.28]	p = 0.211	OR = 0.63 [0.32, 1.25]	p = 0.188
		500m	OR = 0.98 [0.88, 1.09]	p = 0.706	OR = 0.96 [0.86, 1.07]	p = 0.450	OR = 0.95 [0.85, 1.06]	p = 0.331
	fluroxypyr-meptyl	50m	OR = 0.03 [0.00, 1.3e+03]	p = 0.530	OR = 0.06 [0.00, 2.5e+03]	p = 0.608	OR = 0.07 [0.00, 2.8e+03]	p = 0.620
		100m	OR = 0.29 [0.04, 2.40]	p = 0.253	OR = 0.34 [0.04, 2.83]	p = 0.322	OR = 0.36 [0.04, 2.94]	p = 0.339
		250m	OR = 0.8 [0.58, 1.12]	p = 0.197	OR = 0.87 [0.62, 1.21]	p = 0.401	OR = 0.89 [0.63, 1.24]	p = 0.476
		500m	OR = 0.95 [0.90, 1.01]	p = 0.092	OR = 0.99 [0.93, 1.05]	p = 0.626	OR = 1.01 [0.95, 1.08]	p = 0.674
	glufosinate	50m	OR = 0 [0.00, 0.05]	p = 0.030	OR = 0 [0.00, 0.11]	p = 0.034	OR = 0 [0.00, 0.12]	p = 0.034
		100m	OR = 0 [0.00, 0.18]	p = 0.008	OR = 0 [0.00, 0.21]	p = 0.010	OR = 0 [0.00, 0.22]	p = 0.010
		250m	OR = 0.83 [0.45, 1.54]	p = 0.557	OR = 0.86 [0.46, 1.60]	p = 0.635	OR = 0.87 [0.47, 1.61]	p = 0.656
		500m	OR = 0.98 [0.88, 1.09]	p = 0.713	OR = 1 [0.90, 1.12]	p = 0.955	OR = 1 [0.90, 1.12]	p = 0.982
	glufosinate-ammonium	50m	OR = 51000 [12.67, 2.0e+08]	p = 0.010	OR = 40000 [8.68, 1.9e+08]	p = 0.014	OR = 41000 [8.54, 1.9e+08]	p = 0.014
		100m	OR = 18.89 [3.28, 108.82]	p = 0.001	OR = 18.79 [3.19, 110.83]	p = 0.001	OR = 18.83 [3.19, 111.33]	p = 0.001
		250m	OR = 1.16 [0.85, 1.58]	p = 0.362	OR = 1.16 [0.85, 1.58]	p = 0.358	OR = 1.16 [0.85, 1.59]	p = 0.343
		500m	OR = 1.04 [0.99, 1.10]	p = 0.131	OR = 1.05 [0.99, 1.11]	p = 0.084	OR = 1.05 [1.00, 1.11]	p = 0.053
	linuron	50m	OR = 0.03 [0.00, 2.51]	p = 0.122	OR = 0.04 [0.00, 3.27]	p = 0.154	OR = 0.04 [0.00, 3.02]	p = 0.144
		100m	OR = 0.57 [0.26, 1.29]	p = 0.177	OR = 0.62 [0.28, 1.39]	p = 0.243	OR = 0.6 [0.27, 1.35]	p = 0.220
		250m	OR = 0.95 [0.84, 1.06]	p = 0.356	OR = 0.96 [0.85, 1.07]	p = 0.449	OR = 0.95 [0.85, 1.07]	p = 0.404
		500m	OR = 1 [0.98, 1.02]	p = 0.966	OR = 1 [0.98, 1.02]	p = 0.695	OR = 1 [0.98, 1.02]	p = 0.892
	propiconazole	50m	OR = 7300 [0.11, 4.7e+08]	p = 0.116	OR = 12000 [0.14, 1.1e+09]	p = 0.104	OR = 13000 [0.15, 1.1e+09]	p = 0.103
		100m	OR = 2.79 [0.33, 23.48]	p = 0.345	OR = 3.24 [0.37, 28.17]	p = 0.287	OR = 3.26 [0.37, 28.38]	p = 0.284
		250m	OR = 0.89 [0.62, 1.28]	p = 0.531	OR = 0.9 [0.63, 1.30]	p = 0.582	OR = 0.9 [0.63, 1.30]	p = 0.585
		500m	OR = 1.03 [0.96, 1.10]	p = 0.376	OR = 1.03 [0.97, 1.10]	p = 0.330	OR = 1.03 [0.97, 1.10]	p = 0.363

S5 Results of regression analyses with increasing level of adjustment to potential confounders (cont.)

Outcome	Active ingredient (AI)	Buffer size	Basic model		Intermediate model		Full model	
			Risk estimate	p value	Risk estimate	p value	Risk estimate	p value
Low birth weight	thiacloprid	50m	OR = 0.18 [0.00, 1.2e+05]	p = 0.800	OR = 0.12 [0.00, 7.2e+04]	p = 0.751	OR = 0.12 [0.00, 7.5e+04]	p = 0.757
		100m	OR = 1.25 [0.11, 14.61]	p = 0.861	OR = 1.12 [0.10, 13.13]	p = 0.930	OR = 1.14 [0.10, 13.36]	p = 0.918
		250m	OR = 0.94 [0.65, 1.34]	p = 0.715	OR = 0.94 [0.66, 1.34]	p = 0.728	OR = 0.95 [0.66, 1.35]	p = 0.761
		500m	OR = 0.98 [0.93, 1.04]	p = 0.539	OR = 1 [0.94, 1.06]	p = 0.924	OR = 1 [0.94, 1.05]	p = 0.902
	triadimenol	50m	OR = 2.41 [0.00, 6.7e+17]	p = 0.966	OR = 0.05 [0.00, 4.9e+16]	p = 0.890	OR = 0.05 [0.00, 4.8e+16]	p = 0.889
		100m	OR = 377.28 [0.40, 3.6e+05]	p = 0.090	OR = 189.94 [0.17, 2.1e+05]	p = 0.142	OR = 190.83 [0.17, 2.1e+05]	p = 0.141
		250m	OR = 2.54 [0.93, 6.99]	p = 0.070	OR = 2.4 [0.86, 6.66]	p = 0.093	OR = 2.45 [0.88, 6.79]	p = 0.085
		500m	OR = 1.02 [0.84, 1.23]	p = 0.863	OR = 1.05 [0.87, 1.26]	p = 0.642	OR = 1.08 [0.89, 1.30]	p = 0.441
	vinclozolin	50m	OR = 1.8e+185 [0.00, Inf]	p = 0.170	OR = 2e+179 [0.00, Inf]	p = 0.200	OR = 3e+185 [0.00, Inf]	p = 0.181
		100m	OR = 1.8e+28 [0.00, 6.1e+90]	p = 0.376	OR = 9.7e+25 [0.00, 4.9e+90]	p = 0.431	OR = 5.8e+27 [0.00, 1.1e+92]	p = 0.397
		250m	OR = 0 [0.00, 1.0e+08]	p = 0.597	OR = 0 [0.00, 2.0e+07]	p = 0.502	OR = 0 [0.00, 4.1e+07]	p = 0.538
		500m	OR = 0.02 [0.00, 0.52]	p = 0.019	OR = 0.03 [0.00, 0.75]	p = 0.033	OR = 0.02 [0.00, 0.64]	p = 0.026
Large for gestational age	asulam	50m	OR = 0.59 [0.32, 1.08]	p = 0.089	OR = 0.58 [0.31, 1.07]	p = 0.083	OR = 0.57 [0.31, 1.06]	p = 0.077
		100m	OR = 0.91 [0.80, 1.03]	p = 0.132	OR = 0.91 [0.80, 1.03]	p = 0.124	OR = 0.9 [0.80, 1.02]	p = 0.113
		250m	OR = 1 [0.98, 1.02]	p = 0.659	OR = 1 [0.98, 1.02]	p = 0.643	OR = 0.99 [0.98, 1.01]	p = 0.612
		500m	OR = 1 [1.00, 1.01]	p = 0.023	OR = 1 [1.00, 1.01]	p = 0.036	OR = 1 [1.00, 1.01]	p = 0.039
	carbetamide	50m	OR = 1.04 [0.23, 4.64]	p = 0.959	OR = 1.03 [0.23, 4.57]	p = 0.973	OR = 0.97 [0.22, 4.33]	p = 0.967
		100m	OR = 0.96 [0.71, 1.28]	p = 0.759	OR = 0.95 [0.71, 1.27]	p = 0.732	OR = 0.94 [0.70, 1.26]	p = 0.670
		250m	OR = 0.98 [0.94, 1.03]	p = 0.421	OR = 0.98 [0.94, 1.03]	p = 0.376	OR = 0.98 [0.93, 1.02]	p = 0.340
		500m	OR = 0.99 [0.98, 1.00]	p = 0.134	OR = 0.99 [0.98, 1.00]	p = 0.078	OR = 0.99 [0.98, 1.00]	p = 0.087
	cyproconazole	50m	OR = 8.6e+21 [0.00, 3.8e+46]	p = 0.081	OR = 4.1e+21 [0.00, 1.9e+46]	p = 0.086	OR = 8e+20 [0.00, 3.8e+45]	p = 0.097
		100m	OR = 345.08 [0.01, 2.1e+07]	p = 0.299	OR = 268.2 [0.00, 1.7e+07]	p = 0.321	OR = 173.54 [0.00, 1.1e+07]	p = 0.360
		250m	OR = 8.1 [1.59, 41.22]	p = 0.012	OR = 7.59 [1.49, 38.61]	p = 0.015	OR = 7.03 [1.38, 35.85]	p = 0.019
		500m	OR = 0.69 [0.46, 1.04]	p = 0.074	OR = 0.7 [0.47, 1.06]	p = 0.090	OR = 0.69 [0.46, 1.05]	p = 0.081
	epoxiconazole	50m	OR = 0.02 [0.00, 408.96]	p = 0.427	OR = 0.01 [0.00, 352.94]	p = 0.410	OR = 0.04 [0.00, 975.37]	p = 0.534
		100m	OR = 0.57 [0.08, 3.91]	p = 0.569	OR = 0.55 [0.08, 3.79]	p = 0.548	OR = 0.72 [0.11, 4.94]	p = 0.742
		250m	OR = 1 [0.75, 1.34]	p = 0.987	OR = 1 [0.75, 1.33]	p = 0.976	OR = 1.04 [0.78, 1.39]	p = 0.800
		500m	OR = 1.01 [0.96, 1.06]	p = 0.773	OR = 1.01 [0.96, 1.06]	p = 0.780	OR = 1 [0.95, 1.05]	p = 0.881
	fluroxypyr-meptyl	50m	OR = 3.99 [0.04, 449.92]	p = 0.566	OR = 3.49 [0.03, 397.69]	p = 0.605	OR = 1.76 [0.02, 200.58]	p = 0.815
		100m	OR = 1.7 [0.67, 4.32]	p = 0.269	OR = 1.65 [0.65, 4.20]	p = 0.297	OR = 1.37 [0.54, 3.49]	p = 0.513
		250m	OR = 1.15 [0.99, 1.34]	p = 0.060	OR = 1.14 [0.98, 1.33]	p = 0.078	OR = 1.1 [0.95, 1.28]	p = 0.212
		500m	OR = 0.99 [0.96, 1.02]	p = 0.441	OR = 0.99 [0.96, 1.01]	p = 0.314	OR = 0.98 [0.95, 1.01]	p = 0.116

Outcome	Active ingredient (AI)	Buffer size	Basic model		Intermediate model		Full model	
			Risk estimate	p value	Risk estimate	p value	Risk estimate	p value
Large for gestational age	glufosinate	50m	OR = 0 [0.00, 3.80]	p = 0.101	OR = 0 [0.00, 4.96]	p = 0.115	OR = 0 [0.00, 9.75]	p = 0.156
		100m	OR = 0.33 [0.06, 1.85]	p = 0.209	OR = 0.36 [0.06, 1.98]	p = 0.239	OR = 0.41 [0.07, 2.30]	p = 0.312
		250m	OR = 0.96 [0.73, 1.26]	p = 0.755	OR = 0.96 [0.73, 1.26]	p = 0.764	OR = 0.96 [0.73, 1.27]	p = 0.777
		500m	OR = 1 [0.95, 1.05]	p = 0.907	OR = 1 [0.95, 1.05]	p = 0.942	OR = 0.96 [0.91, 1.01]	p = 0.085
	glufosinate-ammonium	50m	OR = 8.85 [0.21, 367.84]	p = 0.252	OR = 9.58 [0.23, 399.16]	p = 0.235	OR = 6.91 [0.17, 287.36]	p = 0.310
		100m	OR = 1.57 [0.68, 3.64]	p = 0.287	OR = 1.6 [0.69, 3.70]	p = 0.269	OR = 1.48 [0.64, 3.42]	p = 0.354
		250m	OR = 0.96 [0.83, 1.11]	p = 0.588	OR = 0.96 [0.83, 1.12]	p = 0.614	OR = 0.95 [0.82, 1.10]	p = 0.514
		500m	OR = 0.96 [0.94, 0.99]	p = 0.006	OR = 0.96 [0.94, 0.99]	p = 0.005	OR = 0.97 [0.95, 1.00]	p = 0.032
	linuron	50m	OR = 4.64 [0.88, 24.41]	p = 0.070	OR = 4.09 [0.78, 21.51]	p = 0.097	OR = 4.19 [0.80, 22.13]	p = 0.091
		100m	OR = 1.46 [1.05, 2.02]	p = 0.023	OR = 1.42 [1.02, 1.97]	p = 0.036	OR = 1.43 [1.03, 1.98]	p = 0.033
		250m	OR = 1.08 [1.02, 1.13]	p = 0.005	OR = 1.07 [1.02, 1.13]	p = 0.008	OR = 1.07 [1.02, 1.13]	p = 0.008
		500m	OR = 1.01 [1.00, 1.02]	p = 0.004	OR = 1.02 [1.01, 1.03]	p < 0.001	OR = 1.01 [1.00, 1.02]	p = 0.005
	propiconazole	50m	OR = 12.01 [0.14, 1.1e+03]	p = 0.276	OR = 14.67 [0.17, 1.3e+03]	p = 0.240	OR = 11.76 [0.13, 1.0e+03]	p = 0.281
		100m	OR = 1.3 [0.48, 3.52]	p = 0.611	OR = 1.37 [0.50, 3.72]	p = 0.536	OR = 1.31 [0.48, 3.57]	p = 0.592
		250m	OR = 0.9 [0.75, 1.07]	p = 0.228	OR = 0.9 [0.76, 1.08]	p = 0.256	OR = 0.9 [0.76, 1.08]	p = 0.251
		500m	OR = 0.94 [0.91, 0.97]	p < 0.001	OR = 0.94 [0.91, 0.97]	p < 0.001	OR = 0.95 [0.92, 0.98]	p = 0.004
	thiacloprid	50m	OR = 0.18 [0.00, 31.54]	p = 0.517	OR = 0.18 [0.00, 30.38]	p = 0.508	OR = 0.13 [0.00, 22.34]	p = 0.434
		100m	OR = 0.81 [0.29, 2.25]	p = 0.693	OR = 0.81 [0.29, 2.24]	p = 0.682	OR = 0.75 [0.27, 2.08]	p = 0.580
		250m	OR = 0.96 [0.82, 1.13]	p = 0.636	OR = 0.96 [0.82, 1.12]	p = 0.606	OR = 0.95 [0.81, 1.11]	p = 0.530
		500m	OR = 0.99 [0.96, 1.01]	p = 0.295	OR = 0.98 [0.96, 1.01]	p = 0.204	OR = 0.99 [0.97, 1.02]	p = 0.521
	triadimenol	50m	OR = 1.06 [0.00, 1.8e+06]	p = 0.993	OR = 0.72 [0.00, 1.3e+06]	p = 0.965	OR = 0.61 [0.00, 1.1e+06]	p = 0.946
		100m	OR = 1.84 [0.09, 37.74]	p = 0.693	OR = 1.62 [0.08, 33.42]	p = 0.753	OR = 1.47 [0.07, 30.47]	p = 0.803
		250m	OR = 1.19 [0.74, 1.93]	p = 0.471	OR = 1.18 [0.73, 1.92]	p = 0.491	OR = 1.14 [0.70, 1.85]	p = 0.591
		500m	OR = 1.04 [0.95, 1.13]	p = 0.372	OR = 1.03 [0.95, 1.13]	p = 0.454	OR = 1.01 [0.93, 1.10]	p = 0.801
	vinclozolin	50m	OR = 0 [0.00, 7.4e+31]	p = 0.183	OR = 0 [0.00, 1.5e+30]	p = 0.172	OR = 0 [0.00, 6.6e+34]	p = 0.203
		100m	OR = 0 [0.00, 1.5e+06]	p = 0.168	OR = 0 [0.00, 5.0e+05]	p = 0.154	OR = 0 [0.00, 4.6e+06]	p = 0.183
		250m	OR = 0.08 [0.00, 131.90]	p = 0.498	OR = 0.07 [0.00, 118.37]	p = 0.480	OR = 0.09 [0.00, 158.84]	p = 0.532
		500m	OR = 4.84 [1.73, 13.52]	p = 0.003	OR = 4.54 [1.62, 12.69]	p = 0.004	OR = 3.49 [1.23, 9.90]	p = 0.019
Small for gestational age	asulam	50m	OR = 0.9 [0.40, 2.02]	p = 0.791	OR = 0.9 [0.40, 2.04]	p = 0.807	OR = 0.93 [0.41, 2.11]	p = 0.870
		100m	OR = 1 [0.85, 1.18]	p = 0.957	OR = 1 [0.85, 1.18]	p = 0.961	OR = 1.01 [0.86, 1.19]	p = 0.879
		250m	OR = 1 [0.97, 1.02]	p = 0.879	OR = 1 [0.97, 1.02]	p = 0.875	OR = 1 [0.97, 1.03]	p = 0.958
		500m	OR = 0.99 [0.99, 1.00]	p < 0.001	OR = 0.99 [0.99, 1.00]	p < 0.001	OR = 0.99 [0.99, 1.00]	p < 0.001
	carbetamide	50m	OR = 1.65 [0.32, 8.47]	p = 0.547	OR = 1.67 [0.32, 8.56]	p = 0.542	OR = 1.81 [0.35, 9.27]	p = 0.478
		100m	OR = 1.16 [0.85, 1.59]	p = 0.340	OR = 1.17 [0.86, 1.61]	p = 0.315	OR = 1.2 [0.88, 1.64]	p = 0.258
		250m	OR = 1.01 [0.96, 1.06]	p = 0.662	OR = 1.02 [0.97, 1.07]	p = 0.521	OR = 1.02 [0.97, 1.07]	p = 0.442
		500m	OR = 1 [0.99, 1.01]	p = 0.789	OR = 1 [0.99, 1.01]	p = 0.522	OR = 1 [0.99, 1.01]	p = 0.498
	cyproconazole	50m	OR = 0 [0.00, 5.4e+18]	p = 0.486	OR = 0 [0.00, 1.1e+20]	p = 0.541	OR = 0 [0.00, 5.7e+19]	p = 0.528
		100m	OR = 0 [0.00, 0.59]	p = 0.041	OR = 0 [0.00, 1.16]	p = 0.053	OR = 0 [0.00, 1.06]	p = 0.051
		250m	OR = 0.07 [0.01, 0.45]	p = 0.005	OR = 0.08 [0.01, 0.50]	p = 0.007	OR = 0.08 [0.01, 0.51]	p = 0.007
		500m	OR = 0.75 [0.49, 1.14]	p = 0.182	OR = 0.74 [0.48, 1.12]	p = 0.155	OR = 0.82 [0.54, 1.24]	p = 0.345

→

S5 Results of regression analyses with increasing level of adjustment to potential confounders (cont.)

Outcome	Active ingredient (AI)	Buffer size	Basic model		Intermediate model		Full model	
			Risk estimate	p value	Risk estimate	p value	Risk estimate	p value
Small for gestational age	epoxiconazole	50m	OR = 0.02 [0.00, 823.29]	p = 0.475	OR = 0.03 [0.00, 1.3e+03]	p = 0.527	OR = 0.02 [0.00, 716.07]	p = 0.454
		100m	OR = 0.78 [0.10, 6.01]	p = 0.812	OR = 0.87 [0.11, 6.75]	p = 0.892	OR = 0.74 [0.09, 5.80]	p = 0.774
		250m	OR = 1.05 [0.77, 1.43]	p = 0.746	OR = 1.05 [0.78, 1.43]	p = 0.736	OR = 1.02 [0.75, 1.38]	p = 0.917
		500m	OR = 1.08 [1.03, 1.14]	p = 0.001	OR = 1.08 [1.02, 1.13]	p = 0.004	OR = 1.06 [1.01, 1.11]	p = 0.029
	fluroxypyr-meptyl	50m	OR = 0.95 [0.01, 148.84]	p = 0.985	OR = 1.08 [0.01, 170.79]	p = 0.976	OR = 1.6 [0.01, 254.46]	p = 0.856
		100m	OR = 0.48 [0.17, 1.32]	p = 0.154	OR = 0.5 [0.18, 1.39]	p = 0.185	OR = 0.56 [0.20, 1.56]	p = 0.267
		250m	OR = 0.82 [0.70, 0.96]	p = 0.015	OR = 0.84 [0.72, 0.99]	p = 0.041	OR = 0.88 [0.75, 1.03]	p = 0.114
		500m	OR = 0.95 [0.92, 0.98]	p < 0.001	OR = 0.96 [0.93, 0.99]	p = 0.006	OR = 1 [0.97, 1.03]	p = 0.935
	glufosinate	50m	OR = 2.57 [0.00, 1.1e+04]	p = 0.826	OR = 1.82 [0.00, 8.0e+03]	p = 0.889	OR = 2.19 [0.00, 9.6e+03]	p = 0.855
		100m	OR = 1.1 [0.18, 6.63]	p = 0.920	OR = 1.03 [0.17, 6.26]	p = 0.971	OR = 1.1 [0.18, 6.64]	p = 0.921
		250m	OR = 1.12 [0.84, 1.48]	p = 0.445	OR = 1.11 [0.84, 1.47]	p = 0.469	OR = 1.12 [0.84, 1.48]	p = 0.437
		500m	OR = 1.01 [0.96, 1.06]	p = 0.646	OR = 1.02 [0.97, 1.07]	p = 0.552	OR = 1.02 [0.96, 1.07]	p = 0.566
	glufosinate-ammonium	50m	OR = 0.03 [0.00, 3.25]	p = 0.144	OR = 0.03 [0.00, 2.76]	p = 0.126	OR = 0.03 [0.00, 2.84]	p = 0.128
		100m	OR = 0.5 [0.19, 1.32]	p = 0.162	OR = 0.48 [0.18, 1.29]	p = 0.145	OR = 0.48 [0.18, 1.29]	p = 0.144
		250m	OR = 1.05 [0.90, 1.23]	p = 0.502	OR = 1.06 [0.91, 1.23]	p = 0.488	OR = 1.06 [0.91, 1.24]	p = 0.431
		500m	OR = 1 [0.97, 1.03]	p = 0.921	OR = 1.01 [0.98, 1.03]	p = 0.634	OR = 1.01 [0.99, 1.04]	p = 0.285
	linuron	50m	OR = 0.85 [0.12, 5.75]	p = 0.864	OR = 1.08 [0.16, 7.35]	p = 0.937	OR = 0.95 [0.14, 6.46]	p = 0.957
		100m	OR = 0.91 [0.63, 1.33]	p = 0.637	OR = 0.96 [0.66, 1.40]	p = 0.850	OR = 0.94 [0.64, 1.36]	p = 0.731
		250m	OR = 0.95 [0.89, 1.00]	p = 0.054	OR = 0.95 [0.90, 1.01]	p = 0.086	OR = 0.95 [0.90, 1.00]	p = 0.069
		500m	OR = 1 [0.99, 1.01]	p = 0.444	OR = 0.99 [0.98, 1.00]	p = 0.070	OR = 0.99 [0.98, 1.01]	p = 0.320
	propiconazole	50m	OR = 0.22 [0.00, 42.79]	p = 0.570	OR = 0.14 [0.00, 29.05]	p = 0.474	OR = 0.18 [0.00, 35.30]	p = 0.520
		100m	OR = 0.64 [0.21, 1.95]	p = 0.433	OR = 0.59 [0.19, 1.81]	p = 0.359	OR = 0.61 [0.20, 1.87]	p = 0.388
		250m	OR = 1 [0.84, 1.20]	p = 0.958	OR = 1 [0.84, 1.19]	p = 0.984	OR = 1 [0.84, 1.19]	p = 0.992
		500m	OR = 1 [0.97, 1.03]	p = 0.956	OR = 1 [0.97, 1.04]	p = 0.837	OR = 1 [0.97, 1.03]	p = 0.958
	thiacloprid	50m	OR = 0.59 [0.00, 226.58]	p = 0.861	OR = 0.56 [0.00, 217.71]	p = 0.849	OR = 0.66 [0.00, 254.20]	p = 0.893
		100m	OR = 1.06 [0.33, 3.39]	p = 0.922	OR = 1.05 [0.33, 3.35]	p = 0.940	OR = 1.1 [0.34, 3.50]	p = 0.876
		250m	OR = 0.94 [0.79, 1.13]	p = 0.507	OR = 0.95 [0.79, 1.14]	p = 0.566	OR = 0.96 [0.80, 1.14]	p = 0.629
		500m	OR = 0.96 [0.93, 0.98]	p = 0.003	OR = 0.97 [0.94, 1.00]	p = 0.027	OR = 0.97 [0.94, 0.99]	p = 0.019
	triadimenol	50m	OR = 0 [0.00, 3.1e+03]	p = 0.363	OR = 0 [0.00, 5.8e+03]	p = 0.405	OR = 0 [0.00, 4.2e+03]	p = 0.384
		100m	OR = 0.57 [0.02, 13.13]	p = 0.726	OR = 0.66 [0.03, 15.27]	p = 0.795	OR = 0.65 [0.03, 14.89]	p = 0.786
		250m	OR = 1.29 [0.79, 2.10]	p = 0.302	OR = 1.32 [0.81, 2.15]	p = 0.265	OR = 1.37 [0.84, 2.23]	p = 0.203
		500m	OR = 1.05 [0.96, 1.14]	p = 0.305	OR = 1.06 [0.97, 1.16]	p = 0.191	OR = 1.11 [1.02, 1.21]	p = 0.018
	vinclozolin	50m	OR = 0 [0.00, 2.2e+121]	p = 0.875	OR = 0 [0.00, 4.2e+122]	p = 0.885	OR = 0 [0.00, 2.0e+127]	p = 0.944
		100m	OR = 38.51 [0.00, 1.7e+28]	p = 0.907	OR = 81.57 [0.00, 5.9e+28]	p = 0.889	OR = 1900 [0.00, 1.3e+30]	p = 0.811
		250m	OR = 5.27 [0.00, 7.9e+04]	p = 0.735	OR = 5.79 [0.00, 9.1e+04]	p = 0.722	OR = 9.34 [0.00, 1.5e+05]	p = 0.651
		500m	OR = 0.2 [0.04, 0.90]	p = 0.036	OR = 0.25 [0.05, 1.11]	p = 0.069	OR = 0.21 [0.05, 0.96]	p = 0.044

CHAPTER 6

GENERAL DISCUSSION

The study of exposure to agricultural pollutants, specifically pesticides and livestock emissions, and health effects among residents in rural areas is a complex. The work presented in this thesis makes its contribution to the topic by using large administrative data sets and modelled exposure to identify potential associations of these exposures to health outcomes. Using large data resources to study this topic is a powerful approach, since large studies help identify small risks that may have a significant impact in rural population's health, given the large number of people potentially exposed to agricultural pollutants. In this chapter, lessons learned from the previous chapters are brought together to provide a picture of what insights were gained and how they fit in past and future research on this topic. The chapter begins by revisiting the main findings and proceeds with a discussion of the strengths and limitations of the approaches taken. These are further put into context by comparing this work with that of other studies. Finally, it provides a perspective for future work on the topic.

SUMMARY OF THE MAIN RESULTS

In general, and contrary to expectation, people living near crops where pesticides were applied presented better physical and mental health (**chapter 4**) and overall lower mortality risks (**chapter 3**) when compared to people living further way but still in rural areas. In contrast to this overall pattern, a signal for an association between living near maize crops and respiratory mortality arose, even when stringent interpretation criteria were applied (**chapter 3**). The study on birth outcomes identified specific active ingredients of concern, namely linuron, fluroxypyr-meptyl, glufosinate-ammonium, vinclozolin and picoxystrobin, raising important questions regarding mechanisms for toxicity and dysregulation that could potentially lead to disease (**chapter 5**). The association found between residential proximity to pig farms and increased respiratory mortality found in the livestock study, albeit referring to a time period of about 20 years ago, indicate a potential concern regarding pig farm emissions that should be investigated, since it is likely that exposures are still qualitatively similar nowadays (**chapter 2**).
The use of administrative databases supplied the studies with a large number of subjects for analyses, providing enough power to detect even small risks but rendering interpretation challenging, especially in distinguishing false positive results and judging what pertains a relevant health effect. It also required exposure assessment methods based on modelling given the impossibility to measure and analyze exposures of such large study populations. Concomitantly, selection of health outcomes was limited to those available in registries at the time the studies were conducted. The use of these data sources shaped the analyses described in this thesis. Strengths, limitations and

methodological aspects of using these administrative datasets for the study of the effect of agricultural pollutants on rural populations health are discussed and put into context with other approaches below.

THE STUDY OF THE EFFECT OF AGRICULTURAL POLLUTANTS ON HEALTH – CONSIDERATIONS ON THE APPROACHES TAKEN IN THIS THESIS

There is currently no gold standard approach to study the effect of agricultural pollutants on health, largely because both "agricultural pollutants" and "health" encompass a great number of aspects. Different approaches regarding data sources and exposure and outcome assessment methods present their own strengths and limitations as well as feasibility to address specific goals (exploratory research to identify potential associations, corroboration of previously reported associations, quantification of effect sizes). As mentioned, the studies comprising this thesis had a common premise: the use of administrative data sets for outcome and exposure assessment. In particular, farmers' surveys and registries and geodatabases were used to compute proxies of exposure to pesticides and to livestock emissions based on residential proximity to crops and farms, respectively, while outcome and covariate information was obtained from national health surveys and registries.

THE USE OF ADMINISTRATIVE DATA

Administrative data can supply large amounts of study subjects, resulting in large data sets for analysis. In **chapters 2** and **3**, virtually all rural residents fulfilling the inclusion criteria (30 years old or older, complete information on residential history) were included, resulting in data sets of 3–4 million subjects. Similarly, in **chapter 5**, virtually all singleton births from mothers living in rural areas that had occurred over 5 years (nearly 340,000) were included in the study. In **chapter 4**, although the study population was comprised of a sample of the Dutch population, the study still relied on over 215.000 subjects. The inclusion of a large number of subjects provides the statistical analysis with large power to detect subtle effects and allows investigation of rare outcomes and rare exposures. Small effect sizes are common in observational epidemiological studies but because a large number of people are potentially exposed there can be important public health and policy consequences. Likewise, outcomes that do not occur often may still be of interest if they have a severe impact on society and/or affect individuals early in life, such as some rarer types of cancer or congenital malformations.

Administrative information is virtually complete across exposures, outcomes and some relevant covariates, available at individual level. Indeed, it was possible to compute exposure proxies for all residences in the Netherlands based on the Basisregistratie Adressen en Gebouwen (BAG), a public geodatabase containing all addresses and buildings in the Netherlands. The Basisregistratie Gewaspercelen (BRP) and the Landelijk Grondgebruik Nederland (LGN) geodatabases provided maps with information on the location and types of crops while farmer's surveys provided information on the type and amounts of pesticides used on those crops. Agricultural registries provided information on the location, type and size of farms. Death and birth registries provided the information on nearly all deaths and births in the Netherlands while the Health Survey 2012 provided a study population based on a representative sample of the Dutch population and self-reported depression and perceived health. Because the registries and the health survey are available as microdata within Statistics Netherlands (CBS) secure environment, it was possible to link them to other sources of administrative data obtaining therefore information on relevant covariates such as sex, age, complete residential history (an essential input for linking individual health data to modelled exposure), household income, migration background, and marital status as well as information on aggregated level such as neighborhood's urbanization degree, and social economic position. The survey provided, in addition, information on lifestyle factors. Furthermore, completeness of data also pertains to completeness of subjects in the databases, that is, the entire Dutch population. Since virtually all people complying with the specific inclusion criteria for each chapter were included in the studies, it was possible to circumvent the limitations of some types of bias such as selection or participation bias.

In the Netherlands, and many other countries, administrative data is routinely collected for a myriad of purposes, consisting of a cost-efficient way of obtaining and storing information. In contrast, setting up and maintaining large cohorts is a laborious and expensive endeavor, especially when follow-up extends for long periods; recruiting participants in case-control studies is also often challenging and slow. This constant availability of data allows for the prompt start of retrospective cohort, case-control and cross-sectional studies at any time, provided that it is possible to assess exposure in the past. It is important to note, however, that consistent and periodic collection of registry data or availability of such data for research is not a reality for all countries.

STRENGTHS

Exposure assessment

Because of the large number of subjects included in the studies covered in this thesis, measurement of exposure concentrations was not feasible and modelling approaches were used to estimate residential exposure to agricultural pollutants. The models relied on registries as input data allowing for an objective exposure assessment at an individual level. An individual level objective assessment for agricultural pollutants is rarely observed in studies where measurements are not a possibility, which usually either rely on questionnaires or aggregated level information for exposure assessment, such as estimates on the amount of pesticides sold or the total number of farms in a region. Another advantage of using models is that it allows for the estimation of exposure levels for very large geographical areas. Indeed, in all chapters exposure was assessed for the whole country, whereas most studies on the topic are often focused on a regional approach. Nationwide exposure assessment not only provides large power in subsequent statistical analysis, but also helps achieve exposure contrasts that may be difficult to obtain when studying specific regions. Table 1 displays succinctly the main approaches to exposure assessment used in studies on the topic, providing a general overview of how these approaches compare in terms of key strengths and limitations.

Outcome assessment

The thesis explored several health endpoints, including birth outcomes, specific causes of death, and depression and perceived health. Birth outcomes and causes of death (**chapters 2, 3** and **5**) were available from registries, which provide complete information on individual level that was objectively determined by a health professional. By including all birth and death from people meeting the inclusion criteria of the specific studies, certain sources of biases were avoided, namely participation selection and exclusion biases. **Chapter 4** focused on (non-clinical) depression and perceived health. While self-reported, and thus conditional on some degree of subjectivity, these outcomes can only be assessed via questionnaire. Information on these outcomes was virtually complete (1–3% missing values).

Information on other relevant covariates

In all chapters, analyses considered several relevant factors that may be relevant for the associations under investigation, in increasingly adjusted models. Overall, information

on age, sex, migration background, household income and marital status, were available at the individual level and objectively obtained from registries. Another advantage of using registries is that this sort of information is nearly complete and systematically updated, with most data sets being produced on a yearly basis. Information, including start and end dates of all sorts of occurrences (such as births, deaths, changes in residence, changes in marital status, etc), is thus continuously assessed, processed and added to the data bases. However, information on lifestyle, such as drug and alcohol use, smoking habits or physical activity is, of course, not included in registries and therefore not included in the studies, except that of **chapter 4.** In this chapter, information on lifestyle factors was available since data in this study came primarily from the 2012 Health Survey. Of note, is that the national Health Surveys, although conducted on a regular basis (about every 4 years), are each time based on a different sample from the population. Consequently, these surveys often pertain to a one-time lifestyle factors assessment that is not updated regularly.

Statistical approach, biases and interpretation

In all chapters, an exploratory approach was used to identify potential associations between residential proximity to crops, amount of pesticides used near the residence or residential proximity to different types of farms, and several health endpoints. This consequently resulted in the estimation of many associations raising concerns about multiple testing and erroneous inferences about the results. Together with the use of big data, the work covered in this thesis is prone to detection of statistically significant findings and raises the chance of obtaining false positive results. To mitigate this, a rather strict interpretation of results was used: results were always interpreted in a holistic manner and, when applicable, the false discovery rate (FDR) method was used. An isolated result was never considered on its own but rather interpreted in context with other results. For example, there was careful verification of whether effects were consistent among different buffer sizes and/or among different exposure metrics. Then again, one can argue that application of so many criteria for interpretation could have hampered the detection of a true finding or may have resulted in overlooking of a signal (false negative). In exploratory research, such as the one featured in this thesis, balancing between false negative and false positive results is key. It is important that false negatives are avoided so as not to miss important signals that can become leads in confirmatory research (which will later "weed [false positive results] out")[1] However, a too lenient approach can come at the cost of obtaining a relatively high rate of false positives and generating spurious leads that would potentially waste research resources.

To address potential sources of bias, such as effect modification and confounding by unmeasured factors, many sensitivity analyses and stratified analyses were conducted across all chapters. Negative control analyses were carried out in chapter 2 as a further step to identify residual sources of bias, using outcomes that were related to poor lifestyle behaviors but not to exposure to livestock pollution. In chapter 3 analyses were re-run using another (smaller) cohort, the AMIGO cohort, where information on lifestyle factors was available. Interestingly, all of these analyses showed no major changes in results or at least not enough that would induce a change in the general interpretation of the results as a whole. Overall, throughout all chapters, rather stringent interpretation criteria were applied and there was thoughtful examination and integrated interpretation of the main and additional analyses. Such an approach borders the concept of triangulation and lends credit to the robustness and trustworthiness to the identified signals. As some authors argue, such pluralistic approaches may be more adequate to assess the impact of specific sources of bias in observational studies than using a deterministic risk of bias assessment method[2] .

LIMITATIONS

Exposure assessment

Despite the advantages of modelled exposure assessment used in this thesis, this approach has its own set of limitations. Logically, taking personal or environmental samples to measure exposure to agricultural pollutants delivers the highest level of detail for (individual) exposure assessment, providing information on specific compounds and their concentrations in the environment (residence) and/or in subjects. Nevertheless, this is usually only feasible in small to medium-scale studies and for a sample of the population of interest, with the added disadvantage that data collection can be affected by participation bias in some study designs (in terms of both exposure and outcome perceptions or concerns). In the studies comprising this thesis, measuring and analyzing samples for exposure assessment was not a possibility, not only because of the large amounts of study subjects but also because exposure occurred in past and no measurements were performed then. Indeed, a big limitation of this approach is that exposure measurements do not reflect past exposures when they are not performed during a study's exposure period.
Estimation of residential exposure to agricultural pollutants was thus obtained from a modelling approach, based on a rather crude computation and with limitations that complicate the interpretation of the results. First, the modelled exposure metrics re-

flected a mixture of components, that is, they reflected the total combination of pesticides used in a crop (**chapters 3** and **4**) and the several pollutants generated in livestock farming (various gases and particles including greenhouse gases, ammonia, fine particles and endotoxins, **chapter 2**). Although some specificity regarding the mixture was achieved by using exposure variables that reflected the area of a specific crop or the number of specific livestock animals, it was not possible to disentangle the individual effects of the compounds in the exposure mixture precluding pinpointing the actual culprits for the findings. The exception was the approach taken in **chapter 5** on birth outcomes, where it was possible to estimate the amount of specific active ingredients used around the residence. In this chapter, annual crop data with high resolution and detailed information on crop types was available (unlike **chapters 2** and **3**), the exposure time window was shorter and clearly defined and the design was longitudinal (unlike **chapter 4**). These conditions allowed for the computation of estimates of the amount of specific active ingredients used around residences and provided insight into possible cause-effect relationships between those active ingredients and birth outcomes. Throughout all chapters, the semi-quantitative proxies used for exposure unavoidably resulted in (difficult to quantify) degree of exposure misclassification. This misclassification is due to several aspects: 1) the inability to account for determinants that influence personal exposure, such as people's time-activity patterns and other sources of exposure to agricultural pollutants (occupation, food contamination), 2) meteorological conditions affecting the dispersion of agricultural pollutants and adoption of exposure mitigation measures by farmers, which can lead to a discussion on whether the use of symmetrical buffers is appropriate (even though it would remain difficult to determine what the optimal buffer shape would be), and 3) difficult identification of the pollutants' source in a more precise manner, especially pertaining to the exact location of livestock animals that are usually raised in grazing systems (such as dairy cattle and sheep). Nevertheless, it is important to mention that exposure assessment based on measurements is also not free of exposure misclassification due to other determinates, especially those related to the temporal variability of the concentrations of the components in the environment and in the body and the ability to take samples at the appropriate timing (Table 1). Still, the modelled exposure proxies used in the studies of this thesis are likely suited for long term exposure assessment. Indeed, area of crop and animal-specific livestock-related characteristics (such as number of farms or number of species-specific animals) within a buffer were found to be among the strongest predictors in exposure models developed for estimating residential exposure to pesticides and livestock-related air pollution, respectively[3,4] Studies conducted in the US have also shown high correlations between residential proximity to crops and pesticide concentrations at home (albeit a translation to the European setting may not

be entirely adequate since different agricultural practices, such as aerial spraying of pesticides, can affect the dispersion of agricultural pollutants)[5]
A major challenge, not only in this thesis' work but in general, is the linkage of exposure assessment data to the relevant induction time period for the outcome of interest, especially when it is unclear which time period of exposure would be of relevance (pregnancy, childhood, adolescence, adulthood) or when the induction time for the disease is unclear or long (a common feature of chronic diseases). Except for **chapter 5** on birth outcomes, where the exposure period considered was pregnancy, a rather loose approach was taken in all chapters, considering exposure periods only during adulthood. Nevertheless, one should not ignore that other time windows could be highly relevant, such as the case of early life exposures that can imply an increased duration of exposure or exposure during crucial developmental periods. A relevant reflection regarding the studies included in this thesis, however, is that looking into several time periods in addition to already large number of tests performed in each study would lead to a compounding of the multiple testing problem.
The study populations from the studies included in this thesis were based on the general population, which naturally includes farmers and other occupations where exposure to pesticides is higher. Because exposure was assessed based on residence, it was not possible to determine the occupational exposures. Nevertheless, since registries have information on working sectors, it was possible to perform sensitivity analyses excluding people working in the agricultural setting and assess if the effects were driven by occupational exposures. In all studies, this sensitivity analysis showed that the effects found remained, after excluding agricultural workers from analyses, albeit in some cases with a slightly lower magnitude. This reinforces that the findings were driven by residential exposures.

Outcome assessment

The quality of outcome definition across the chapters was variable. In **chapter 5**, analyses were based on quite detailed and complete data where birth outcomes from all births occurring after 22 weeks of gestation were registered by health professionals (midwives and obstetricians). In **chapters 2** and **3**, outcome was defined as the primary cause of death, which often results in missing cases for diseases that are usually not (registered as) direct causes of death, such as neurologic diseases (for example, Parkinson's disease and Alzheimer's disease). This may have hampered the possibility of identifying associations between exposure to agricultural pollutants and these diseases. **Chapter 4** was based on data from the national health survey from 2012, where information on (chronic) health conditions is obtained from a sample of the population. Outcome information was

based on self-report and did not include the date of diagnosis. Such information allows for an inference on the prevalence of the inquired diseases but hampers the possibility of performing a longitudinal study, since it precludes assignment of exposure prior to developing the outcome. A general overview of the advantages and disadvantages of several data sources for outcome assessment is displayed in Table 2.
Finally, because this work was based on registries and other administrative data, it was not possible to investigate other outcomes of scientific interest or of public health relevance (for example, cognitive problems in children). In the Netherlands, communicable and non-communicable diseases are registered at the level of hospitals, general practitioners (GPs), laboratories and surveillance institutes, such as the National Institute for Public Health and the Environment (RIVM), and then reported to the municipalities' Public Health Department (Gemeentelijke Gezondheidsdiensten, GGD).

Information on other relevant covariates

Although it was possible to include information on several factors that are featured in registries, information on lifestyle factors that could be potential confounders was not available (except in **chapter 4**). However, it is unlikely that lifestyle factors are associated with living in proximity to a specific type of livestock farm (e.g. pigs but not poultry), a specific crop type (e.g. summer barley but not winter wheat) or a specific active ingredient (boscalid but not fluopyram). In addition, it is debatable whether one should expect lifestyle-related confounding differentially affecting results depending on whether someone lives within the short distances analysed in this thesis. This would entail lifestyle differences e.g. of people living within 50m as compared to living between 50–100m from a crop. Further analyses were conducted to evaluate the robustness of the results to uncontrolled potential confounders across all studies. Given that contrasts in lifestyle factors clearly exist between rural and urban areas, all chapters included a sensitivity analysis restricted to the most rural areas.
In this analysis, people living in semi-urban areas that could potentially have lifestyles more similar to those living in urban areas were exclude. No major changes in effect estimates, compared to the main analyses, were observed. Other analyses, such as the negative control analysis previously mentioned, and an analysis using a small data set for which lifestyle factors were included in the models, also did not provide clear evidence that lifestyle factors were associated with exposure to agricultural pollutants and thus markedly bias the association of these exposures with the health outcomes investigated. Naturally, there is still potential for residual confounding and there are other approaches to assess confounding such as indirect adjustments to the models and computing E-values, which are revisited later in the outlook section of this chapter.

Biases and statistical approach

The exposure misclassification resulting from the rather simple exposure models used was likely non-differential therefore resulting in an overestimation of exposed people and biasing the estimates towards the null, i.e., an underestimation of the effect size. Consequently, there is the possibility of signals having been missed due to the similarity in exposure levels of exposed and unexposed groups, and the identified signals, if true, may in fact correspond to higher risks for the population.

While using big data provides large power to detect small effect sizes, these are tendentially statistically significant and may not correspond to relevant health effects or clinical significance. An example of this is the small increase in gestational age or birth weight observed in babies whose mothers were exposed to some active ingredients during pregnancy (**chapter 5**). Mothers who were highly exposed to linuron (90^{th} percentile of exposure to linuron) had babies 12 (500m buffer) to 26 (50m buffer) grams heavier than babies whose mothers were not exposed. These are small effects when compared to the decrease in birth weight in mothers that smoke during pregnancy (around 200 grams). However, it is important to bear in mind that, although the clinical relevance these small effects may not be impactful on an individual level, they can translate in important shifts in the distribution of the outcomes among the exposed population. In the case of birth weight, this could translate into a shift in the mean birth weight exposed population that could result in a higher proportion of babies being classified as large for gestational age. As mentioned before, exposure misclassification may have led to an underestimation of the effects and careful interpretation of the results is warranted. The approach taken in this thesis has an exploratory character with no attempt to quantify effects *per se* and rather than using clinical minimal values, all results that were robust (that is, indicative of potential associations) were reported. Clinical relevance and impact in public health should be considered in subsequent (confirmatory) analyses.

On the other hand, to mitigate the tendency to consider individual statistically significant result as findings, results were interpreted in a careful and holistic manner. While this gives confidence about the signals found, it is arguable that criteria for interpretation used were rather stringent and that associations pointing to potentially harmful effects were overlooked.

Another point that could raise some critique regarding the methods applied in the chapters using area of a specific crop as the exposure metric is the fact that these exposure metrics were used as continuous variables in the models and it is arguable that their relationships with the outcomes are not linear and/or that a comparison should be made between the different exposure levels. This was in fact explored in a report

for the Ministry of Health, Welfare and Sport, where it was determined that for some specific crop and outcome pairs, the relationship was indeed not linear, but the same conclusions were overall reached: only living near maize crops showed evidence of being at higher risk of death from chronic lower respiratory diseases.[6]

THIS THESIS IN THE CONTEXT OF WORK ON THE TOPIC OF AGRICULTURAL POLLUTANTS AND HEALTH

The observations of lower mortality, lower risk for depression and lower risk of bad perceived health in the people living near crops are in contrast to the expected direction of effects, although inverse associations between adverse health outcomes and other types of environmental exposures, including air pollution and livestock emissions, have previously been observed among the Dutch rural populations.[7,8] It remains unclear what is driving the observed overall lower risks, but possible explanations lie with potential for residual confounding likely biasing the effects towards the null. Results that differed from this general tendency were all the more noteworthy.

Chapter 3 revealed indications for an association between living near maize crops and higher risk of chronic lower respiratory mortality. It is unclear whether this effect is related to exposure to the pesticides used in maize crops or to other characteristics specific to this crop that could adversely affect respiratory health. Regardless, results are in line with findings from both occupational studies identifying higher risks for COPD and asthma in people exposed to pesticides and from a study identifying higher mortality in maize farmers due to production related air pollution (namely ammonia from nitrogen fertilizers).[9] In **chapter 2** people living near pig farms were also at higher risk of death from chronic lower respiratory diseases and associations between pig farm emissions and respiratory diseases have been reported in studies assessing occupational exposures[10] Results from **chapter 5** showed that linuron and vinclozolin, pesticides with antiandrogenic effects, were associated with higher birth weight, higher risk of being large for gestational age (linuron) and longer gestational age (vinclozolin), especially in girls. It is unclear why we observed these effects especially in females, but it is likely that the balance in sex hormones during gestation is important for fetal growth. Interestingly, even though results across studies are inconsistent, most indicate higher risks for lower birth weight when pregnant women are exposed to pesticides and endocrine disrupting chemicals[11–13]. In this chapter, we also observed that glufosinate-ammonium was associated with higher risk of having babies with low birth weight, fluroxypyr-meptyl was associated with longer gestational age and picoxystrobin was associated with longer gestational age and with higher risk of having

large for gestational age babies. The underlying mechanisms for these effects remain unknown as well, especially since toxicological studies either show no evicence of reproductive and/or developmental effects or point to adverse birth outcomes that are not in line with the observed effects, such as foetal death and premature deliveries and reduced body weight[14–16]

Overall, the effect sizes of the signals identified throughout the chapters were small, but even a small change in risks can have significant impact given the large number of people affected. In other words, environmental factors, such as those explored in this thesis, causing small upward shifts in disease risk or burden distributions in an exposed population translates into a substantial increase of the proportion of individuals with the disease. Generalization of the results of this work to other countries is not straightforward. The Netherlands is a densely populated country where people live closer to agricultural plots and farms than in most other countries. International generalization of the results from the chapters using crops as proxy for exposure to a mixture of pesticides is especially hampered by differences in agricultural settings. For example, differences in climatic zones determine the type of crops grown, the type of pesticides used and the timing of their application. European countries are bound to regulations that differ from those applied in other countries regarding agricultural techniques (such as prohibition of aerial spraying) and approved active ingredients. Results from the livestock study (**chapter 2**) may be more generalizable. Albeit relying on data from nearly 20 years ago and farming practices having evolved since then, the type of emissions produced by livestock farming is likely qualitatively similar today and similar to those in other countries (except specific pathogens). Given that adverse respiratory health effects are currently observed in people living near farms, the result on higher risk of respiratory mortality in people living near pig farms should be a research priority. Similarly, the results from the chapter on birth outcomes are more easily generalizable to other countries since specificity of exposure assessment was relatively higher (specific active ingredients in a short but specific time window of exposure (pregnancy)). It follows that it is important to replicate this result and to elucidate the biological mechanisms underpinning the observed effect.

Many of the studies on the topic of the effects of agricultural pollutants on health have taken a range of different, yet complementary, approaches, with the limitations of one approach covering the limitations of another. On one hand, and similarly to the approach taken in this thesis, there are studies covering large areas (states or countries) and using complete data from several administrative sources. For example, a study in California, USA, estimated residential exposure to relevant active ingredients based on land use maps and the pesticide usage in each field from the California's Pesticide Use Report (PUR) to assess their association to adverse birth outcomes (using information from the birth registry)[17] Another example is the Dutch cross-sectional study on the ef-

fects of residential proximity to livestock farms on COPD, where researchers assessed the association between distance to livestock farms and use of COPD medication, using medication purchase information registered by health insurance companies[18] In these large studies, exposure assessment was based on modelled estimates of exposure, with the limitations discussed above, and information on potential confounders was limited, while simultaneously struggling with interpreting statistically significant results of very small effects that are possibly clinically irrelevant. Complementary to the caveats of large epidemiological studies are smaller studies that are able to use (personal) measurement for a more precise exposure assessment and to collect detailed information on several potential confounders. However, their lower statistical power or cross-sectional design can limit their capacity to provide statistically significant results for the small effect sizes that are expected or to provide strong evidence of causality. In the end, the fact that different types of studies provide different types of information underscores the importance of interpreting results from all approaches in a holistic manner. Such is the approach taken on the EXPANSE project (and, at a smaller scale, in **chapter 3** with the use of a smaller cohort study to investigate the impact of adjusting for lifestyle factors). In EXPANSE, researchers investigate how the urban exposome, i.e., all environmental factors people are exposed to in cities, affects health. They have triangulated the result from studies using large administrative cohorts with very limited phenotyping, prospective cohorts with less subjects but better phenotyping, and molecular studies including just a few people where very deep phenotyping was performed, to understand what and how environmental factors pose potential health risks. Finally, of note is the fact that the main route of exposure to pesticides is food intake, which was not considered in any of the chapters concerning health effects of pesticide exposure. It is difficult to assess what the impact of this route of exposure can be among the rural population. On one hand, people in rural areas may be more likely to grow their own vegetables, a practice that can be even more common close to agricultural fields. Their food gardens may be contaminated by pesticides applied in fields nearby or people might use pesticides themselves, possibly ending up using pesticides and fertilizers in a less efficient way than in intensive crop productions. On the other hand, one can also argue that it is unlikely that there are systematic differences in food intake between people living within and beyond certain buffer sizes from a specific type of crop and that, consequently, confounding by food intake may be neglectable.

OUTLOOK

With the results and the limitations encountered in this work in mind, clear directions for future work for the study of agricultural pollutants and health in the Netherlands can be identified. These belong to three major steps discussed in this subsection.

Important steps in exposure assessment

The natural next step in exposure assessment is the use of more precise exposure models that take the physicochemical characteristics of the compounds including the exposure mixtures and their dispersion patterns into account. Such models are fundamental tools for residential exposure assessment in large studies and there have been efforts to develop such models in the last years.

Deterministic exposure models have been developed by Figueiredo et al. for pesticides used in flower bulbs in the Netherlands[19] The models were developed using data from downward spraying events and were validated for such applications. They are probably suited for lateral spraying but their performance in crops where upward spraying of pesticides is used (for example in fruit tree crops) must be assessed. A practical limitation of using these models for exposure assessment is that input data is not registered in necessary detail on a routinely manner. Those data refer to information on exact time and location of spraying, and exact amounts and types of pesticides used on a specific field. In the Netherlands, farmers must note what pesticides are used in their fields but this information is not accessible, hence the need to rely on quadrennial surveys to estimate the amounts used in specific types of crops. However, given that farmers already register these activities, a solution could be the use of application programming interfaces (APIs) to retrieve and automatically pre-fill the already established yearly agricultural census (Landbouwtelling). This would, in turn, allow the construction a pesticide usage database that would not only monitor pesticide usage but could also help identify challenges farmers face in their production and effective alternatives to pesticides. A prototype of such a software has been recently developed for information regarding crop harvesting but the use of different farm management software among farmers is a challenge regarding APIs development, protection of sensitive data and harmonization of data.

de Rooij et al. developed a land-use regression (LUR) model and a dispersion model for livestock emitted PM_{10} (particulate matter ≤10µm) and its endotoxin content. Both models were developed using data in the context of the Livestock Farming and Residential Health (VGO) program, which is set on a region of the Netherlands with high density of livestock farming. These models are specific for the regions covered in the

VGO program and lack validation in other parts of the Netherlands. A first step is then, naturally, validation of these models for national level use. There is also space for improvement of these spatial models, by including for example temporality (an important feature in the study of the effect of short-term and long-term exposures) or developing a hybrid model that combines both the LUR and dispersion modelling techniques[20] Spatio-temporal models require detailed information not only on the number, type and location of livestock animals, but also on time-varying factors such as farm practices (housing system, type ventilation systems and hygiene practices) and health status. Therefore, a crucial step would be the creation of a national database with such information, similarly to what was proposed above on the registration of pesticide usage in crops and an extension of the national agricultural census.

Finally, there are other common limitations regarding the use of these advanced pesticide and livestock emissions exposure models. Firstly, these models require various input information (such as properties of the active ingredients in the pesticide model or particles' emission data for different sources in the livestock dispersion model) which may be missing and estimated or published values must be used instead, adding to the uncertainty of the output of the models. Secondly, it is difficult to assess whether they are suited to estimate distant past emissions when measures for reduction of agricultural activities pollution were not as widely implemented nor regulated as today. Thirdly, they don't consider all pathways that lead to residential exposure (such as take-home pathways). Lastly, estimating exposure to several compounds for studies with large populations can be computationally challenging. Together, the described limitations restrict current use of these models in studies but their further development and improvement would greatly contribute to the assessment of residential exposure to agricultural pollutants in future research.

Important steps in health research

A natural follow-up from the work of this thesis is the investigation of the signals identified, namely whether these signals correspond to true effects. These include research on the associations between pollutants from maize production and respiratory mortality, pollutants from pig farms and respiratory mortality and maternal exposure to fluroxypyr-meptyl, glufosinate-ammonium, vinclozolin and picoxystrobin and several potentially adverse birth outcomes, namely the higher birth weight, longer gestational age and increased risk of having large for gestational age babies. Of note is that, except for fluroxypyr-meptyl, none of these pesticides are currently approved for use in the European Union (EU) but there are other active ingredients belonging to the same chemical group or having the same mode of action that are nowadays used in the EU and should

be investigated. For example, linuron is a carbamide that is banned in the EU and, although many other carbamides are banned, some, such as chlorotoluron and metobromuron, are currently approved for usage. Investigating such compounds, for example in *in vitro* studies in a read-across approach can be useful in filling in knowledge gaps in biological mechanisms and pathways leading to (potentially adverse) effects.
There are also other health outcomes that have been reported to be linked to pesticide exposure that were not investigated in this work, such as incidence of Parkinson's disease (mortality was addressed in this thesis), congenital malformations, incidence of COPD, incidence of leukemia in both adults and children, cognitive problems in children, acute health problems (for example acute respiratory or neurologic symptoms, dermatologic or ocular lesions), and exacerbation or faster progression of underlying diseases such as COPD and Parkinson's disease.
There has been recent discussion on other health problems being related to agricultural pollutants, such as the link between exposure to pesticides and heart arrythmias or how the loss of biodiversity due to current pesticide use and livestock farming practices may indirectly affect human health. This is an indication that there is much on the effects of agricultural pollutants that remains unknown and unexplored, and it is crucial that we learn more about them. Since the agricultural sector is growing to ensure food production for a worldwide growing population, it is important that we have a deeper knowledge on safe practices and on safe compounds and their usage.

Important steps in linkage and statistical methods

Data of many of the abovementioned outcomes is available from various institutions in the Netherlands. For example, data on incidence of adult and childhood leukemia can be obtained from the Netherlands Cancer Institute (NKI). Many communicable and non-communicable diseases are often registered at the level of hospitals, general practitioners (GPs), surveillance institutes and health insurance companies. The Vektis and the Dutch Hospital Data (DHD) are examples of how data from health care claims and data from hospitals and medical centers, respectively, can be collected, wrangled, stored and later used for research purposes. These databases are good candidates for obtaining information on the incidence of most of the chronic and neurologic diseases previously mentioned. Information on acute outcomes or exacerbation of symptoms can be obtained from medical registries (GPs) or ongoing studies. It is possible to link most of the available databases and data sets from institutions listed above as well as newly created data sets to administrative data and obtain information on residential history and other important covariates that may not be otherwise available (for example, household wealth, marital status, migration background, and ecological varia-

bles at neighborhood level such urbanization degree). Linkage of several (large) data sources provides important advantages in epidemiological studies, by supplying data on medically assessed health outcomes and objectively assessed and systematically updated relevant covariate information. Statistics Netherlands (CBS) provides a digital infrastructure that allows for such linkages and that guarantees compliance with GDPR and ethics regulations, provided that appropriate agreements and contracts between all institutions involved are established. There have also been other large efforts to harmonize data from several cohorts nationally, such as the Netherlands Cohorts Consortium (NCC), and internationally, such as the European Human Exposome Network (EHEN), which further strive to make data FAIR (Findable, Accessible, Interoperable and Reusable). Future studies on the topic of health effects of agricultural pollutants, in particular research on the outcomes not featured in this thesis, would greatly benefit of such linkage approaches.

The databases mentioned above do not contain information on lifestyle factors, which could modify the association between agricultural pollutants and several of the outcomes of interest. A solution to overcome the lack of information on such potential confounders when using administrative sources of data might be the use of indirect adjustments methods using information from health surveys or questionnaires. These methods rely heavily on sometimes unverifiable assumptions, such as that the association between exposure and the unmeasured confounder is linear or that the prevalence of the unmeasured confounder is similar across strata of the exposure and of other measured confounders[21–23] The latter assumption implies thus that the registry and the ancillary survey populations are comparable but comparability between these populations is more often than not difficult to achieve. However, it may be (at least partially) overcome by random stratified sampling of the registry population to create groups that resemble the survey population[24] Another less cumbersome option to evaluate the risk of confounding is computing E-values, a method that helps quantifying how substantial the unmeasured confounding would have to be (how strongly associated with both the exposure and the outcome) that the observed effect estimate could be explained away (i.e., negate the finding). As the authors of this method put it, "the higher the E-value is, the stronger the unmeasured confounding must be to explain the observed association"[25] E-value computation is an intuitive and easy tool that does not require any assumptions about the unmeasured confounder. Reporting E-values in large observational studies, where there is a tendency to find statistically significant but spurious associations, may be very useful as they help characterize the evidence strength for causality[26] However, their interpretation depends on knowledge about the strength of the associations between the unmeasured confounders and the outcome and the exposure, so that the researcher can compare them to the E-value. This is often

unknown, such as in the case of this thesis, where the associations between unmeasured confounders and the several agricultural pollutants is unknown.
There is also much to explore regarding statistical analyses of agricultural pollutants and of environmental exposures in general. So far, most epidemiological studies focus on exposure to a single compound or to a few compounds of the same source or nature (air pollutants, pesticides, industrial chemicals). While it is still important to evaluate the effects of individual compounds, it crucial to evaluate their interaction and combined effects to characterize real-life exposures more accurately. People are exposed to a range of chemicals and evidence from several studies support the dose-additivity concept, that is, the combined effect of simultaneous exposure to several chemicals with similar modes of action is larger than the effect of one mixture component alone[2727] Concomitantly, research also points to important adverse effects from (low dose) exposure to compounds that have dissimilar modes of action but that produce the same adverse outcome.[28,29] This highlights the importance of expanding research focus from groups of compounds with similar modes of action to considering several compounds that act through different pathways to cause an adverse health outcome. Accordingly, there has been growing interest in the "chemical mixture problem" in environmental epidemiology, with an expanding body of literature focused assessing and quantifying the combined effects of exposure to chemicals. In the SPRINT-project, for instance, efforts are made to develop methods to characterize and quantify the effect of pesticide mixtures have impact on biological systems (human, animal, plant and environmental health)[30] Statistical methods that can be used to examine mixtures range from classical Principal Component Analysis (PCA) to more complex techniques using supervised learning, such as Weighted Quantile Sum (WQS) and Bayesian Kernel Machine (BKMR) regressions. These methods are not only able to provide leads for future research by detecting relevant compounds among the mixture but are also capable of estimating the overall mixture effect.

Policy implications

By identifying signals pointing to adverse health effects in people exposed to agricultural pollutants, this work raises important questions on whether there are specific agricultural pollutants causing these effects or whether the culprit is the mixture of compounds itself. It is crucial that these questions are answered to implement appropriate preventive actions. But, perhaps more importantly, it is imperative to obtain deeper understanding on the relevance of agricultural pollutants not just to these direct health effects but to broader issues such as biodiversity, economic growth and social equity, much of which we currently cannot foresee. Politicians in Europe seem to have now

given the topic of agricultural practices management its due attention, and strategies and actions from the European Green Deal, such as Farm to Fork or Zero Pollution, are first steps into finding solutions to balanced agricultural activities[31,32] It is still unknown how these actions will impact European agriculture. On the one hand, efforts are being made to reduce agricultural pollutants, such as legislation to reduce the use of pesticides and fertilizers in crops and of antimicrobials in livestock animals, or campaigns to raise awareness of the health and environmental benefits of reducing the amount of animal products in the diet. On the other hand, the predicted increase in global human population in the next decades raises concerns about food security and pesticides may become more relevant in ensuring food production. Consequently, there may be increased health risks regarding pesticide usage and knowledge on safe compounds and safe usage becomes the more crucial. In the end, the aim of these strategies is to aid in the transition to a more sustainable food system that will significantly contribute to improvements in public health, economic growth, food security and a reversal of loss of biodiversity.

CONCLUSION

The study of the effects of residential proximity to sources of agricultural pollutants on health is a complex topic, with scientific, societal and political challenges. From a scientific point of view, this complexity pertains not only to the unknown magnitude of the health impacts these pollutants have but also the uncertainties inherent to the methods currently being used for exposure assessment and statistical analysis. The work in this thesis constitutes a starting point, providing valuable insights on potential health issues related to agricultural pollutants, namely pesticides and livestock emissions, but it admittedly uncovers only a small section of the iceberg. The common approach between the studies comprising this thesis was the use of administrative registries and surveys for both exposure and outcome assessment. These databases are useful to explore associations and detect weak signals, as well as investigating rarer outcomes with potentially high impact to society. However, due to limitations on exposure assessment and information on relevant outcomes and covariates, these databases become less suited for providing quantification of effects. To have a deeper understanding of these effects, combining different approaches to the topic is required. Only by combining complementary research, possibly in a triangulation approach within a large (inter)national project, will we be able to have a complete and accurate image of all aspects that this complex topic involves, from exposure, to disease pathways, to appropriate epidemiological designs and statistical methodologies.

The exploratory approach taken in this work identified a few signals pointing to associations to (adverse) health outcomes. Specifically, analysis showed higher risk of respiratory mortality in people living near maize crops and near pig farms, and larger baby (girls) when mothers were exposed to linuron, fluroxypyr-meptyl, glufosinate-ammonium, vinclozolin and picoxystrobin during pregnancy. This work also provided valuable information to determine priorities for future work. First, the identified methodological and statistical challenges have shown the need for high quality exposure assessment and use of statistical approaches that handle the effects of mixtures. Second, the studies included in the thesis took steps for interpretation that reduced the detection of false positive results and therefore the identified signals constitute a strong basis for hypothesis-based research. In this sense, these results are helpful in steering future work toward identification of culprit compounds or mixtures, so that appropriate measures can be taken. Third, the experience gained while conducting these studies will help setting up new studies on relevant health outcomes that were not featured in this thesis, such as neurologic diseases, incidence of chronic diseases, acute effects and cognitive problems in children. Finally, to truly understand the effects of agricultural pollutants on health it is important that research is conducted in a holistic manner and thus includes efforts to quantify the impacts of loss of biodiversity due to agricultural activities on cornerstones of human health (food security, balanced nutrition, wellbeing and quality of life, and infectious diseases spread).

TABLES

Table 1: State of the art approaches for environmental exposure assessment in the study of the effect of agricultural pollutants on health.

Type	Sample/method	Route[a]	Period	Strenghts[c]	Limitations[c]
PERSONAL MEASUREMENTS	Urine	All routes	Current (Days)	Reflects recent exposure to all routes	Large temporal (almost daily) variability
	Blood	All routes	Current (Days)	Reflects recent exposure to all routes	Large temporal variability + invasive
	Hair	All routes	Past & Current (days to months)	Easy to collect, store and transport + cost-effective	Difficult to translate concentrations in hair to actual exposure
	Handwipe	Dermal contact	Current	Easy to collect + cost-effective	Temporal variability (probability of hand being contaminated)
	Wristbands	All routes	Current & Future (variable)	Easy to deploy, store and transport + captures all routes + cost-effective	Highly dependent on diffusion rates
	Questionnaire [d]	Nd	Past & Current	Easy to collect information	Recall bias and difficult to quantify
	Expert assessment [d]	Nd	Past & Current	Complementary tool, important to fill in data gaps, standardize and integrate expert knowledge	Uncertainty in exposure ranking
ENVIRONMENTAL MEASUREMENTS	Active air sampling (AAS)	Inhalation	Current	Accuracy + High temporal resolution	Cost and difficult deployment
	Passive air sampling (PAS)	Inhalation	Current	Economically viable	Not as accurate as AAS
	Vacuumed floor dust	Dermal contact & dust ingestion	Past & Current		
	(Weeks to years)	Economically viable	Difficult to ascertain exposure time-frame		
	Dust from doormat [placed clean]	Dermal contact & dust ingestion	Current & Future (Defined time-frame)	One can chose which exposure time-frame to capture	Does not capture all exposure routes
	Electrostatic Dust Collector (EDC)	Inhalation & dermal contact	Nd	Easy to collect and cost-effective	Only captures settable dust
	Wipe – indoor surfaces	Dermal contact	Nd	Easy to collect and cost-effective	Large concentration variability between indoor surfaces
	Soil from residential garden	Ingestion and dermal contact	Past & Current (Weeks to months)	Assess a very specific exposure route	Mainly captures home outdoor exposure

Type	Sample/method	Route[a]	Period	Strenghts[c]	Limitations[c]
MODELLING OF ENVIRONMENTAL EXPOSURE	Proximity to source (presence, area, distance or counts)	All routes	Past and current (weeks to years)	Cost effective, no sampling needed	Relies on several assumptions Computes proxies for exposure
	Dispersion modelling and land use regression modelling	All routes	Past and current (weeks to years)	Cost effective, no sampling needed	Requires availability of several sources of data May require intensive computations depending on amount compounds investigated, time period assessed and number of observations

a This is the main exposure route captured according to literature. The sampled matrix might capture other exposure routes.
b If a sample is collected multiple times, then it is possible to assess chronic exposure.
c Some of the most important strengths and limitations.
d These are qualitative methods. The route cannot be determined given that it is dependent on the questions asked (for the questionnaire) and on the type of expert assessment performed.
Nd Not possible to determine given the current literature.
This table is an extension of a table created by Daniel Figueiredo, 2022.

Table 2: Main approaches for outcome and potential confounders assessment in the study of the effect of agricultural pollutants on health

Data source	Strengths	Limitations
Questionnaires / self-report	Easy to collect information Lifestyle information	Participation bias Diagnosis not confirmed by health professional
Biometrics (measurements)	Objective assessment Information collected for the purpose of the study	Expensive and time-consuming Potential for participation bias May require contact between researchers and study subjects
Hospital registries General Practitioners' data bases Birth and death registries	Already available data Routinely collected, complete and updated data Cost-efficient No contact between researchers and study subjects	GDPR restraints Not all outcomes and potential confounders of interest are registered

REFERENCES

1. Dirnagl U. Resolving the Tension Between Exploration and Confirmation in Preclinical Biomedical Research BT – Good Research Practice in Non-Clinical Pharmacology and Biomedicine. In: Bespalov A, Michel MC, Steckler T, editors. Cham: Springer International Publishing; 2020. p. 71–9. Available from: https://doi.org/10.1007/164_2019_278
2. Savitz DA, Wellenius GA, Trikalinos TA. The Problem With Mechanistic Risk of Bias Assessments in Evidence Synthesis of Observational Studies and a Practical Alternative: Assessing the Impact of Specific Sources of Potential Bias. Am J Epidemiol. 2019 Sep;188(9):1581–5.
3. Figueiredo DM, Duyzer J, Huss A, Krop EJM, Gerritsen-Ebben MG, Gooijer Y, et al. Spatio-temporal variation of outdoor and indoor pesticide air concentrations in homes near agricultural fields. Atmos Environ [Internet]. 2021;262:118612. Available from: https://www.sciencedirect.com/science/article/pii/S1352231021004349
4. de Rooij M, Heederik D, van Nunen E, van Schothorst I, Maassen C, Gerard H, et al. Spatial Variation of Endotoxin Concentrations Measured in Ambient PM10 in a Livestock-Dense Area: Implementation of a Land-Use Regression Approach. Environ Health Perspect [Internet]. 2018 Nov 13;126(1):17003. Available from: https://doi.org/10.1289/EHP2252
5. Ward MH, Lubin J, Giglierano J, Colt JS, Wolter C, Bekiroglu N, et al. Proximity to crops and residential to agricultural herbicides in Iowa. Environ Health Perspect. 2006;114(6):893–7.
6. Simões M, Brouwer M, Krop E, Huss A, Vermeulen RCH, Baliatsas C, et al. Gezondheidsverkenning omwonenden van landbouwpercelen [Internet]. Bilthoven; 2018. Available from: https://www.rivm.nl/publicaties/gezondheidsverkenningomwonenden-van-landbouwpercelen
7. Fischer PH, Marra M, Ameling CB, Velders GJM, Hoogerbrugge R, de Vries W, et al. Particulate air pollution from different sources and mortality in 7.5 million adults – The Dutch Environmental Longitudinal Study (DUELS). Sci Total Environ. 2020 Feb;705:135778.
8. Smit LAM, Hooiveld M, van der Sman-de Beer F, Opstal-van Winden AWJ, Beekhuizen J, Wouters IM, et al. Air pollution from livestock farms, and asthma, allergic rhinitis and COPD among neighbouring residents. Occup Environ Med [Internet]. 2014 Feb 1;71(2):134 LP – 140. Available from: http://oem.bmj.com/content/71/2/134.abstract
9. Hill J, Goodkind A, Tessum C, Thakrar S, Tilman D, Polasky S, et al. Air-quality-related health damages of maize. Nat Sustain [Internet]. 2019;2(5):397–403. Available from: https://doi.org/10.1038/s41893-019-0261-y
10. Guillien A, Soumagne T, Dalphin J-C, Degano B. COPD, airflow limitation and chronic bronchitis in farmers: a systematic review and meta-analysis. Occup Environ Med [Internet]. 2019 Jan 1;76(1):58 LP – 68. Available from: http://oem.bmj.com/content/76/1/58.abstract
11. Roberts JR, Karr CJ, Health COE. Pesticide exposure in children. Pediatrics [Internet]. 2012/11/26. 2012 Dec;130(6):e1765–88. Available from: https://pubmed.ncbi.nlm.nih.gov/23184105
12. Khoshhali M, Davoodi S, Ebrahimpour K, Shoshtari-Yeganeh B, Kelishadi R. The association between maternal exposure to organophosphate pesticides and neonatal anthropometric measures: A systematic review and meta-analysis. J Res Med Sci Off J Isfahan Univ Med Sci. 2020;25:79.

13. Street ME, Bernasconi S. Endocrine-Disrupting Chemicals in Human Fetal Growth. Int J Mol Sci. 2020 Feb;21(4).
14. European Food Safety Authority (EFSA). Conclusion regarding the peer review of the pesticide risk assessment of the active substance glufosinate. EFSA J [Internet]. 2005 Apr 1;3(4):27r. Available from: https://doi.org/10.2903/j.efsa.2005.27r
15. European Food Safety Authority (EFSA). Conclusion on the peer review of the pesticide risk assessment of the active substance fluroxypyr (evaluated variant fluroxypyr-meptyl). EFSA J [Internet]. 2011 Mar 1;9(3):2091. Available from: https://doi.org/10.2903/j.efsa.2011.2091
16. The European Food Safety Authority (EFSA). Peer review of the pesticide risk assessment of the active substance picoxystrobin. EFSA J [Internet]. 2016 Jun 1;14(6):e04515. Available from: https://doi.org/10.2903/j.efsa.2016.4515
17. Ling C, Liew Z, Von Ehrenstein OS, Heck JE, Park AS, Cui X, et al. Prenatal Exposure to Ambient Pesticides and Preterm Birth and Term Low Birthweight in Agricultural Regions of California. Vol. 6, Toxics. 2018.
18. Post PM, Houthuijs D, Sterk HAM, Marra M, van de Kassteele J, van Pul A, et al. Proximity to livestock farms and exposure to livestock-related particulate matter are associated with lower probability of medication dispensing for obstructive airway diseases. Int J Hyg Environ Health. 2021 Jan;231:113651.
19. Figueiredo DM, Vermeulen RCH, Jacobs C, Holterman HJ, van de Zande JC, van den Berg F, et al. OBOMod – Integrated modelling framework for residents' exposure to pesticides. Sci Total Environ [Internet]. 2022;825:153798. Available from: https://www.sciencedirect.com/science/article/pii/S0048969722008907
20. de Rooij MMT. Livestock-related microbial air pollution : residential exposure and health effects. Universiteit Utrecht; 2019.
21. Strak M, Janssen N, Beelen R, Schmitz O, Karssenberg D, Houthuijs D, et al. Associations between lifestyle and air pollution exposure: Potential for confounding in large administrative data cohorts. Environ Res. 2017 Jul;156:364–73.
22. Shin HH, Cakmak S, Brion O, Villeneuve P, Turner MC, Goldberg MS, et al. Indirect adjustment for multiple missing variables applicable to environmental epidemiology. Environ Res [Internet]. 2014;134:482–7. Available from: https://www.sciencedirect.com/science/article/pii/S0013935114001704
23. Erickson AC, Brauer M, Christidis T, Pinault L, Crouse DL, van Donkelaar A, et al. Evaluation of a method to indirectly adjust for unmeasured covariates in the association between fine particulate matter and mortality. Environ Res. 2019 Aug;175:108–16.
24. Klompmaker JO, Janssen NAH, Bloemsma LD, Marra M, Lebret E, Gehring U, et al. Effects of exposure to surrounding green, air pollution and traffic noise with non-accidental and cause-specific mortality in the Dutch national cohort. Environ Heal [Internet]. 2021;20(1):82. Available from: https://doi.org/10.1186/s12940–021–00769–0

25. Haneuse S, VanderWeele TJ, Arterburn D. Using the E-Value to Assess the Potential Effect of Unmeasured Confounding in Observational Studies. JAMA [Internet]. 2019 Feb 12;321(6):602–3. Available from: https://doi.org/10.1001/jama.2018.21554
26. Mathur MB, Ding P, Riddell CA, VanderWeele TJ. Web Site and R Package for Computing E-values. Epidemiology [Internet]. 2018 Sep;29(5):e45–7. Available from: https://pubmed.ncbi.nlm.nih.gov/29912013
27. Martin O, Scholze M, Ermler S, McPhie J, Bopp SK, Kienzler A, et al. Ten years of research on synergisms and antagonisms in chemical mixtures: A systematic review and quantitative reappraisal of mixture studies. Environ Int [Internet]. 2021;146:106206. Available from: https://www.sciencedirect.com/science/article/pii/S0160412020321619
28. Kortenkamp A, Faust M, Scholze M, Backhaus T. Low-level exposure to multiple chemicals: reason for human health concerns? Environ Health Perspect [Internet]. 2007 Dec;115 Suppl(Suppl 1):106–14. Available from: https://pubmed.ncbi.nlm.nih.gov/18174958
29. EFSA Panel on Plant Protection Products and their Residues (PPR). Scientific Opinion on the relevance of dissimilar mode of action and its appropriate application for cumulative risk assessment of pesticides residues in food. EFSA J [Internet]. 2013 Dec 1;11(12):3472. Available from: https://doi.org/10.2903/j.efsa.2013.3472
30. SPRINT, Sustainable Plant Protection Transition [Internet]. [cited 2022 Oct 13]. Available from: https://sprint-h2020.eu/
31. European Union Commission. Farm to Fork strategy [Internet]. [cited 2022 Oct 13]. Available from: https://food.ec.europa.eu/horizontal-topics/farm-fork-strategy_en
32. European Union Commission. Zero pollution action plan [Internet]. [cited 2022 Oct 13]. Available from: https://environment.ec.europa.eu/strategy/zero-pollution-action-plan_en#:~:text=The zero pollution vision for,creating a toxic-free environment.

APPENDICES

SUMMARY

With growing political, scientific and public interest on the health effects of agricultural pollutants among the general population, it has become important to gain evidence-based knowledge about the potential association between these pollutants and health. This is especially relevant for rural populations, who are the most exposed to agricultural pollutants given their residential proximity to these pollutants´ sources. In the Netherlands, a very densely populated country, about 35% of the population lives in rural areas and their residences are located close to livestock farms and crop fields where pesticides are applied. Because of the large number of exposed individuals, even the typically small effects of environmental exposures, such as agricultural exposures, can have important public health impacts. This thesis, "Residential exposure to agricultural pollutants – the use of administrative data to explore the health of the Dutch rural populations", makes use of large national administrative databases to explore associations between exposure to agricultural pollutants, namely pesticides used in crops and livestock farming emissions, and health outcomes among the Dutch rural population. The aim was to identify possible associations between living close to sources of these pollutants and cause-specific mortality, birth outcomes, psychological distress and perceived health that could point to more specific (groups of) exposures for further confirmatory research.

Chapter 2 of this thesis assesses the potential associations between respiratory mortality and residential proximity to (intensive) livestock farms, considering both the species and the number of animals being raised. From the Dutch Environmental Longitudinal Study (DUELS), a national census-based cohort, we selected about 4 million study subjects living in rural areas and that did not change address in the period 1999–2003 to build a retrospective cohort. For each of these individuals´ residences, we computed the average number of cattle, pigs, chicken, and mink present in 500m, 1000m, 1500m and 2000m. We followed the individuals over 8 years for respiratory mortality. Using Cox proportional hazards regression and adjusting for potential confounders at individual and neighborhood level, we found evidence that living up to 2000m of pig farms was associated with increased respiratory mortality, namely from chronic lower respiratory diseases.

Chapter 3 uses the same cohort, but further restricted to people that did not change addresses in the period 1995 to 2003, which amounted to a study population of ~3.1 million. In this chapter, we studied the potential effects of living near pesticide-treated agricultural land and cause-specific mortality. We estimated the area of specific crop groups cultivated within buffers of 50m, 100m and 250m around each individual's resi-

dence and the amount of fungicides, herbicides and insecticides used within the same buffers. We followed the individuals for 25 primary causes until 2012. Using the same statistical approach as in the previous chapter we observed overall decreased mortality risks. In contrast to this overall trend, we noted an increased risk for chronic lower respiratory diseases mortality and proximity to maize cultivation. The analysis focused on amount of pesticides used near residences did not provide evidence of an association with cause-specific mortality. Together, these results indicate that specific pesticides or practices in maize cultivation that may lead to the observed increased risk.

Chapter 4 investigated the association between residential proximity to crops and mental and perceived health in a cross-sectional study. Using the Public Health Monitor national survey from 2012, we selected 216 932 participants who lived in rural areas of the Netherlands. As in the previous chapter, we estimated area of specific crop groups cultivated within buffers of 50m, 100m, 250m and 500m around each individual's residence in the period 2009–2012. Psychological distress (depression) was assessed via the Kessler Psychological Distress scale (K10) and participants were asked to assess their own health. Unlike other chapters using solely registry data, here we were able to include information on lifestyle factors, such as alcohol and drug use, smoking and physical activity. These, and other individual and area-level confounders were included in logistic regression analyses which later revealed no evidence that residential proximity to pesticide treated-crops was associated with psychological distress or poorer perceived health.

In chapter 5 we used the Dutch birth registry to explore associations between residential exposure to specific pesticides during pregnancy and birth outcomes. From all singleton births registered in 2009–2013, we selected those whose mothers lived in rural areas (N=339 947). We estimated the amount (kg) of 139 active ingredients (AIs) used within buffers of 50m, 100m, 250m and 500m around each mother's home during pregnancy. First, we focused on 12 AIs for which there was *a priori* toxicological evidence of reproductive and developmental adverse effects. We investigated their associations with birth outcomes using generalized linear models. These analyses showed that maternal residential exposure to linuron was associated with higher birth weight and higher odds of having large for gestational babies, glufosinate-ammonium was associated with higher risk of having babies with low birth weight, and both fluroxiypyr-meptyl and vinclozolin were associated with longer gestational age. Second, we explored other potentially relevant associations by applying a variable selection method to the pool of the remainder 127 AIs. From these analyses, we found that picoxystrobin was associated with having large for gestational age babies. Although the underlying mechanism driving this effect is unclear, this finding warrants more research into the effects of (non-occupational) exposure to the identified AIs on human health.

Administrative data provide a potentially highly useful resource. With near-complete population coverage (no selection or participation biases), they contain detailed information at individual level regarding exposures, outcomes and demographic characteristics. Big study populations grant sufficient statistical power for detection of subtle effects and for investigation of rare exposures and rare outcomes. Furthermore, administrative data allow for time and cost efficient set-up of retrospective or cross-sectional studies that provide valuable information before setting up huge target studies, which is an important aspect for policy-relevant research. However, a major challenge in using such data sources is that exposure cannot be measured, either because exposure occurred in the past or because it is unfeasible to do so for such large numbers of subjects. Alternatively, as was done in this thesis, exposure was modelled, specifically by computing proxies reflecting the number of animals, the area of specific crops or the amount of active ingredients used around residences was used. These proxies are, however, prone to some degree of exposure misclassification that could have resulted in underestimation of the effects obtained in this thesis. Additionally, administrative data often do not include information on non-communicable diseases, incidence of relevant other health outcomes nor possibly relevant additional information (such as lifestyle factors). Despite these limitations, use of large administrative data is useful in exploratory approaches, such as the one taken in this thesis, that aim to identify relevant associations and narrow the scope of agricultural pollutants candidates for more targeted studies. In smaller targeted studies it is possible to have more detailed information on exposure, outcome and potential confounders and to conduct confirmatory research. Only by combining complementary research approaches will we have a complete and accurate image of all aspects involved in this complex topic of the effects of agricultural pollutants on the health of rural populations. Important next steps in research on this topic include the improvement of exposure models for agricultural pollutants that have been recently developed for the Dutch scenario and further investigation of identified signals in this thesis.

SAMENVATTING

Door de toenemende politieke, wetenschappelijke en publieke interesse naar de gezondheidseffecten van landbouwgerelateerde verontreinigende stoffen onder de algemene bevolking, is het belangrijk geworden om empirisch bewijs te verzamelen over de potentiële associatie tussen deze verontreinigende stoffen en gezondheid. Dit is in het bijzonder relevant voor de populatie die op het platteland woont, omdat zij het meest blootgesteld zijn, gezien hun nabijheid tot de bronnen van verontreinigende stoffen. In het dichtbevolkte Nederland leeft ongeveer 35% van de populatie op het platteland, gebieden die in de buurt liggen van veehouderijen en gewassen waar bestrijdingsmiddelen worden gebruikt. Zelfs de relatief kleine effecten van milieuvervuiling, waaronder landbouwverontreiniging, kunnen een enorme impact hebben omdat het gaat om een groot aantal blootgestelde individuen. In dit proefschrift, getiteld "Residential exposure to agricultural pollutants – the use of administrative data to explore the health of the Dutch rural populations", worden grote nationale administratieve databases gebruikt om de associaties tussen blootstelling aan landbouwverontreiniging en gezondheidsuitkomsten in kaart te brengen, onder de Nederlandse plattelandsbevolking. Met de term landbouwverontreiniging bedoelen we hier namelijk verontreinigende stoffen uit de veehouderijen en bestrijdingsmiddelen die bij gewassen worden gebruikt. Het doel was om mogelijke associaties vast te stellen tussen wonen in de buurt van deze bronnen van verontreinigende stoffen en oorzaak-specifieke sterfte, geboorteuitkomsten, psychisch lijden en de zelf-ervaren gezondheid. De resultaten kunnen als basis dienen voor nader bevestigend onderzoek naar effecten van meer specifieke (groepen van) blootstellingen.

Hoofdstuk 2 van dit proefschrift onderzoekt de mogelijke associatie tussen sterfte aan luchtwegaandoeningen en de nabijheid van (grote) veehouderijbedrijven om de woning, waarbij rekening wordt gehouden met de soort en het aantal gehouden dieren. Uit de Dutch Environmental Longitudinal Study (DUELS), een nationaal census-gebaseerd cohort, selecteerden we vier miljoen personen die op hetzelfde adres op het platteland leefden van 1999–2003 om een retrospectief cohort op te zetten. For alle woonadressen van de deelnemers, berekenden wij het gemiddelde aantal runderen, varkens, kippen en nertsen in categorieën tot 500 meter (m), 1000m, 1500m en 2000m. We volgden deze deelnemers gedurende acht jaar voor sterfte aan luchtwegaandoeningen. Met een Cox proportionele hazards regressie model gecorrigeerd voor mogelijke verstorende variabelen op individueel- en wijkniveau, vonden wij bewijs dat tot 2000m afstand wonen van varkensboerderijen geassocieerd was met verhoogde kans op sterfte aan luchtwegaandoeningen, namelijk door chronische aandoeningen van de onderste luchtwegen.

In hoofdstuk 3 gebruikten wij ditzelfde cohort maar nu verder geselecteerd op personen die op hetzelfde adres op het platteland leefden van 1995–2003, waardoor ongeveer 3.1 miljoen personen overbleven. In dit hoofdstuk onderzochten we de mogelijke effecten van wonen nabij gewassen (die met bestrijdingsmiddelen wordt behandeld) en oorzaak-specifieke sterfte. We schatten het oppervlak van bepaalde typen gewassen die gecultiveerd werden in buffers van 50m, 100m en 250m rondom de woning van participanten. We schatten ook de hoeveelheid bestrijdingsmiddelen d.w.z. schimmel-, onkruid- en insectendodende middelen in dezelfde buffer categorieën. We volgden de participanten tot 2012 voor 25 verschillende oorzaken voor sterfte. Met hetzelfde statistische model als het vorige hoofdstuk vonden wij in het algemeen lagere risico's op sterfte voor participanten die op een afstand tot 500m van de gewassen woonden. In contrast met deze algemene trend, vonden we een verhoogd risico op sterfte door chronische aandoeningen van de onderste luchtwegen voor participanten nabij maïsteelt. We vonden geen associaties tussen de hoeveelheid bestrijdingsmiddelen en oorzaak-specifieke sterfte. Samengenomen geven deze resultaten aan dat bestrijdingsmiddelen of praktijken specifieke voor de maïs-teelt verantwoordelijk kunnen zijn voor het gevonden verhoogd risico.

In hoofdstuk 4 onderzochten wij in een dwarsdoorsnede onderzoek de associatie tussen nabijheid tot gewassen en de mentale en zelf-ervaren lichamelijke gezondheid. Uit een landelijke onderzoek uit 2012, de Gezondheidsmonitor Volwassenen en Ouderen 2012, selecteerden we 216 932 participanten die op het Nederlandse platteland woonden. Net als in het vorige hoofdstuk schatten we het oppervlakte van specifieke groepen gewassen die gecultiveerd werden binnen buffers van 50m, 100m, 250m en 500m rond ieders woning in de periode 2009–2012. Psychisch lijden (depressie) was vastgesteld met de Kessler Psychological Distress scale (K10) en participanten werden ook gevraagd hun eigen gezondheid te beoordelen. In tegenstelling tot andere hoofdstukken met alleen registergegevens konden we hier ook informatie over leefstijlfactoren meenemen zoals alcohol- en drugsgebruik, roken en lichamelijke activiteit. Deze factoren, inclusief de eerdere factoren op individueel- en wijkniveau, werden meegenomen in het logistische regressiemodel. We vonden geen associaties tussen nabijheid van met bestrijdingsmiddelen-behandelde gewassen en psychisch lijden of een slechtere zelf-ervaren gezondheid.

In hoofdstuk 5 gebruikten we data uit het Nederlandse geboorteregister om associaties te onderzoeken tussen blootstelling van omwonenden aan specifieke bestrijdingsmiddelen gedurende de zwangerschap en geboorteuitkomsten. We selecteerden moeders die op het platteland woonden die een eenling baarden en deze registreerde in de periode van 2009–2013 (N = 339 947). We schatten het gewicht (in kilogram, kg) van 139 actieve ingrediënten (AIs) die gebruikt werden in buffers van 50m, 100m, 250m

en 500m afstand van de woning van de moeders gedurende hun zwangerschap. Eerst richtten we ons tot de 12 AIs waar al enig toxicologisch bewijs voor was reproductieve en ontwikkelingsbijwerking kunnen hebben. We onderzochten hun associaties met geboorteuitkomsten met behulp van generieke lineaire modellen. Deze analyses toonden aan dat moederlijke residentiele blootstelling aan linuron geassocieerd was met een hoger geboortegewicht en hogere kans op een groot voor de zwangerschapsduur baby, glufosinate-ammonium was geassocieerd met een hoger risico op baby's met een laag geboortegewicht, en zowel fluroxiypyr-meptyl als vinclozolin waren geassocieerd met een langere zwangerschapsduur. Als tweede stap onderzochten we andere mogelijke associaties voor de overgebleven 127 AIs met een selectiemethode. In deze analyses vonden we dat picoxystrobin geassocieerd was met een hogere kans op een groot voor de zwangerschapsduur baby te hebben. Hoewel het onderliggende mechanismes achter deze effecten onduidelijk zijn, geven deze bevindingen aan dat er verder onderzoek nodig is naar de effecten van (niet-beroepsmatige) blootstelling aan de geïdentificeerde AIs en de menselijke gezondheid.

Administratieve data (zoals nationale registratie databases) zijn belangrijke bronnen voor wetenschappelijk onderzoek. Ze omvatten doorgaans de complete populatie (er vindt geen selectie of participatie plaats) en gedetailleerde informatie op individueel niveau voor blootstellingen, uitkomsten en demografische gegevens. Grote populaties geven de statistische power die vereist is om ook kleinere effecten op de gezondheid aan te kunnen tonen, of onderzoek te doen naar zeldzame blootstellingen of uitkomsten. Verder zijn administratieve data een kosten-efficiënte manier om explorerend onderzoek te doen, die vervolgens bevestigd kunnen worden in ander, doelgericht onderzoek. Een uitdaging in het gebruik van deze databronnen is dat blootstelling vaak niet gemeten kan worden, enerzijds omdat blootstelling in het verleden plaats vond, of omdat het niet haalbaar is om te meten voor een populatie van miljoenen personen. Een oplossing, zoals veelvuldig toegepast in dit proefschrift, is om blootstelling te modelleren, hier door gebruik te maken van proxy's die het aantal dieren, oppervlakte van gewassen en de hoeveelheid AIs rondom de woningen van de participanten schatten, wat vervolgens de verwachte blootstelling reflecteert. Deze proxies blijven echter schattingen en die zijn niet zonder fouten, waardoor misclassificatie van blootstelling plaatsvindt, wat mogelijkerwijs een onderschatting oplevert van de gevonden associaties in dit proefschrift. Administratieve databases verzamelen doorgaans ook geen informatie over niet-overdraagbare ziektes, de incidentie van andere belangrijke gezondheidsuitkomsten of overige belangrijke informatie voor individuen zoals leefstijlfactoren. Ondanks deze beperkingen blijft het gebruik van administratieve databases nuttig voor exploratief onderzoek, zoals uitgevoerd in dit proefschrift, om een selectie te maken van de landbouwgerelateerde verontreinigende stoffen die verder onder-

zocht zouden moeten worden. In kleinere, gerichte studies kan vervolgens op meer detail informatie worden verzameld voor alle belangrijke factoren waaronder blootstellingen, uitkomsten en mogelijke verstorende variabelen om de initiële verdenking te bevestigen. Alleen door deze aanpakken te combineren, kunnen we een volledig en accuraat beeld krijgen van alle aspecten in het complexe onderwerp van de effecten van landbouwverontreiniging op de gezondheid van de plattelandsbevolking. Belangrijke stappen voor vervolgonderzoek voor dit onderwerp zijn onder meer het verbeteren van de modellen om blootstelling aan verontreiniging te bepalen, waar aan gewerkt wordt voor een Nederland-specifiek scenario, en nader onderzoek naar de signalen die zijn gevonden in dit proefschrift.

ACKOWLEDGMENTS

I was about 8 years old when, at dinner table with my family, I had a realization. I observed that while my slim mother and skinny little brother were quite volatile, had a short temper and were prone to throw fits, my father and I were chubby and had a good, calm and pleasant disposition. I promptly concluded and generalized to the world population that skinny people are ill-tempered and fat people are nice. This was my first observational study, n=4. Later, in high school, I believed that learning statistics was a waste of my time and refused to study it. Further down my educational path, during my first Master's, my father advised me to use R for the statistical computations of my final report. Having no coding experience whatsoever, it took me a day to produce a boxplot. I swore there and then that I would never use R again in my life, to which my father replied "Then, try Python".

It is safe to say I have learned some humbling life lessons in the past years. Moving to Netherlands and doing the Master's in Epidemiology at Utrecht University was the turning point for me, personally and professionally. But it was thanks to my supervisors, who took a leap of faith and decided to hire an unexperienced but eager to learn PhD candidate, that allowed me to grow the most. Prof. dr. ir. Roel Vermeulen, Dr. Anke Huss and Dr. Nicole Janssen – thank you for the privilege of learning from you, for your guidance and for teaching me how to become a researcher. I truly enjoyed working alongside all of you these past years on this project, where no corners were cut (and so the buffers remained squared).

Roel, thank you for giving me all the space to conduct my research, allowing me to explore and experiment at my own pace while instilling a great sense of responsibility in me. In our meetings, your input and feedback were always immensely valuable in interpreting and connecting the vast amount of output from my analyses, guiding me on the next steps my work needed. Your vast knowledge and the incredible attention you pay to all projects you are involved in are truly motivating and energizing.

Anke, you have been a mentor and a friend. I have learned a great deal about research and the field of environmental epidemiology from you. I always look forward to our weekly meetings – they are a space for discussion and ideas and have been motivators to continue my work. You often provided me with the (much needed) break from my ramping ideas to do more and more analyses, figures or tables and were able to steer my focus to what maters the most. I appreciate immensely how kind, patient and supportive you are and I would like to underline how important this was for me during my PhD. Unfortunately, I am yet to learn that «"No" is a full sentence», but I do appreciate your efforts to teach me this important lesson.

Nicole, from you I learned not only about environmental epidemiology but also how to navigate the CBS environment. Thank you also for your support throughout these years; I felt that you were always on my corner, as a "knight in shining armor" whenever work was piling up on me. Your astounding ability to look at all those enormous tables full of numbers without being overwhelmed and actually making sense of them will forever amaze me. I admire your passion for environmental epidemiology and the efforts you make every day to make this a better world.

I would also like to extend my gratitude to the assessment committee, Prof. dr. A.M. May, Prof. dr. V. Geissen, Prof. dr. J.A. Stegeman, Dr. Stafoggia and Prof. dr. ir. J. Kromhout, for the time and effort spent in evaluating my thesis.

To co-authors on the RIVM reports and papers featured in this thesis – Lützen, Maartje, Esmeralda, Lidwien, Dick, Jan-Paul, Anton, Maarten, Christos, Daniel and all the other co-authors – thank you for the fruitful meetings, discussions, input and feedback on this joint effort. They were cornerstones of my learning path during the PhD and absolutely essential to produce the work that resulted in this thesis. Daniel, a special acknowledgment to you – what a serendipity to have started our PhD tracks at the same time on sister projects and being conterrâneos. I am only sorry that there were not more opportunities to work together until the very last phase of our tracks. Still, I learned quite a lot from our informal conversations on pesticides.

I could not have had a better working environment than IRAS – a welcoming and open place to work, where high quality research is done. And this is thanks to the wonderful people working there! Ingrid, Christina, Louisa, Petra, Djoeke, and Mieke – thank you for your friendliness and help in all sorts of practical matters. To my former roommates at 320c, Liese, Erik, Jeroen, Fleur, Warner, Yujia – thank you for the *gezellig* environment, for trying to keep my plants alive, for the cheerful lunches (haloumi!) and coffee breaks. Luuk, thanks for your friendship, support and sharing of cat and baby tips (yes, it´s ok to put them in the same category). Of course the list of good colleagues and friends I must thank extends to all IRAS'ers. To name a few, Myrna, Marije, Gijs, Lizan, Jochem, Alejandro, George, Calvin, Joseph, José, Maciek, Ilse, Jules, Samuel, Susan, Inge, Inka, Daniella, Jelle, Roosmarijn and many more – thank you for the walks, lunches, outings, chats and work- and non-work-related help. Kees van Eijden, thank you for introducing me to the supercomputer and putting an end to my suffering of using of 4 workstations for computations.

Living abroad was made easier because of the friends I made in the Netherlands. Patrícia, Andreia, Maria Sofia & Michael, Tajana & João, Honey & Ivo, Levi, Cláudia & Ricardo, Marta, Isabel & Albino, Ana Sofia & family, Stef, Sofia & Gijs - you were my family during my expat years.

Romin and Rik: from talking about food while eating food and playing board games while still thinking about food, to sidetracking into discussions about Epi while we were phi-

losophizing about life, thank you for being there for me since 2013 – that's ten years of friendship! I hope we continue to have plenty of opportunities to "release the chickens" and may we "Mariana face" through life!
To my dear paranimfen, Bernice and Nahid: thank you for your support, friendship and help with the logistics of the defense! Bernice, thank you for the yoga&lunch sessions and for listening to my unsolicited ramps about life (although some of seemingly innocent questions do tend to make one 'spill the beans'…). I only wish we had become friends much sooner, as I feel we I "lost" time we could have spent doing many more fun things! Nahid, I am so glad we ended up the same room and that you became one of my closest friends. Thank you for your dedicated support, listening without judgment and, of course, all the dinners, lunches and talks we shared – I still think they were simply not enough!
To Davidzinho, thank you for having a (very thourogh) read at the introduction.
Vitor, this thesis would not exist if it weren't for your bold move of leaving Portugal and going on an adventure in another country. Thank you for the wonderful 14 years that we have been together, nine of which were spent in the Netherlands. It was here that I feel we grew and learned the most, the country where we built our first home and where our children were born. Maybe we can look at this thesis as another product of our combined efforts, since you also were pivotal for accomplishing this professional milestone. Although I don't think I will ever manage to be a stoic, I will always look up to you for inspiration.
Finally, I would like to thank my family. Who I am today is very much shaped by who I was and how you all prepared me on many aspects of life, allowing me to take this challenge and complete this work. And a special acknowledgment and token of gratitude goes to you, Pai, who are probably more excited than anyone that I am getting my PhD, who patiently watched me make several mistakes, only to end up where you always knew I belonged. This thesis is for you.

Mariana Simões
Figueira da Foz, August 2023

ABOUT THE AUTHOR

Mariana Lobo Simões was born on February 22nd 1985, in Lisbon, Portugal. She attended Colégio Moderno in Lisbon from pre-school to high-school. Starting in 1998, she attended Escola de Música do Conservatório Nacional, following a supplementary music program to that of regular school. In 2003 she started her studies in Veterinary Medicine and for her internship and Master´s thesis, she moved to Chaves, Portugal, for 5 months to work with the Official Veterinary Services on sampling, data wrangling and descriptive statistics regarding brucellosis in small ruminants. While studying she worked as cello teacher and musician, playing at several events. She graduated both the Master´s in Veterinary Medicine and the Secondary Music Course with a major in Cello in 2012. In the summer of that year she moved to the Netherlands where she worked at a call center and as a nanny for a family of 3 boys, while taking courses to learn Dutch. In 2013 she started her second Master´s program in Epidemiology at Utrecht University (UU) and did her internship at the National Institute for Public Health and the Environment (RIVM), supervised by f Dr. Kees van den Wijngaard. There, she worked with data from tekenradar.nl to investigate risk factors for Lyme disease following a tick bite and to evaluate the association between complaints reported after a tick bite and tick infection by *B. burgdorferi* and/or *B. miyamotoi*. In 2016 Mariana started her PhD at the Institute for Risk Assessment Sciences (IRAS) at the UU under the supervision of Prof. Roel Vermeulen, Dr. Anke Huss and Dr. Nicole Janssen. Her initial PhD program focused on the use of administrative databases, namely Microdata from Statistics Netherlands (CBS) and geodatabases on land use, to explore associations between residential proximity to crops where pesticides are applied and several health outcomes. Later, she also investigated the association between residential proximity to livestock farms and respiratory mortality. Her PhD thesis is a combination of studies on the health effects of exposure to these agricultural pollutants. Meanwhile she collaborated on a feasibility study on residential pesticide exposure and relevant health outcomes that were not investigated in her thesis. At the start of 2022, having moved back to Portugal, she began working remotely as a postdoctoral researcher at IRAS on residential exposure to pesticides and incidence of Parkinson´s disease, in collaboration with the Radboud University, and on the project "Air pollution and COVID-19", in collaboration with the RIVM. She also currently provides support to colleagues at IRAS regarding the use of Microdata within CBS´s secure environment.