Abstract

Background: Pesticide exposure might increase risk of lung cancer. The purpose of this study was to investigate the association between the historical use of pesticides commonly found in Thailand, and lung cancer.

Methods: This case-control study compared a lifetime pesticide exposure of 233 lung cancer cases, and 447 healthy neighbours matched for gender, and age (±5 years). Data on demographic, pesticide exposure and other related factors were collected using a face-to-face interview questionnaire. Associations between lung cancer and types of pesticides as well as individual pesticides were analyzed using logistic regression adjusted for gender (male, female), age (≤54, 55-64, 65-74, ≥75), cigarette smoking (never smoked, smoked < 109,500, smoked ≥109,500), occupation (farmer, non-farmer), cooking fumes exposure (yes, no), and exposure to air pollution (yes, no).

Results: It was found that lung cancer was positively associated with lifetime use of herbicides, insecticides, and fungicides. Compared to people in the lowest quartile of number of days using the herbicides and insecticides, those in a higher quartile had an elevated risk of lung cancer, with odds ratio (OR) between 3.31 (95% confidence interval (CI) 1.49–7.34), and 12.58 (95% CI 5.70-27.75) (p < 0.001). For fungicides, only the most exposed group had a significant risk (OR = 4.25; 95% CI
For individual pesticides, those presenting a significant association with lung cancer were dieldrin (OR = 2.56; 95% CI 1.36-4.81), chlorpyrifos (OR = 3.29; 95% CI 1.93-5.61), and carbofuran (OR = 2.10; 95% CI 1.28-3.42).

**Conclusions:** The results showed that lung cancer among Thai people in Nakhon Sawan province is associated with previous pesticide use. This finding was consistent with previous studies in other parts of the world. Further study should focus on identifying more individual compounds that may cause lung cancer, as well as other types of cancer.

**Keywords**
Lung cancer, Pesticides exposure, Herbicides, Insecticides, Fungicides
Introduction
Lung cancer is a common and deadly type of cancer. In 2018, there were 2.1 million people around the world diagnosed with lung cancer, and 1.8 million died of the disease. In 2018, Thailand had 170,495 incidences, and 114,199 deaths of lung cancer. Besides genetic factors, a major risk factor of lung cancer is cigarette smoking. However, lung cancer was also related to other risk factors, including asbestos, crystalline silica, radon, polycyclic aromatic hydrocarbons, diesel engine exhaust particles, chromium, and nickel. Previous studies have also linked cooking fumes to lung cancer.

Pesticide exposure might also cause lung cancer. The association between pesticides and lung cancer were presented around 50 years ago among grape farmers. A large study in the United States found that lung cancer cases increase with the number of years working as a licensed pesticide applicator. Another study in USA reported an increased risk of lung cancer among acetochlor herbicide users (RR = 1.74, 95% CI 1.07-2.84). In Pakistan, a study also found a strong association between pesticide exposure and lung cancer (OR = 5.1, 95% CI 3.1-8.3). Some studies can also link individual pesticides to lung cancer. In the USA, a study evaluated 50 pesticides and found that seven—dicamba, metolachlor, pendimethalin, carbofuran, chlorpyrifos, diazinon, and dieldrin—to be positively associated with lung cancer. Another study also showed a significantly increased risk of lung cancer among applicators who had been exposed to dieldrin. Jones et al. reported an increased lung cancer incidence among male pesticide applicators with the highest exposure category of diazinon (odds ratio (OR) = 1.6, 95% confidence interval (CI) 1.11-2.31). Other individual pesticides that had been associated with lung cancer were chlorpyrifos, diazinon, pendimethalin and carbofuran.

To our knowledge, the association between lung cancer and pesticides has never been studied before among Thai people. The objective of this study was to investigate associations between pesticide exposure and lung cancer among people living in Nakhon Sawan province, Thailand. The results can be used for the prevention of lung cancer, and to support the global literature.

Methods
Study participants
This study is a population-based case-control study. Cases referred to people diagnosed with primary lung cancer during the period of January 1, 2014 to March 31, 2017, and having at least ten years residence in Nakhon Sawan province, Thailand. Cases were selected from the database of Thai Cancer Based Program (TCB) operated by Thai National Cancer Institute. The TCB program requires every hospital to register cancer patients and to provide related information, e.g. types of cancer, diagnostic method, treatment information, etc.

From 299 living cases registered during the study period, 32 died during the year, 20 cases were in stage IV (undifferentiated) cancer, and the other 14 not willing to participate in the study. After exclusion of those cases, 233 (participation rate = 77.9%) were contacted in person, and participated in this study. From 233 cases, 126 were confirmed by Computerized Tomography scan (CT scan)/ Magnetic Resonance Imaging (MRI)/ ultrasound of the whole abdomen/ Chest X-ray (CXR), 62 by histology of primary or metastasis, 21 by cytology of hematology, and 24 by history and physical exam.

Controls were neighbors who did not have lung or any other cancer, but were of the same gender, and age (±5 years) as the cases. In each case, two controls were selected by the interviewer using convenience sampling. In this study, data from 458 controls were used as a comparison group.

The minimum sample size was determined to be 229 for cases and 458 for controls using Kelsey’s formula (unmatched population base case-control study). The assumptions used were as follows: proportion of case with pesticide exposure was 0.5, proportion of control with exposure was 0.4, and the ratio of case to control was 1:2.

Questionnaire
Data on pesticide exposure and other risk factors were collected using a questionnaire previously used in a study on pesticide exposure and diabetes. The questionnaire has two major parts (provided as Extended data in English). Part 1 is about demographic data. We collected data on gender, age, marital status, education, occupation, living duration in the community, distances between home and farmland, exposure to air pollution (i.e., cooking smoke, working in a factory with air pollution; asbestos, diesel engine exhaust, silica, wood dust, painting and welding exposure). Data on smoking status, number of cigarettes smoked per day, and the total number of years smoked was also collected. Number of cumulative cigarettes smoked over a lifetime was calculated by multiplying the number of cigarettes smoked per day by the number of years. Those who smoked <109,500 cigarettes were considered a light smoker, while those who smoked ≥109,500 cigarettes were a heavy smoker.

In Part 2, information on the historical use (mix or spray) of pesticides were collected. In this study, pesticides were categorized into five groups: insecticides (organochlorine, organophosphate, carbamate, and pyrethroid), herbicides, fungicides, rodenticides, and molluscicides. For each group of pesticides, we collected data on the numbers of years and days using pesticides. The data of lifetime pesticide exposure days were then computed by multiplying the total years of exposure by the number of days. Data were reanalyzed using new information on cigarette smoke. More information on lung cancer patients has been provided, including morphology of lung cancer and diagnostics tool for confirmation of the case. On study limitation, bias from exposure misclassification has been further discussed.
days per year. This study also collected data on the use of 35 individual pesticides commonly found in Thailand.

Pesticide exposure data were collected by the researcher and two village health volunteers working full-time for the project. Both case and control were interviewed by the same interviewer. Prior to data collection, all interviewers were trained on how to interview and properly use the questionnaire.

Statistical analysis
Collected data were analysed using IBM SPSS Statistics (version 25) and OpenEpi (version 3.5.1). P values <0.05 were considered statistically significant. Demographic data was analysed using descriptive statistics. The associations were determined between lung cancer and groups of pesticides (herbicides, insecticides, fungicides, and molluscicides), between lung cancer and 17 individual compounds. Both crude and adjusted ORs with 95% confidence intervals (CIs) were presented. Adjusted ORs were analyzed using multiple logistic regressions controlled for gender (male, female), age (≤54, 55–64, 65–74, and ≥75), cigarette smoking (never smoked, smoked < 109,500, smoked ≥ 109,500), occupation (farmer, non-farmer), cooking fumes exposure (yes, no), and exposure to air pollution i.e., working in factories with air pollution (yes, no). In addition to the fundamental confounding factors, variables with statistically difference between cases and controls were included in a regression model.

Cumulative exposure days on groups of pesticides were categorized into quartiles (Q1-Q4; Q1 being the lowest exposure and Q4 the highest). The lung cancer risk was then predicted, using quartile 1 as a reference. For each specific pesticide, exposure data was categorized only to “ever used” and “never used”, but not the cumulative exposure days because number of subjects who reported using each pesticide was too small.

Ethical considerations
This study was approved by the Ethics Board of Naresuan University (project number 550/60). Written informed consent was obtained from each subject before the interviewing process.

Results
Demographic information
In this study, most of study participants were male with a mean age of around 65. Both cases and controls have similar gender, age, marital status, education, occupation, period of residence, distances, pollution exposure, and cigarette smoking. More detailed demographic data among case and control groups were in Table 1 and in Underlying data².

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Case</th>
<th>Control</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (any) (N = 680)</td>
<td>233</td>
<td>447</td>
<td>0.693</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>135</td>
<td>266</td>
<td>0.891</td>
</tr>
<tr>
<td>Female</td>
<td>98</td>
<td>181</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤54</td>
<td>34</td>
<td>71</td>
<td>0.891</td>
</tr>
<tr>
<td>55–64</td>
<td>72</td>
<td>128</td>
<td></td>
</tr>
<tr>
<td>65–74</td>
<td>72</td>
<td>146</td>
<td></td>
</tr>
<tr>
<td>≥75</td>
<td>55</td>
<td>102</td>
<td></td>
</tr>
<tr>
<td>Mean age (years) ± SD</td>
<td>66.04 ± 10.63</td>
<td>65.37 ± 10.88</td>
<td>0.644</td>
</tr>
<tr>
<td>Median age (min-max)</td>
<td>65.00 (37–98)</td>
<td>66.00 (31–92)</td>
<td></td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>17</td>
<td>27</td>
<td>0.295</td>
</tr>
<tr>
<td>Married</td>
<td>188</td>
<td>357</td>
<td></td>
</tr>
<tr>
<td>Divorced/Separated</td>
<td>28</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>Education completed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary school (Grade 1–6)</td>
<td>217</td>
<td>402</td>
<td>0.899</td>
</tr>
<tr>
<td>Secondary school (Grade 7–12)</td>
<td>13</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Characteristic</td>
<td>Case</td>
<td>Control</td>
<td>P value**</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>---------------</td>
<td>-----------------</td>
<td>-----------</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Undergraduate or higher</td>
<td>3</td>
<td>1.3</td>
<td>5</td>
</tr>
<tr>
<td>Occupation</td>
<td>0.970</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farmer</td>
<td>131</td>
<td>56.3</td>
<td>252</td>
</tr>
<tr>
<td>Non-farmer</td>
<td>102</td>
<td>43.7</td>
<td>195</td>
</tr>
<tr>
<td>Period of residence (years)</td>
<td>0.913</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;21</td>
<td>25</td>
<td>10.7</td>
<td>45</td>
</tr>
<tr>
<td>21–30</td>
<td>32</td>
<td>13.7</td>
<td>66</td>
</tr>
<tr>
<td>&gt;30</td>
<td>176</td>
<td>75.6</td>
<td>336</td>
</tr>
<tr>
<td>Distances (m)</td>
<td>0.814</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;500</td>
<td>102</td>
<td>43.8</td>
<td>197</td>
</tr>
<tr>
<td>500–1,000</td>
<td>32</td>
<td>13.7</td>
<td>54</td>
</tr>
<tr>
<td>&gt;1,000</td>
<td>99</td>
<td>42.5</td>
<td>196</td>
</tr>
<tr>
<td>Pollution exposure†</td>
<td>0.636</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>116</td>
<td>49.8</td>
<td>214</td>
</tr>
<tr>
<td>No</td>
<td>117</td>
<td>50.2</td>
<td>233</td>
</tr>
<tr>
<td>Cooking fumes exposure</td>
<td>0.390</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>75</td>
<td>32.2</td>
<td>143</td>
</tr>
<tr>
<td>No</td>
<td>158</td>
<td>7.8</td>
<td>304</td>
</tr>
<tr>
<td>Cigarette smoking</td>
<td>0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never smoked</td>
<td>144</td>
<td>61.8</td>
<td>298</td>
</tr>
<tr>
<td>Smoked (current smoker or ex-smoker)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 109,500</td>
<td>34</td>
<td>14.6</td>
<td>88</td>
</tr>
<tr>
<td>≥ 109,500</td>
<td>55</td>
<td>23.6</td>
<td>61</td>
</tr>
<tr>
<td>Mean (cigarettes) ± SD</td>
<td>175,733±168,868</td>
<td>111,339±107,645</td>
<td></td>
</tr>
<tr>
<td>Median (min-max)</td>
<td>109,500(5,475- 876,000)</td>
<td>87,600(5,475- 812,500)</td>
<td></td>
</tr>
<tr>
<td>Histology</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>114</td>
<td>48.9</td>
<td></td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>17</td>
<td>7.3</td>
<td></td>
</tr>
<tr>
<td>Small cell carcinoma</td>
<td>21</td>
<td>9.0</td>
<td></td>
</tr>
<tr>
<td>Large cell carcinoma</td>
<td>9</td>
<td>3.9</td>
<td></td>
</tr>
<tr>
<td>Neoplasm, malignant</td>
<td>68</td>
<td>29.2</td>
<td></td>
</tr>
<tr>
<td>Other and unspecified</td>
<td>4</td>
<td>1.7</td>
<td></td>
</tr>
</tbody>
</table>

*N was 233 for case and 447 for control unless otherwise indicated.

**χ² test for categorical data; t-test for continuous data with statistically significant (p<0.05).

†Working in factories with air pollution.
Lung cancer and pesticide exposure

After adjusting for confounding factors, lung cancer was positively associated with historical exposure of study participants to herbicides, insecticides and fungicides (Table 2). The adjusted variables included in the analysis were gender (male, female), age (<54, 55–64, 65–74, ≥75), cigarette smoking (never smoked, smoked < 109,500, smoked ≥ 109,500), occupation (farmer, non-farmer), cooking fumes exposure (yes, no), and exposure to air pollution, i.e., working in factories with air pollution (yes, no). Compared with people in the lowest quartile (Q1) of number of days using herbicides, those in Q2-Q4 days of using herbicides had an elevated risk of lung cancer with odds ratio (OR) between 3.31 (95% CI 1.49–7.34) for people with Q2 exposure, and 12.58 (95% CI 5.70–27.75) for Q4 exposure (p < 0.001). A similar association was also found for days of insecticide use and lung cancer (OR = 3.96 for Q3, and OR = 4.13 for Q4, p 0.002). For fungicides, only the Q4 group had a significant risk (OR = 4.25; 95% CI 1.23–14.72).

For individual compounds, lung cancer was statistically associated with a historical use dieldrin (OR = 2.56; 95% CI 1.36–4.81), chlorpyrifos (OR = 3.29; 95% CI 1.93–5.61), and carbofuran (OR = 2.10; 95% CI 1.28–3.42) (Table 3).

Discussion

The study results showed a positive association between lung cancer and the historical use of herbicides, insecticides, and fungicides (Table 3). Herbicides and insecticides have a stronger association to lung cancer than fungicides. Compared to the Q1 group who used pesticides for less than 160 days, demonstrated lower risks of lung cancer than those in Q2 and Q3, who showed elevated risks, while Q4 had the highest risk. (Q4) (OR = 12.58, 95% CI 5.70–27.75). A similar pattern was also observed among the users of insecticides. The risk of lung cancer exponentially increased due to extended periods of using insecticides. (Q3-Q4) with OR between 3.96 (95% CI 1.21–12.97) and 4.13 (95% CI 1.27–15.14). The highest

Table 2. Associations between type of pesticides use and lung cancer.

<table>
<thead>
<tr>
<th>Pesticides use</th>
<th>Case</th>
<th>Control</th>
<th>OR (95% CI)</th>
<th>Adjusted OR (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>233</td>
<td>100.0</td>
<td>447</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Pesticides (any) (N = 490)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Herbicides (N = 347)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>129</td>
<td>54.4</td>
<td>218</td>
<td>48.8</td>
</tr>
<tr>
<td>No</td>
<td>104</td>
<td>44.6</td>
<td>229</td>
<td>51.2</td>
</tr>
<tr>
<td><strong>Number of years using herbicides (N = 347)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;30</td>
<td>23</td>
<td>17.8</td>
<td>51</td>
<td>23.4</td>
</tr>
<tr>
<td>11–30</td>
<td>76</td>
<td>58.9</td>
<td>107</td>
<td>49.1</td>
</tr>
<tr>
<td>1–10</td>
<td>30</td>
<td>23.3</td>
<td>60</td>
<td>27.5</td>
</tr>
<tr>
<td>P-value***</td>
<td>0.192</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of days using herbicides (N = 347)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q4 (&gt;960)</td>
<td>52</td>
<td>40.3</td>
<td>30</td>
<td>13.8</td>
</tr>
<tr>
<td>Q3 (501–960)</td>
<td>45</td>
<td>34.9</td>
<td>50</td>
<td>22.9</td>
</tr>
<tr>
<td>Q2 (160–500)</td>
<td>23</td>
<td>17.8</td>
<td>50</td>
<td>22.9</td>
</tr>
<tr>
<td>Q1 (&lt;160)</td>
<td>9</td>
<td>7.0</td>
<td>88</td>
<td>40.4</td>
</tr>
<tr>
<td>P-value***</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Insecticides (N = 305)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>116</td>
<td>49.8</td>
<td>189</td>
<td>42.3</td>
</tr>
<tr>
<td>No</td>
<td>117</td>
<td>50.2</td>
<td>258</td>
<td>57.7</td>
</tr>
<tr>
<td><strong>Number of years using the insecticides (N = 305)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;30</td>
<td>37</td>
<td>31.9</td>
<td>43</td>
<td>22.7</td>
</tr>
<tr>
<td>11–30</td>
<td>63</td>
<td>54.3</td>
<td>96</td>
<td>50.8</td>
</tr>
</tbody>
</table>
category of years using insecticides (Q4) also showed a positive association with lung cancer (OR = 4.13; 95% CI 1.27–15.14). For fungicides, a significant association was found only among fungicide users in Q4 (>500 days) (OR = 4.25; 95% CI 1.23–14.72).

These results were consistent with literature indicating the potential carcinogenicity of pesticides. In an experimental study, exposure to pesticides caused the production of reactive oxygen species (ROS), an oxygen-containing species containing an unpaired electron, such as superoxide, hydrogen peroxide, and hydroxyl radical, which are highly unstable and may cause DNA damage, protein damage, mutagenicity, necrosis, and apoptosis. Pesticides may also increase the risk of cancer via other mechanisms including genotoxicity, tumor promotion, epigenetic effects, hormonal action and immunotoxicity. In epidemiological study, evidence linked pesticide exposure to lung cancer are increasing, and the issue will be further discussed in the following section.

For individual pesticides, the study found lung cancer to be statistically associated with dieldrin (OR = 2.56; 95% CI 1.23–14.72) and hydroxyl radical, which are highly unstable and may cause DNA damage, protein damage, mutagenicity, necrosis, and apoptosis. Pesticides may also increase the risk of cancer via other mechanisms including genotoxicity, tumor promotion, epigenetic effects, hormonal action and immunotoxicity. In epidemiological study, evidence linked pesticide exposure to lung cancer are increasing, and the issue will be further discussed in the following section.

For individual pesticides, the study found lung cancer to be statistically associated with dieldrin (OR = 2.56; 95% CI 1.23–14.72).
### Table 3. Associations between individual pesticide use and lung cancer.

<table>
<thead>
<tr>
<th>Pesticide</th>
<th>Case</th>
<th>Control</th>
<th>OR (95% CI)</th>
<th>Adjusted OR (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>233</td>
<td>100.0</td>
<td>447</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Herbicides</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Glyphosate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>105</td>
<td>45.1</td>
<td>176</td>
<td>39.4</td>
</tr>
<tr>
<td>No</td>
<td>128</td>
<td>54.9</td>
<td>271</td>
<td>60.6</td>
</tr>
<tr>
<td><strong>Paraquat</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>89</td>
<td>38.2</td>
<td>150</td>
<td>33.6</td>
</tr>
<tr>
<td>No</td>
<td>144</td>
<td>61.8</td>
<td>297</td>
<td>66.4</td>
</tr>
<tr>
<td><strong>2, 4-Dichlorophenoxy acetic acid</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>47</td>
<td>20.2</td>
<td>70</td>
<td>15.7</td>
</tr>
<tr>
<td>No</td>
<td>186</td>
<td>79.8</td>
<td>377</td>
<td>84.3</td>
</tr>
<tr>
<td><strong>Butachlor</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>14</td>
<td>6.0</td>
<td>24</td>
<td>5.4</td>
</tr>
<tr>
<td>No</td>
<td>219</td>
<td>94.0</td>
<td>423</td>
<td>94.6</td>
</tr>
<tr>
<td><strong>Propanil</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>11</td>
<td>4.7</td>
<td>21</td>
<td>4.7</td>
</tr>
<tr>
<td>No</td>
<td>222</td>
<td>95.3</td>
<td>426</td>
<td>95.3</td>
</tr>
<tr>
<td><strong>Alachlor</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>14</td>
<td>6.0</td>
<td>28</td>
<td>6.3</td>
</tr>
<tr>
<td>No</td>
<td>219</td>
<td>94.0</td>
<td>419</td>
<td>93.7</td>
</tr>
<tr>
<td><strong>Insecticides</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Organochlorine insecticides</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Endosulfan</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>35</td>
<td>15.0</td>
<td>44</td>
<td>9.8</td>
</tr>
<tr>
<td>No</td>
<td>198</td>
<td>85.0</td>
<td>403</td>
<td>90.2</td>
</tr>
<tr>
<td><strong>Dieldrin</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>24</td>
<td>10.3</td>
<td>20</td>
<td>4.5</td>
</tr>
<tr>
<td>No</td>
<td>209</td>
<td>89.7</td>
<td>427</td>
<td>95.5</td>
</tr>
<tr>
<td><strong>DDT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>13</td>
<td>5.6</td>
<td>37</td>
<td>8.3</td>
</tr>
<tr>
<td>No</td>
<td>220</td>
<td>94.4</td>
<td>410</td>
<td>91.7</td>
</tr>
<tr>
<td><strong>Organophosphate insecticides</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Chlorpyrifos</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>40</td>
<td>17.2</td>
<td>30</td>
<td>6.7</td>
</tr>
<tr>
<td>No</td>
<td>193</td>
<td>82.8</td>
<td>417</td>
<td>93.3</td>
</tr>
</tbody>
</table>
1.36–4.81), chlorpyrifos (OR = 3.29; 95% CI 1.93–5.61), and carbofuran (OR = 2.10; 95% CI 1.28–3.42) (Table 3). Dieldrin is an extremely persistent organic pollutant linked to many health problems, e.g., Parkinson’s disease, breast cancer, affecting the immunity system, the reproductive, and nervous systems33. In the USA, the Agricultural Health Study found seven pesticides including dicamba, metolachlor, pendimethalin, carbofuran, chlorpyrifos, diazinon, and dieldrin to be positively associated with lung cancer14. Further studies found dieldrin exposure to relate to the highest tertile of days use (RR = 5.30; 95% CI 1.50–18.60)34. In Thailand, 688 tons of dieldrin was used in 1981–1990, before it was banned on May 16, 1990. A study among pest control workers in Florida, USA found a long term exposure to organophosphate and carbamate insecticides to increase mortality risk of lung cancer (OR = 1.4; 95% CI 0.7–3.0) for subjects licensed from 10–19 year; OR = 2.1; 95% CI 0.8–5.5 for those licensed 20 year or more11.

<table>
<thead>
<tr>
<th>Pesticide</th>
<th>Case</th>
<th>Control</th>
<th>OR (95% CI)</th>
<th>Adjusted OR (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Folidol</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>40</td>
<td>17.2</td>
<td>14.3</td>
<td>1.24 (0.80–1.90)</td>
</tr>
<tr>
<td>No</td>
<td>193</td>
<td>82.8</td>
<td>383</td>
<td>85.7</td>
</tr>
<tr>
<td>Mevinphos</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>16</td>
<td>6.9</td>
<td>22</td>
<td>4.9</td>
</tr>
<tr>
<td>No</td>
<td>217</td>
<td>93.1</td>
<td>425</td>
<td>95.1</td>
</tr>
<tr>
<td><strong>Carbamate insecticides</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carbaryl/Savin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>14</td>
<td>6.0</td>
<td>34</td>
<td>7.6</td>
</tr>
<tr>
<td>No</td>
<td>219</td>
<td>94.0</td>
<td>413</td>
<td>92.4</td>
</tr>
<tr>
<td>Carbofuran</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>43</td>
<td>18.5</td>
<td>42</td>
<td>9.4</td>
</tr>
<tr>
<td>No</td>
<td>190</td>
<td>81.5</td>
<td>405</td>
<td>90.6</td>
</tr>
<tr>
<td><strong>Pyrethoid insecticides</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abamectin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>44</td>
<td>18.9</td>
<td>90</td>
<td>20.1</td>
</tr>
<tr>
<td>No</td>
<td>189</td>
<td>81.1</td>
<td>357</td>
<td>79.9</td>
</tr>
<tr>
<td><strong>Fungicides</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Armure/Propiconazole</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>29</td>
<td>12.4</td>
<td>51</td>
<td>11.4</td>
</tr>
<tr>
<td>No</td>
<td>204</td>
<td>87.6</td>
<td>396</td>
<td>88.6</td>
</tr>
<tr>
<td>Methyl aldehyde</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>15</td>
<td>6.4</td>
<td>28</td>
<td>6.3</td>
</tr>
<tr>
<td>No</td>
<td>218</td>
<td>93.6</td>
<td>419</td>
<td>93.7</td>
</tr>
</tbody>
</table>

*Logistic regression adjusted for gender, age (≤ 54, 55–64, 65–74, and ≥ 75), cigarette smoking (never smoked, smoked <109,500, smoked ≥ 109,500), occupation (farmer and non-farmer), cooking fumes exposure (yes, no) and exposure to air pollution (working in factories with air pollution) (yes, no).

**Statistically significant (p < 0.05).
In the Agricultural Health Study, a dose response relationships was found between lung cancer and chlorpyrifos (RR = 2.18; 95% CI 1.31–3.64) and diazinon (RR = 3.46; 95% CI 1.57–7.65). Similar results were also replicated in later studies of the Agricultural Health Study cohort for chlorpyrifos (RR = 1.80; 95% CI 1.00–3.23), which are referring to applicators in the lowest category of exposure. At this time, chlorpyrifos still not banned by Thai government. On the other hand, chlorpyrifos was the primary insecticide imported to Thailand (1,193,302 kilograms in 2013).

For carbofuran, it has demonstrated mutagenic properties in laboratory studies. In the Agricultural Health Study, lung cancer risk of carbofuran for those with >109 days of lifetime exposure (RR = 3.05; 95% CI 0.94–9.87) compared with those with < 109 lifetime exposure days. In Thailand, a study reported 87 different commercial brands of insecticides which were used for 202 rice fields in Suphanburi Province (abamectin 40%, followed by chlorpyrifos 30%, and carbofuran 20%), 93 brands of plant hormones, and 56 brands of chemicals for the control of plant diseases.

The potential limitations of this type of study were the recall bias where cases and controls can recall past exposure differently. Cases tend to memorize exposure better, particularly when they know or are aware of what caused their illness. However, with limited available information on the issues in Thailand, we did not expect participants to be aware of pesticides as causal factor for lung cancer. It was very likely that the study could have exposure misclassification due to the participants not being able to recall or name the pesticides they used in the past. In addition, data on pesticide exposure were obtained solely from the interview questionnaire without any exposure measurement. However, this information bias usually occurs evenly across a case and control group, and only has a negative effect on the association. Selection bias might also occur when using non-random sampling or when the study sample did not represent the study population. However, in this study, data from all of the lung cancer patients except those who were severely ill, were collected to minimize the bias.

Conclusion
This study found that the occurrence of lung cancer among people in Nakhon Sawan province, Thailand is associated with pesticide use. Out of 17 individual pesticides investigated, dieldrin, chlorpyrifos, and carbofuran showed significant associations with lung cancer incidence. These results are consistent with the literature from other parts of the world. Further studies should focus on identifying more individual pesticides that could cause lung cancer, as well as other types of cancer.

Data availability
Underlying data

This project contains the following underlying data:
- Dataset_pesticide and lung cancer (SAV and CSV). (All underlying data gathered in this study.)
- Data Dictionary (DOCX).

Extended data

This project contains the following extended data:
- Questionnaire-pesticide and lung cancer Thailand (DOCX). (Study questionnaire in English.)

Data are available under the terms of the Creative Commons Zero “No right reserved” data waiver (CC0 1.0 Public domain dedication).

Acknowledgments
First, our gratitude goes to the study participants, as without them, this study would not have been possible. We want to thank the village health volunteers for their help in data collection. We appreciate support by Dr. Adisorn Vatthanasak, chief of Nakhon Sawan Provincial Public Health Office. Thank you also to Mr. Kevin Mark Roebl of Naresuan University’s Writing Clinic for editing assistance.

References

Matthew R Bonner
Department of Epidemiology and Environmental Health, School of Public Health and Health Professions, University at Buffalo, Buffalo, NY, USA

Overall, the revised manuscript is responsive to many of my comments. However, there are several outstanding issues that require additional attention.

1) Response: Matching of a few variables can be considered loose-matching, therefore, it is more appropriate to analyze using unconditional logistic regression. Kuo and team (2018) said that “There is a presumption that matched data need to be analyzed by matched methods. Conditional logistic regression has become a standard for matched case–control data to tackle the sparse data problem. The sparse data problem, however, may not be a concern for loose-matching data when the matching between cases and controls are not unique, and one case can be matched to other controls without substantially changing the association. Data matched on a few demographic variables are clearly loose-matching data, and we hypothesize that unconditional logistic regression is a proper method to perform.” (Kuo, Duan & Grady, 2018)*

The characterization of matching on weak confounding factors as “loose-matching” seems to overlook the bias introduced by matching on confounders. As Neil Pearce states: “In essence, the matching process makes the controls more similar to the cases not only for the matching factor but also for the exposure itself. This introduces a bias that needs to be controlled in the analysis” (Pearce N. Analysis of matched case-control studies. BMJ (2016)).

This is a well-recognized consequence of matching in case-control studies and mitigating this introduced bias in the analysis is fundamental to conducting a matched case-control study. The sparse data problem that is emphasized in the response is a common problem for most studies regardless of whether matching procedures were used. The use of conditional logistic regression can help with both these issues and should not be ignored because of a presumption of “loose-matching.” A comparison between the proportional hazards risk estimates with the unconditional logistic regression risk estimates seems to indicate that little if any bias was introduced by matching. This is likely due to a weak association between the matching factors and exposure. The methods section should state that conditional and unconditional logistic regression were used, that the results were similar, and results from the unconditional logistics regression are reported.
2) **Response:** In this study, we actually collected data from 35 individual pesticides, but 17 of them were excluded due to small sample size (less than 5 in each cell). Therefore, the OR groups may be larger than the individual OR ones.

This issue is unresolved, and depending on which table or exposure metric, one can render different conclusions about the association between pesticides and lung cancer. In table 2, herbicide days of use has a strong monotonic exposure response gradient, but in table 3, none of the reported herbicides have an association with lung cancer.

**Additional Comment:**
1) The discussion section on chlorpyrifos and carbofuran describes results from the Agricultural Health Study. However, cited reports are not the most recent reports on these pesticides and lung cancer. A more recent report from 2017, found no association with either chlorpyrifos or carbofuran. The discussion should be updated in light of this more recent report ([https://ehp.niehs.nih.gov/doi/10.1289/EHP456](https://ehp.niehs.nih.gov/doi/10.1289/EHP456)).

**References**

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Cancer epidemiology, pesticides, air pollution, occupational and environmental epidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

---

**Author Response 10 Feb 2021**

**Chudchawal Juntarawijit,** Naresuan University, Muang District, Phitsanulok, Thailand

**Comment:**
Overall, the revised manuscript is responsive to many of my comments. However, there are several outstanding issues that require additional attention.

1) **Response:** Matching of a few variables can be considered loose-matching, therefore, it is more appropriate to analyze using unconditional logistic regression. Kuo and team (2018) said that “There is a presumption that matched data need to be analyzed by matched methods.”
Conditional logistic regression has become a standard for matched case–control data to tackle the sparse data problem. The sparse data problem, however, may not be a concern for loose-matching data when the matching between cases and controls are not unique, and one case can be matched to other controls without substantially changing the association. Data matched on a few demographic variables are clearly loose-matching data, and we hypothesize that unconditional logistic regression is a proper method to perform." (Kuo, Duan & Grady, 2018)*

The characterization of matching on weak confounding factors as “loose-matching” seems to overlook the bias introduced by matching on confounders. As Neil Pearce states: “In essence, the matching process makes the controls more similar to the cases not only for the matching factor but also for the exposure itself. This introduces a bias that needs to be controlled in the analysis” (Pearce N. Analysis of matched case-control studies. BMJ (2016)).

This is a well-recognized consequence of matching in case-control studies and mitigating this introduced bias in the analysis is fundamental to conducting a matched case-control study. The sparse data problem that is emphasized in the response is a common problem for most studies regardless of whether matching procedures were used. The use of conditional logistic regression can help with both these issues and should not be ignored because of a presumption of “loose-matching.” A comparison between the proportional hazards risk estimates with the unconditional logistic regression risk estimates seems to indicate that little if any bias was introduced by matching. This is likely due to a weak association between the matching factors and exposure. The methods section should state that conditional and unconditional logistic regression were used, that the results were similar, and results from the unconditional logistics regression are reported.

**Response:**
The statement that both conditional and unconditional analyses were performed has been added to the method section.

**Comment:**
2) **Response:** In this study, we actually collected data from 35 individual pesticides, but 17 of them were excluded due to small sample size (less than 5 in each cell). Therefore, the OR groups may be larger than the individual OR ones.

This issue is unresolved, and depending on which table or exposure metric, one can render different conclusions about the association between pesticides and lung cancer. In table 2, herbicide days of use has a strong monotonic exposure response gradient, but in table 3, none of the reported herbicides have an association with lung cancer.

**Response:**
Taking the suggestion, we decided to recheck the data and found some individual pesticides to have enough sufficient participants to calculate cumulative exposure days. Reanalysis of the data using quartile of exposure days, found good results. Those exposed to glyphosate and paraquat in Q4 and Q3 had a significant association with lung cancer. So, the issue was solved. Thank you very much for your thoughtful suggestions. Without it, we would miss this important finding.

Also, to make it more consistent, we decided to use ‘nonexposed’ as a reference instead of
Q1, and the reanalysis of the data. This change caused only a minor change in the results.

All parts of the paper (abstract, results, and discussion section) were updated.

**Comment:**

**Additional Comment:**

1) The discussion section on chlorpyrifos and carbofuran describes results from the Agricultural Health Study. However, cited reports are not the most recent reports on these pesticides and lung cancer. A more recent report from 2017, found no association with either chlorpyrifos or carbofuran. The discussion should be updated in light of this more recent report (https://ehp.niehs.nih.gov/doi/10.1289/EHP456).²

**References**


**Response:**

The reference has been updated. Thank you for the information.

**Competing Interests:** No competing interests were disclosed.
inquiring about days and years of pesticide use. Logistic regression, adjusting for potential confounders, was used to estimate the odds ratio and 95% confidence intervals. Pesticide classes (herbicides and organophosphate) and select specific pesticides were positively associated with lung cancer in this study. The authors conclude that “...lung cancer among Thai people in Nakhon Sawan province is associated with previous pesticide use.” Overall, the study seems to be designed well, but crucial information regarding several specific details are missing from the report. These details, and other concerns, are described below.

Comments:
1. Crucial information about the lung cancer cases is missing. Specifically, were the cases comprised of 1st primary lung cancer or were lung cancer cases with a prior history of another cancer, including lung, eligible to participate in the study. Were the lung cancer cases’ diagnosis histologically confirmed? What was the stage and grade of these lung cancer cases? On average, how long after their diagnosis were lung cancer cases interviewed?

2. Were potential controls excluded if they had a prior history of cancer?

3. The methods state that two neighbor controls were selected randomly. This implies that there a sampling frame of some sort. That sampling frame for random selection needs to be adequately described.

4. The controls were matched on age and sex to the cases. This necessitates a conditional logistic regression to account for the selection bias introduced by matching. Unconditional logistic regression is inappropriate for a matched case-control study. Breaking the matching and adjusting for the matching factors may not bias the odds ratios, but this should be confirmed by comparing ORs estimated with unconditional logistic regression and conditional logistic regression.

5. In table 2, there is striking qualitative confounding for the pesticide classes (yes vs. no). For instance, the crude OR organophosphates is 0.63 (95% CI = 0.46-0.87) while the adjusted OR is 1.77 (95% CI = 1.22-2.57). The use of unconditional logistic regression might explain this as the crude estimate for such a regression is inappropriate. That notwithstanding, a number of other variables were included in the regressions. Given this qualitative confounding, additional analyses to identify variables or combination of variables is driving this confounding is warranted.

6. Tobacco smoking is a recognized strong risk factor for lung cancer and a known potential confounder in studies of other exposures and lung cancer. As such, substantial efforts to mitigate confounder are often employed. In this study, smoking was a binary (ever vs. never) variable that may not adequately capture the interrelationship between pesticide use and lung cancer to control confounding. More detailed smoking information, if available, should be explored to determine the potential for residual confounding of the reported associations.

7. In table 1, smoking is not associated with lung cancer. This suggests that selection forces in the recruitment of cases and controls are biasing the study results. It seems unusual that 61% of lung cancer cases were never smokers. Is this a typical feature of lung cancer in...
Thailand?

8. The results reported in tables 2 and 3 seem to be internally inconsistent. For instance, the ORs for organophosphates depicted in table 4 indicate a strong association with lung cancer with days of use (Q4 vs. Q1 OR = 28.43 (95% CI = 11.11-72.76); an extremely large magnitude. However, the ORs for specific organophosphate insecticides are much more modest, although the statistically significant associations with chlorpyfos and dielrin. A similar pattern is evident for herbicide as well. This lack of internal consistency really points to the methodological limitations as a likely explanation for the observed association.

9. Recall bias was discussed as a limitation, but nothing is mentioned about other threats to internal validity. For instance, the potential for selection bias to arise from the recruitment strategies. As mentioned above, the lack of an association with smoking seems to indicate something is awry. In addition, exposure misclassification is undoubtedly present and should be discussed in the Discussion along with the other potential limitations.

10. References 20 and 35 are the same report.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
No

If applicable, is the statistical analysis and its interpretation appropriate?
No

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Cancer epidemiology, pesticides, air pollution, occupational and environmental epidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.
Chudchawal Juntarawijit, Naresuan University, Muang District, Phitsanulok, Thailand

**Comment:**
1. Crucial information about the lung cancer cases is missing. Specifically, were the cases comprised of 1st primary lung cancer or were lung cancer cases with a prior history of another cancer, including lung, eligible to participate in the study. Were the lung cancer cases' diagnosis histologically confirmed? What was the stage and grade of these lung cancer cases? On average, how long after their diagnosis were lung cancer cases interviewed?

**Response:**
The cases comprised of 1st primary lung cancer. The cases were confirmed by Computerized Tomography scan (CT scan), Magnetic Resonance Imaging (MRI), ultrasound of the whole abdomen, and chest radiography or Chest X-ray (CXR), and histology of primary and metastasis. More information was added to Table 1, and provided in the Table below.

On average, the patients were interviewed approximately 1 year after they had been diagnosed with lung cancer.

**Table.** More information on morphology and stage of the study cases.

**Morphology** of lung cancer cases:

<table>
<thead>
<tr>
<th>Morphology</th>
<th>Count (Percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenocarcinoma</td>
<td>114 (48.9)</td>
</tr>
<tr>
<td>Small cell carcinoma</td>
<td>21 (9.0)</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>17 (7.3)</td>
</tr>
<tr>
<td>Large cell carcinoma</td>
<td>9 (3.9)</td>
</tr>
<tr>
<td>Neoplasm/malignant</td>
<td>68 (29.2)</td>
</tr>
<tr>
<td>Unspecified</td>
<td>4 (1.7)</td>
</tr>
</tbody>
</table>

**Stage**

<table>
<thead>
<tr>
<th>Stage</th>
<th>Count (Percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IA, IB</td>
<td>14 (6.0)</td>
</tr>
<tr>
<td>IIA, IIB</td>
<td>38 (16.3)</td>
</tr>
<tr>
<td>IIIA, IIIB</td>
<td>57 (24.5)</td>
</tr>
<tr>
<td>IV (Distant metastasis)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Unknown/unspecified</td>
<td>124 (53.2)</td>
</tr>
</tbody>
</table>
Comment:
2. Were potential controls excluded if they had a prior history of cancer?

Response:
Yes, potential controls were excluded if they had a prior history of cancer.

Comment:
3. The methods state that two neighbor controls were selected randomly. This implies that there a sampling frame of some sort. That sampling frame for random selection needs to be adequately described.

Response:
The control was, in fact, selected using convenience sampling. The information in the manuscript has been revised.

Comment:
4. The controls were matched on age and sex to the cases. This necessitates a conditional logistic regression to account for the selection bias introduced by matching. Unconditional logistic regression is inappropriate for a matched case-control study. Breaking the matching and adjusting for the matching factors may not bias the odds ratios, but this should be confirmed by comparing ORs estimated with unconditional logistic regression and conditional logistic regression.

Response:
Matching of a few variables can be considered loose-matching, therefore, it is more appropriate to analyze using unconditional logistic regression. Kuo and team (2018) said that “There is a presumption that matched data need to be analyzed by matched methods. Conditional logistic regression has become a standard for matched case–control data to tackle the sparse data problem. The sparse data problem, however, may not be a concern for loose-matching data when the matching between cases and controls are not unique, and one case can be matched to other controls without substantially changing the association. Data matched on a few demographic variables are clearly loose-matching data, and we hypothesize that unconditional logistic regression is a proper method to perform.” (Kuo, Duan & Grady, 2018)*

We gained interesting information by analyzing some of the data using Cox regression. The comparison between unconditional analysis and the Cox regression, yielded similar results. (see Table below). Comparing OR between Cox regression and logistic regression).

Table. Comparing OR between Cox regression and logistic regression.

<table>
<thead>
<tr>
<th>Logistic regression</th>
<th>Cox</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Regression

<table>
<thead>
<tr>
<th>Pesticide</th>
<th>OR (crude)</th>
<th>95% CI</th>
<th>OR (adjusted)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endosulfan</td>
<td>1.61</td>
<td>(1.00–2.60)</td>
<td>1.60</td>
<td>(0.97–2.63)</td>
</tr>
<tr>
<td>(1.12-3.63)</td>
<td></td>
<td></td>
<td>(0.99-3.33)</td>
<td></td>
</tr>
<tr>
<td>Dieldrin</td>
<td>2.45</td>
<td>(1.32–4.53)</td>
<td>2.56</td>
<td>(1.36–4.81)</td>
</tr>
<tr>
<td>(1.57-6.85)</td>
<td></td>
<td></td>
<td>(1.70-7.68)</td>
<td></td>
</tr>
<tr>
<td>Chlorpyrifos</td>
<td>2.88</td>
<td>(1.74–4.76)</td>
<td>3.29</td>
<td>(1.93–5.61)</td>
</tr>
<tr>
<td>(1.91-6.41)</td>
<td></td>
<td></td>
<td>(2.05-7.22)</td>
<td></td>
</tr>
<tr>
<td>Carbofuran</td>
<td>2.18</td>
<td>(1.37–3.45)</td>
<td>2.10</td>
<td>(1.28–3.42)</td>
</tr>
<tr>
<td>(1.43-3.89)</td>
<td></td>
<td></td>
<td>(1.41-4.01)</td>
<td></td>
</tr>
</tbody>
</table>


**Comment:**

5. In table 2, there is striking qualitative confounding for the pesticide classes (yes vs. no). For instance, the crude OR organophosphates is 0.63 (95% CI = 0.46-0.87) while the adjusted OR is 1.77 (95% CI = 1.22-2.57). The use of unconditional logistic regression might explain this as the crude estimate for such a regression is inappropriate. That notwithstanding, a number of other variables were included in the regressions. Given this qualitative confounding, additional analyses to identify variables or combination of variables is driving this confounding is warranted.
Response:
After re-categorizing the smoking variable, and correcting a mistake on the variable coding, the new analysis yielded more consistent results with crude OR at 1.35 (95%CI 0.98-1.86) and adjusted OR at 1.40 (95%CI 0.97-2.02).

6. Tobacco smoking is a recognized strong risk factor for lung cancer and a known potential confounder in studies of other exposures and lung cancer. As such, substantial efforts to mitigate confounder are often employed. In this study, smoking was a binary (ever vs. never) variable that may not adequately capture the interrelationship between pesticide use and lung cancer to control confounding. More detailed smoking information, if available, should be explored to determine the potential for residual confounding of the reported associations.

We actually collected data on the amount of cigarettes and smoking duration of the study participants, and then the information was used to compute number of cigarettes smoked by them in their life time. After grouping smoking status into “never smoked”, “smoked <109500 cigarettes”, and “smoked ≥109500 cigarettes”, a significant difference between case and control was found. The data was then used for the analysis of the odds ratio.

Comment:
7. In table 1, smoking is not associated with lung cancer. This suggests that selection forces in the recruitment of cases and controls are biasing the study results. It seems unusual that 61% of lung cancer cases were never smokers. Is this a typical feature of lung cancer in Thailand?

Response:
Yes, 61% of the cases never smoked is acceptable. It was reported that smoking prevalence of Thai males decreased from 60% to 39%, and from 5% to 2.1% in females between 1991 and 2014 [1]. While a survey in 2017 reported a smoking prevalence of 20.7% of the total adult population over 15 years old [2].

It was interesting to note that in this study, 49.2% of the cases were adenocarcinoma lung cancer which has a limited relation to cigarette smoking, whereas squamous cell, and small cell lung carcinoma are highly related to smoking [3, 4].


Comment:
8. The results reported in tables 2 and 3 seem to be internally inconsistent. For instance,
ORs for organophosphates depicted in table 4 indicate a strong association with lung cancer with days of use (Q4 vs. Q1 OR = 28.43 (95% CI = 11.11-72.76); an extremely large magnitude. However, the ORs for specific organophosphate insecticides are much more modest, although the statistically significant associations with chlorpyfos and dielrin. A similar pattern is evident for herbicide as well. This lack of internal consistency really points to the methodological limitations as a likely explanation for the observed association.

Response:
In this study, we actually collected data from 35 individual pesticides, but 17 of them were excluded due to small sample size (less than 5 in each cell). Therefore, the OR groups may be larger than the individual OR ones.

Comment:
9. Recall bias was discussed as a limitation, but nothing is mentioned about other threats to internal validity. For instance, the potential for selection bias to arise from the recruitment strategies. As mentioned above, the lack of an association with smoking seems to indicate something is awry. In addition, exposure misclassification is undoubtedly present and should be discussed in the Discussion along with the other potential limitations.
Response:
The problems of selection bias and exposure misclassification has been further discussed in the manuscript as suggested. The problem of lack of association with smoking has already been solved.

Comment:
10. References 20 and 35 are the same report.
Response:
The error has been corrected.

Competing Interests: No competing interests were disclosed.
I was pleased to review your paper that describes a case-control study in Thailand including 233 incident lung cancer cases and 458 controls focusing on exposures to pesticides. Please find enclosed my comments for your consideration. In the:

**Introduction**, first paragraph, I think you mean “Polycyclic aromatic hydrocarbons”? I would not call a paper from 1999 “Recent studies.....” because it’s >20 years old.

**Methods**, first paragraph, it’s a case-control study, not a case-controlled study; it would be good to include a few more details such as any time limit for having resided in the province?; from where did the TCB receive cases?; were the diagnosis confirmed by some diagnostic tool? Please clarify if the 299 were contacted and 229 accepted, or if only 229 were contacted and accepted. I would be surprised if the latter, and wonder why the other were not contacted. We also wish to know the “participation rate” among the control subjects. Neighbours are not a random sample it’s a convenience sample. If you mean that the interviewer randomly selected control subjects among all neighbours you need to explain how this was done, e.g. within a distance from the house or “snowball” technique. Why do you adjust for farming (yes/no)? Please explain your rational. It does not make sense to me.

**Questionnaire**, the English questionnaire does not indicate that the number of days of pesticide use is per year, so it seems strange that lifetime exposure is calculated by multiplying years with days. Please also clarify if “exposure” refers to “personally mix or apply pesticides” only, or if it also includes working in the fields? Provide more details regarding the data collection e.g. were the interviewers employed for the study full-time, or were they students?, were there any quality control measures implemented, e.g. double interviews of a proportion of subjects, were the interviewers interviewing both cases and controls?

**Results**, it is very strange that there is not difference between cases and controls regarding smoking, if you have an explanation for this please discuss it later.

**Discussion**, I don't think that “the association were closer for herbicides and insecticides”, possibly “stronger” or “more pronounced”, and I prefer “more days” rather than “higher days”.

Among the limitations I think there is more to information bias, e.g. it is commonly difficult to assess exposure to specific chemicals because people don't know the names or don't recognize exposure. I must admit that I get suspicious that there are no missing in the data and no category for “don't know” in the questionnaire. I would add potential selection bias to the discussion; although we don't really know the participation rate among controls or how neighbours were selected, they are generally not an ideal control population.

**Disclaimer:**

Where authors/reviewers are identified as personnel of the International Agency for Research on Cancer / World Health Organization, the authors/reviewers alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer / World Health Organization

**Is the work clearly and accurately presented and does it cite the current literature?**
Partly

**Is the study design appropriate and is the work technically sound?**
Partly

**Are sufficient details of methods and analysis provided to allow replication by others?**
Partly

**If applicable, is the statistical analysis and its interpretation appropriate?**
Not applicable

**Are all the source data underlying the results available to ensure full reproducibility?**
Yes

**Are the conclusions drawn adequately supported by the results?**
Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Occupational cancer epidemiology, lung cancer, case-control studies, cohort studies

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 02 Dec 2020

**Chudchawal Juntarawijit,** Naresuan University, Muang District, Phitsanulok, Thailand

**Comment: Introduction,** first paragraph, I think you mean “Polycyclic aromatic hydrocarbons”? I would not call a paper from 1999 “Recent studies.....” because it's >20 years old.

**Response:** The term was changed to polycyclic aromatic hydrocarbons.

**Comment: Methods,** first paragraph, it's a case-control study, not a case-controlled study; it would be good to include a few more details such as any time limit for having resided in the province?; from where did the TCB receive cases?; were the diagnosis confirmed by some diagnostic tool? Please clarify if the 299 were contacted and 229 accepted, or if only 229 were contacted and accepted. I would be surprised if the latter, and wonder why the other were not contacted. We also wish to know the “participation rate” among the control subjects. Neighbours are not a random sample it's a convenience sample. If you mean that the interviewer randomly selected control subjects among all neighbours you need to explain how this was done, e.g. within a distance from the house or “snowball” technique. Why do you adjust for farming (yes/no)? Please explain your rational. It does not make sense to me.
**Response:** The mistake was corrected and more information on residency, TCB, and diagnostic confirmation was added to the Methods. The information of the number of cases was clarified; and information on the participation rate was also provided.

For the question, why did we adjust for farming? Actually, we tried to adjust for occupations since it is very likely that they will be exposed to pesticides differently. At first, there were several types of occupations, but due to the small number of participants in each category, the groups were limited to “farmer” and “none-farmer”. These two groups tended to have different risks of exposure to environmental pesticides, due to the nature of their work and physical health.

**Comment:**
**Questionnaire,** the English questionnaire does not indicate that the number of days of pesticide use is per year, so it seems strange that lifetime exposure is calculated by multiplying years with days. Please also clarify if “exposure” refers to “personally mix or apply pesticides” only, or if it also includes working in the fields? Provide more details regarding the data collection e.g. were the interviewers employed for the study full-time, or were they students?; were there any quality control measures implemented, e.g. double interviews of a proportion of subjects, were the interviewers interviewing both cases and controls?

**Response:**
More information was added and the mistakes were corrected. In this study, “exposure” refers to “personally mixed and/or applied pesticides” only, not working in the field. More information of interviewers was added to the methods. There were no other quality control measures implemented.

**Comment:**
**Results,** it is very strange that there is not difference between cases and controls regarding smoking, if you have an explanation for this please discuss it later.

**Response:** Data on cigarette smoke was reanalyzed and the difference was observed using a new category.

**Comment:**
**Discussion,** I don't think that “the association was closer for herbicides and insecticides”, possibly “stronger” or “more pronounced”, and I prefer “more days” rather than “higher days”.

**Response:** The term “closer” was changed to “stronger”

**Comment:** Among the limitations, I think there is more to information bias, e.g. it is commonly difficult to assess exposure to specific chemicals because people don't know the names or don't recognize exposure. I must admit that I get suspicious that there are no missing in the data and no category for “don't know” in the questionnaire. I would add potential selection bias to the discussion; although we don't really know the participation rate among controls or how neighbours were selected, they are generally not an ideal control population.
Response:
Yes, we agree that it was likely that some of participants could not recall or know the name of the pesticides used. If this type of bias occurs it would be equal between both the case and control groups, and minimize the association between exposure to pesticides and lung cancer. More information regarding bias was added to the Discussion section, and more information about the control group was added to the Methods. Those who could not recall or “don’t know” the name of the pesticides were categorized as “not used”.

Competing Interests: No competing interests were disclosed.