Annals of Epidemiology 27 (2017) 281–289

Contents lists available at ScienceDirect

Annals of Epidemiology

journal homepage: www.annalsofepidemiology.org

Review

# 2,4-dichlorophenoxyacetic acid (2,4-D) and risk of non-Hodgkin lymphoma: a meta-analysis accounting for exposure levels

Adam M. Smith MPH [a], Martyn T. Smith PhD [a], Michele A. La Merrill PhD [b], Jane Liaw MPH [c], Craig Steinmaus MD [c,*]

[a] School of Public Health, University of California at Berkeley, Berkeley
[b] Department of Environmental Toxicology, Comprehensive Cancer Center, University of California at Davis, Davis
[c] Arsenic Health Effects Research Program, UC Berkeley School of Public Health, Berkeley, CA

ARTICLE INFO

*Article history:*
Received 9 July 2016
Accepted 19 March 2017
Available online 31 March 2017

*Keywords:*
2,4-D
Lymphoma
Non-Hodgkin's
Meta-analysis
Occupational

ABSTRACT

2,4-Dichlorophenoxyacetic acid (2,4-D) is one of the most commonly used selective herbicides in the world. A number of epidemiology studies have found an association between 2,4-D exposure and non-Hodgkin lymphoma (NHL) but these results are inconsistent and controversial. A previous meta-analysis found no clear association overall but did not specifically examine high-exposure groups. We conducted a systematic review and meta-analysis of the peer-reviewed epidemiologic studies of the associations between 2,4-D and NHL, with a particular focus on high-exposure groups, and evaluations of heterogeneity, dose-response, and bias. A total of 12 observational studies, 11 case-control studies, and one cohort study, were included. The summary relative risk for NHL using study results comparing subjects who were ever versus never exposed to 2,4-D was 1.38 (95% confidence interval (CI), 1.07–1.77). However, in analyses focusing on results from highly exposed groups, the summary relative risk for NHL was 1.73 (95% CI, 1.10–2.72). No clear bias based on study design, exposure assessment methodology, or outcome misclassification was seen. Overall, these findings provide new evidence for an association between NHL and exposure to the herbicide 2,4-D.



2,4-D was the first successful selective herbicide ever developed and its commercial release came in 1946 [1]. As an auxinic herbicide, its main mechanism of herbicidal action is the induction of cell growth and division in broad–leaf weeds to the point of abnormal growth that results in weed death [2]. It increased yields for various cereal crops such as wheat, maize, and rice, and its low manufacturing costs led to its continued use and global dissemination [3–5]. It is currently listed as an active ingredient in hundreds of commercially available products [6].

Non-Hodgkin lymphomas (NHLs) are a heterogeneous group of cancers arising from lymphocytes and their precursors in the immune system. In 2012, there were an estimated 217,643 new cases, and it was the eighth most commonly diagnosed cancer worldwide [7]. Descriptive epidemiology reveals that NHL incidence and mortality have been rising in the developed world faster than the vast majority of other cancer types for the last several decades [8]. NHL is usually more common after age 60 and in men. The causes of most cases are unknown but risk factors may include certain chemicals such as benzene or chemotherapy agents, immune deficiencies caused by immunosuppressant drugs, infection with human immunodeficiency virus, various autoimmune diseases such as Sjogrens syndrome and ionizing radiation [9].

In its most recent 2007 review, the US Environmental Protection Agency concluded that there was a lack of sufficient evidence to establish a link between 2,4-D exposure, and cancer [10]. In 2015, the World Health Organization's International Agency for Research on Cancer (IARC) confirmed its 1987 classification of 2,4-D as a group 2B, possible human carcinogen, after concluding that there was not sufficient human epidemiological evidence to list 2,4-D as a group 1 carcinogen in humans, although a substantial minority considered that the evidence was limited [11,12]. While IARC identified a number of human epidemiologic studies showing an elevated risk of NHL in groups exposed to 2,4-D, a number of other studies showed no association. A recent meta-analysis on 2,4-D exposure and NHL reported finding no association but focused their analyses on results comparing groups with any exposure to 2,4-D to groups with no known exposure to this agent [13].

* Corresponding author. Arsenic Health Effects Research Program, UC Berkeley School of Public Health, 2470 Telegraph Ave., Suite 301, Berkeley, CA 94704. Tel.: (510) 990-8354.
*E-mail address:* craigs@berkeley.edu (C. Steinmaus).

http://dx.doi.org/10.1016/j.annepidem.2017.03.003
1047-2797/

However, the potential risks in those groups who most likely had the highest exposures were not specifically evaluated. This is important since, if a true association does exist, higher exposures will usually result in higher relative risks (RRs), and higher RRs are generally less likely to be due to chance, bias, or confounding. As described in the first of the Bradford–Hill causal inference considerations and elsewhere, smaller increases in RR generally have lower statistical power (all else being equal) and can more likely be solely caused by minor bias and confounding than larger increases in RR [14–17]. Because the primary purpose of this work was to evaluate whether the current literature supports an association between 2,4-D and NHL, rather than attempt to define specific dose-response relationships, our focus here was on evaluating those groups with the highest 2,4-D exposure.

## Methods

PubMed, Scopus, and TOXLINE databases were searched for peer-reviewed observational epidemiology studies that evaluated exposure to 2,4-D and NHL. Key search words included "2,4-dichlorophenoxyacetic acid," "2,4-D," "herbicides," or "pesticides," and "neoplasms," "cancer," "carcinogenesis," "lymphoma," "non-Hodgkin lymphoma," or "NHL." Review articles and the bibliographies of all included articles were also searched for relevant studies. The results of our literature search are detailed in Supplement Figure 1 and include articles indexed up to April 25th, 2016.

This meta-analysis includes only case-control and cohort studies that provided RR estimates for exposure to 2,4-D specifically, and only data published in peer-reviewed scientific journals. Ecologic studies were excluded [18]. Studies that only reported RRs of NHL based on broader exposure categories such as herbicides or phenoxyacetic acids, or only reported RRs for combinations of chemicals or only by job type (e.g., farmers) were also excluded. Cross-sectional studies that measured 2,4-D exposure after cancer diagnosis were also excluded [19]. Studies that calculated RRs for either NHL incidence and mortality were included in the meta-analysis.

For each selected study, the following information was abstracted: authors, year of publication, year(s) of the study, study design (case-control vs. cohort), location of the study, number and sources of NHL cases and non-cases, age ranges, gender distribution, exposure metrics used, and how exposures were assessed, participation rates, confounding variables assessed, exposure levels when available, and the RR estimates for each exposure metric, and exposure level assessed with their corresponding confidence intervals. Most studies we identified gave RRs for several different metrics of 2,4-D exposure, including cumulative exposure, exposure intensity, duration of exposure, exposure with and without personal protective equipment (PPE), and time since first exposure. When this occurred, we attempted to identify the metric with the most comprehensive evaluation of exposure and the metric most likely to capture the highest exposure group. As such, when RRs were given for multiple exposure metrics, we selected a single RR for the metric in the following order: cumulative exposure (intensity $\times$ duration), exposure intensity (usually expressed as the number of days used, mixed, or applied 2,4-D per year), the longest duration (expressed as number of years using, mixing, or applying 2,4-D per year), and ever exposed to 2,4-D versus never exposed to 2,4-D. This order was determined *a priori*. If separate RRs were provided for subjects using or not using PPE, the results for subjects not using PPE were selected. Several studies provided separate RRs for different levels of exposure (e.g., low, medium, and high cumulative exposure). When this occurred, because our focus was on groups with the highest exposure, we chose the RR for the highest exposure level. In some instances, multiple publications reported results from the same population or cohort. When this occurred, one publication was selected based on the following criteria in the following order: the publication that presented results for the most likely highest exposure group (i.e., highest level of cumulative exposure, exposure intensity, or longest duration); the publication with the largest number of NHL cases; and the most recent publication. Summaries of our selection criteria are provided in Supplemental Tables 1 and 2. All age groups were included in this analysis, although the majority of studies included looked exclusively at adult populations.

Two measures of association were reported, odds ratios (ORs) for the case-control studies and a standardized incidence ratio for the cohort study. Since NHL is a relatively rare disease, odds ratios and standardized incidence ratios were considered equivalent for this meta-analysis [20]. Summary RRs were calculated using the fixed effects inverse variance weighting method [21] and the random effects method [22]. All results listed in this publication are for the random effects model, unless otherwise specified. Heterogeneity among studies was assessed using the general variance–based method developed by Petitti [23]. Heterogeneity was also quantified using the $I^2$ values presented by Higgins et al. and was calculated using equation $I^2 = 100\% \times (X^2 - df)/X^2$, where df is the degrees of freedom (number of studies minus one) and $X^2$ is the chi-square heterogeneity statistic. The $I^2$ value describes the percentage of total variation across studies due to between study heterogeneity rather than chance. An $I^2$ percentage of 75 or higher was considered to be high heterogeneity and not attributable to chance [24].

Publication bias was assessed using funnel plots and Begg's and Egger's tests. Egger's test assesses the asymmetry in the funnel plot with a simple linear regression comparing each study's precision to its effect size divided by the standard error. The funnel plot is a graphical presentation of each study's effect size compared with an estimate of its precision and can exhibit asymmetry if publication bias is present [25]. Begg's test uses Kendall's rank order test to assess the correlation between the studies effect size's and their precision [26]. We assessed the quality of each study included in the quantitative analysis using the Newcastle–Ottawa assessment scales for cohort and case-control studies.

The impact of each study on summary RRs was evaluated by recalculating summary RRs after removing each study one at a time. As aforementioned, our focus was on assessing RRs in the highest exposure groups. To evaluate how this focus might have impacted our results, we performed a separate meta-analysis in which the RR for broadest exposure group or metric was selected. In general, this involved RRs for groups who were ever exposed to 2,4,-D (despite their exposure level) to groups who were never known to be exposed to 2,4-D.

## Results

A total of 12 observational studies meeting our inclusion criteria were identified. This included 11 case-control studies [27–37], two of which were nested in cohorts [31,34], and one cohort study [38]. Table 1 shows a summary of the data extracted from each study, including author and publication year, location, study design, number of exposed cases, demographic characteristics of each study population, the exposure groups and categories, exposure metric used, as well as outcome assessed and sources of outcome data. Six of the 12 studies included were done in the United States [27,30,34,36–38], including the only cohort study [38] and one of the nested case-control studies [34]. Two studies were conducted in Europe [28,33], two in Sweden [29,35], one in Canada [32], and one was conducted using an international cohort from 10 countries [31]. Eleven of the studies reported RRs for NHL incidence [27–30,32–38], and one study included both incident and

**Table 1**
Studies used in the meta-analysis of 2,4-D and non-Hodgkin lymphoma

| Author (year) | Location | Study design | Cases [+] | Sex | Age | Exposure group | Exposure category[†] | RR (95% CI) | Outcome[*] | Source | Exposure assessment |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Burns *et al.* [38] 2011 | Michigan | Cohort | 3 | Men | Unk | Production | Cumulative exposure years ≥ 5 unit-years | 2.16 (0.45–6.31) | Incident NHL | Cancer registry | Workplace history on jobs, and IH measurements at production facility |
| Cantor *et al.* [27] 1992 | Iowa/ Minnesota | CC | 89 | Men | ≥30 | Farmers | Ever handled, no PPE | 1.2 (0.9–1.7) | Incident NHL[*] | Cancer registry, hospitals | Self-reported information on occupations and exposure to specific pesticides |
| Cocco *et al.* [28] 2013 | Europe | CC | 2 | Both | ≥17 | Various occupations | Cumulative exposure | 0.6 (0.1–3.5) | Incident B cell lymphoma[*] | Unk | Self-reported information on occupations and pesticides used, and expert evaluation |
| Hardell *et al.* [29] 1994 | Umea | CC | 3 | Men | 25–85 | Various Occupations | ≥1 week continuously or 1 month total exposed | 13 (1.2–360) | Incident NHL[*] | Hospital | Self-reported information on work history and exposures |
| Hoar *et al.* [30] 1986 | Kansas | CC | 5 | Men | ≥21 | Farmers | ≥21 days/year | 7.6 (1.8–32.3) | Incident NHL[*] | Cancer registry | Self-reported information on occupations and pesticides used |
| Kogevinas *et al.* [31] 1994 | Global | NCC | 2 | Both | Unk | Production or Spraying | Cumulative exposure score ≥ 10 | 0.69 (0.11–4.55) | NHL mortality and incidence | Cancer registry, death registry | Company exposure questionnaires and company records and expert evaluation |
| McDuffie *et al* .[32] 2001 | Canada | CC | 11 | Men | ≥19 | Any exposure | >7 days/year | 1.22 (0.60–2.49) | Incident NHL[*] | Cancer registry, Quebec hospitals | Self-reported information on occupations and pesticides used |
| Miligi *et al.* [33] 2006 | Italy | CC | 9 | Both | 20–74 | Various occupations | Probability > low, no PPE | 4.4 (1.1–29.1) | Incident NHL | Hospitals, varese Cancer registry | Self-reported information on agricultural work, crops, and pesticide use, and expert evaluation |
| Mills *et al.* [34] 2005 | California | NCC | 60[++] | Both | Unk | Farmers | "High" use | 3.80 (1.85–7.81) | Incident NHL | Cancer registry | Employee records and pesticide usage database |
| Nordstrom *et al.* [35] 1998 | Sweden | CC | 2 | Men | 28–86 | Any exposure | ≥8 hours | 1.6 (0.3–8.3) | Incident HCL | Cancer registry | Self-reported information on occupations and pesticides used at home or work |
| Woods [36] 1989 | Washington | CC | 576[++] | Men | 20–79 | Farmers | ≥1 or 2 days/year | 0.73 (0.4–1.3) | Incident NHL | Cancer registry | Self-reported information on occupations and specific chemicals used |
| Zahm *et al.*[37] 1990 | Nebraska | CC | 3 | Men | ≥21 | Farmers | ≥21 days/year | 3.3 (0.5–22.1) | Incident NHL[*] | Hospitals, study group | Self-reported information on agricultural practices and pesticides used |

CC = case-control; CI = confidence interval; HCL = hairy cell leukemia; IH = industrial hygiene; NCC = nested case control; NHL = non-Hodgkin's lymphoma; PPE = personal protective equipment; RR = relative risk; Unk = known or not stated.
+ Exposed cases.
++ Total number of cases.
* Studies with histologically confirmed cases.
† Category used in main analysis.

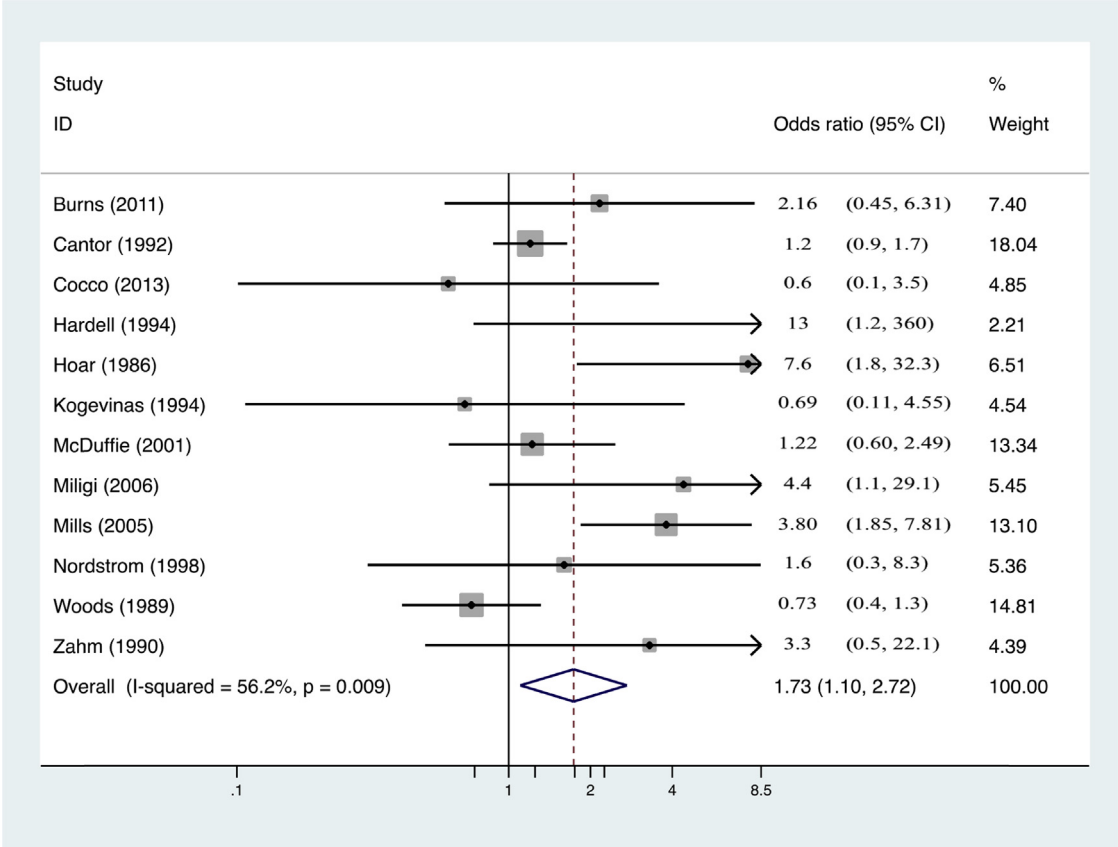



**Fig. 1.** Forest plot of the random effects results from the main analysis involving higher exposure groups.

mortality cases [31]. Eight of the studies exclusively looked at male populations [27,29,30,32,35–38], whereas four studies included both genders [28,31,33,34]. Ten of the studies examined occupational exposures [27–31,33,34,36–38], whereas two studies included nonoccupational and leisure time exposures [32,35]. Farming occupations comprised the exposure group in five of the twelve studies [27,30,34,36,37], and two studies included 2,4 D production workers [31,38]. Nine of the twelve studies (75%) reported RR estimates >1.0 for 2,4-D and NHL [27,29,30,32–35,37,38]. In four of these (33%), the RR estimate was statistically significant (Fig. 1) [29,30,33,34]. Of the three studies with RR estimates less than 1.0, none of these results were statistically significant [28,31,36].

The summary RR estimate in our analysis selecting the highest exposure categories was 1.73 (95% CI 1.10–2.72; *P*-value = .02, *n* = 12 studies; Table 2, Fig. 1). The highest RR included was 13 and the lowest RR was 0.6. The $X^2$ value was 25.12 (*P*-value for heterogeneity = .009; $I^2$ = 56%). No individual study received more than 19% of the total weight assigned in the random effects model. In the meta-analysis where RRs were preferentially selected for the broadest exposure category from each study (e.g., ever vs. never exposed), the summary RR estimate was lower but still statistically significant (RR = 1.38; 95% CI 1.07–1.77; *P* = .01, *n* = 12 studies; $I^2$ = 50%; Table 2, Supplement Fig. 2).

To explore potential sources of heterogeneity, a variety of subgroup analyses were performed. In analyses confined to studies conducted in the United States, the summary RR estimate was 1.96 (95% CI 1.03–3.76, *n* = 6) and the $X^2$ value was 19.33 (*P*-value for heterogeneity = .002; $I^2$ = 74%). For studies that gave a RR estimate that assessed the relationship between 2,4-D and NHL in an occupational cohort specified to be farmers, the summary RR was 1.97 (95% CI 0.95–4.08, *n* = 5). The $X^2$ value was 18.87 (*P*-value for heterogeneity = .001; $I^2$ = 79%).

To evaluate possible exposure misclassification, we conducted subgroup analyses based on the method used to classify subject's exposure. For studies in which the exposure classification was based on the subject's self-reported exposure to 2,4-D, where subjects were asked to recall past exposures, the summary RR was 1.47 (95% CI 0.89–2.44; *P*-value = .14, *n* = 7) with a $X^2$ value of 12.83; *P*-value for heterogeneity = .05, $I^2$ = 53%. In studies in which exposure was based on industrial hygienists or agronomists assessment of likelihood of exposure based job titles, company records, questionnaire, and industrial hygiene monitoring, the summary RR was 2.17 (95% CI 1.03–4.58; *P*-value = .04, *n* = 5) with a $X^2$ value of 6.07; *P*-value for heterogeneity = .19, $I^2$ = 34%. To assess outcome misclassification, a subgroup analysis was conducted for studies that had an independent pathology review to confirm the diagnosis of NHL. Here the summary RR was 1.73 (95% CI 0.92–3.25; *P*-value = .09, *n* = 6) with a $X^2$ value of 10.10; *P*-value for heterogeneity = .07, $I^2$ = 51%. The Newcastle Ottawa scores (highest possible score = 9) ranged from 6 to 9, with a median score of 7 (Supplement Table 3). The summary RR for the four studies with scores ≥8 was 2.86 (95% CI 0.99–8.23; $I^2$ = 65.9%) [27,30,33,37]. For those eight studies with a quality score of ≤7 [28,29,31,32,34–36,38], the summary RR was 1.48 (95% CI 0.80–2.73; $I^2$ = 57.1%).

Publication bias was assessed using the funnel plot, and Egger's and Begg's tests. The funnel plot was relatively symmetrical indicating little to no obvious publication bias (Fig. 2). The Egger's test yielded a bias coefficient of 1.086 with a *P*-value of .14 indicating no

**Table 2**
Summary results of the meta-analysis on 2,4-D exposure and non-Hodgkin lymphoma

| Group | N | Fixed effects | | Random effects | | Heterogeneity | | |
|---|---|---|---|---|---|---|---|---|
| | | RR | 95% CI | RR | 95% CI | $X^2$ | *P*-value | $I^2$ |
| Exposure metric | | | | | | | | |
| Highest exposure [27–38] | 12 | 1.38 | 1.10–1.73 | 1.73 | 1.10–2.72 | 25.12 | .009 | 56% |
| Ever versus never [27–38] | 12 | 1.31 | 1.13–1.52 | 1.38 | 1.07–1.77 | 22.08 | .024 | 50% |
| Study design | | | | | | | | |
| Case-control [27–37] | 11 | 1.36 | 1.09–1.71 | 1.72 | 1.06–2.78 | 24.66 | .006 | 59% |
| Nested case-control [31,34] | 2 | 3.04 | 1.54–5.97 | 2.03 | 0.40–10.17 | 2.80 | .094 | 64% |
| Cohort [38] | 1 | 2.16 | 0.58–8.03 | — | — | — | — | — |
| Region | | | | | | | | |
| US[27,30,34,36–38] | 6 | 1.38 | 1.08–1.77 | 1.96 | 1.03–3.76 | 19.33 | .002 | 74% |
| Europe* [28,29,31,33,35] | 5 | 1.69 | 0.74–3.87 | 1.72 | 0.65–4.60 | 5.44 | .245 | 27% |
| Canada [32] | 1 | 1.22 | 0.60–2.47 | — | — | — | — | — |
| Exposure group/assessment | | | | | | | | |
| Occupational [27–31,33,34,36–38] | 10 | 1.40 | 1.10–1.78 | 1.91 | 1.09–3.34 | 24.96 | .003 | 64% |
| Farming cohort [27,30,34,36,37] | 5 | 1.36 | 1.05–1.74 | 1.97 | 0.95–4.08 | 18.87 | .001 | 79% |
| Production [31,38] | 2 | 1.48 | 0.51–4.32 | 1.48 | 0.51–4.32 | 0.96 | .326 | 0% |
| Self-reported [27,29,30,32,35–37] | 7 | 1.21 | 0.95–1.55 | 1.47 | 0.89–2.44 | 12.83 | .046 | 53% |
| Occupational records/hygienist assessed [28,31,33,34,38] | 5 | 2.57 | 1.50–4.41 | 2.17 | 1.03–4.58 | 6.07 | .194 | 34% |
| Outcome assessment | | | | | | | | |
| Incidence [27–30,32–38] | 11 | 1.40 | 1.11–1.75 | 1.82 | 1.14–2.92 | 24.57 | .006 | 59% |
| Mortality [31] | 1 | 0.69 | 0.11–4.44 | — | — | — | — | — |
| Population based [27,29,30,32,33,35–37] | 8 | 1.25 | 0.98–1.59 | 1.62 | 0.97–2.68 | 15.14 | .034 | 54% |
| Pathology review [27–30,32,37] | 6 | 1.32 | 1.00–1.73 | 1.73 | 0.92–3.25 | 10.10 | .073 | 51% |
| Other | | | | | | | | |
| Cantor *et al.* excluded [27] | 11 | 1.60 | 1.16–2.22 | 1.94 | 1.10–3.44 | 23.52 | .009 | 58% |
| Males only [27,29,30,32,35–38] | 8 | 1.24 | 0.97–1.58 | 1.50 | 0.94–2.39 | 13.55 | .06 | 48% |
| Newcastle–Ottawa scale | | | | | | | | |
| 8 or more [27,30,33,37] | 4 | 1.39 | 1.03–1.87 | 2.86† | 0.99–8.23 | 8.79 | .032 | 65.9% |
| 7 or less [28,29,31,32,34–36,38] | 8 | 1.37 | 0.97–1.94 | 1.48 | 0.80–2.73 | 16.32 | .022 | 57.1% |

CI = confidence interval; $I^2$ = Higgins heterogeneity statistic; N = number of studies; RR = risk ratio; $X^2$ = chi-squared value.
* Kogevinas et al. was included in the analysis despite having two cases from Australia.
† The difference between the fixed and random effects results was primarily due to the large study by Cantor et al. 1992 (OR = 1.2, 95% CI: 0.9–1.7) receiving 89.7% of the weight in the fixed effect model and 38.7% of the weight in the random effects model.

clear publication bias, and the Begg's test gave an adjusted Kendall score of 14 with a *P*-value of .34, also indicating no clear publication bias. The Cantor et al. study [27] was the most heavily weighted study (RR = 1.20, 95% CI, 0.9–1.7) receiving 18% of the weight in the random effects model, and its removal caused the summary RR to increase from 1.73 (95% CI 1.10–2.72) to 1.94 (95% CI 1.10–3.44; *P*-value = .02, $n$ = 11) with a $X^2$ value of 23.52; *P*-value for heterogeneity = .009, $I^2$ = 58%. Removal of the only cohort study [38] in the analysis left the summary RR almost unchanged at 1.72 (95% CI 1.06–2.78; *P*-value = .03, $n$ = 11) with a $X^2$ value of 24.66; *P*-value for heterogeneity = .006, $I^2$ = 59%.

The presence of statistically significant modest heterogeneity according to the $I^2$ value [24] led us to speculate how many studies could be removed from the meta-analysis while still showing a statistically significant summary RR for the association between 2,4-D exposure and NHL. To assess this, we sequentially removed studies based on the amount of weight they contributed to the final estimate, in a stepwise and additional fashion (i.e., those given the largest weight were removed first; Table 3). In this analysis, the six studies receiving the largest weight needed to be removed for the summary RR to no longer be statistically significant.

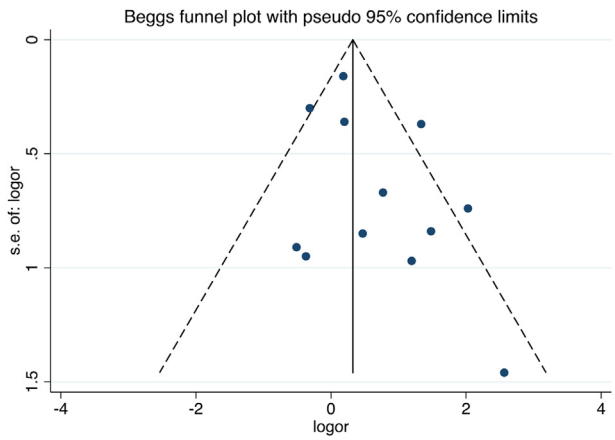



**Fig. 2.** Beggs funnel plot with pseudo 95% confidence intervals of the studies of 2,4 D and non-Hodgkin lymphoma. The x-axis is the natural log value of the relative risk included from each study and the y-axis is the standard errors of the natural log value for each relative risk.

## Discussion

Overall, in our analyses focused on the highest exposure group from each study, we identified a statistically significant association between 2,4-D exposure and increased RRs of NHL. Evidence of moderate heterogeneity was identified ($X^2$ = 25.12; *P*-heterogeneity = .009, $I^2$ = 56%). However, given the wide range in exposure scenarios, exposure assessment methods, outcome assessment methods, statistical methods, study populations, and other factors across studies at least some heterogeneity was expected. In addition, despite the presence of some heterogeneity, 75% of the studies included in our high-exposure analysis reported RR estimates >1.0. In addition, while four studies reported statistically significant positive associations, none of the studies included in our main meta-analysis reported statistically significant negative associations. Finally, the 95% confidence intervals of the large majority of individual study results included our summary RR. Overall, these findings suggest that despite the wide differences in many aspects of

**Table 3**
Sensitivity analysis: removing one study at a time from the main meta-analysis beginning with the studies receiving the greatest weight

| Removed | N | Fixed effects | | | Random effects | | | Heterogeneity | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | RR | 95% CI | *P*-value | RR | 95% CI | *P*-value | $X^2$ | *P*-value | $I^2$ |
| Cantor [27] (1992) | 11 | 1.60 | 1.16–2.22 | .004 | 1.94 | 1.10–3.44 | .023 | 23.52 | .009 | 58% |
| Woods [36] (1989) | 10 | 2.25 | 1.53–3.32 | <.001 | 2.31 | 1.36–3.93 | .002 | 13.66 | .135 | 34% |
| McDuffie [32] (2001) | 9 | 2.93 | 1.85–4.65 | <.001 | 2.80 | 1.64–4.79 | <.001 | 9.51 | .301 | 16% |
| Mills [34] (2005) | 8 | 2.46 | 1.35–4.47 | .003 | 2.44 | 1.24–4.80 | .01 | 8.68 | .276 | 19% |
| Burns [38] (2011) | 7 | 2.54 | 1.30–4.98 | .007 | 2.50 | 1.10–5.67 | .028 | 8.64 | .195 | 31% |
| Hoar [30] (1986) | 6 | 1.88 | 0.88–4.02 | .104 | 1.90 | 0.83–4.34 | .129 | 5.84 | .322 | 14% |

CI = confidence interval; N = number of studies included; RR = relative risk.

study design across the different studies, the overall trend in study results was consistent with the positive summary RR we identified.

The findings of our meta-analysis differed from those of a previous meta-analysis on this same topic [13], which did not identify a clear association (summary RR = 0.97, 95% CI, 0.77–1.22). The differences between our meta-analysis and this previous meta-analysis are detailed in Table 4. The primary reason for the different findings was our focus on selecting RR estimates for those groups most likely to be highly exposed to 2,4-D from each study. This is in contrast to the previous meta-analysis, which generally selected RRs for groups with any 2,4-D exposure, regardless of whether these exposures were high or low. For example, the previous meta-analysis used the RR from the analysis of De Roos et al. [39] which pooled data from three of the studies used in our meta-analysis [27,30,37]. We used data from the original three studies since each provided a RR estimate for a high exposure group. This is in contrast to the RR from the De Roos et al. study used in the previous meta-analysis which was simply for "exposed" compared with "not exposed" groups. As mentioned previously, the reason we focused on higher exposure groups is that if true associations are present, higher exposures generally result in higher RRs, and all else being equal, higher RRs are typically associated with greater statistical power and are less likely to be solely caused by major bias or confounding than RRs closer to 1.0 [14]. This is in contrast to results for groups with any 2,4-D exposure, regardless of whether the exposure was high or low. When subjects with any 2,4-D exposure are considered as the exposed group, subjects with fairly low exposures can be included in the "exposed" group, and the overall average exposure in the group is likely to be lower than that in groups solely composed of highly exposed workers. Because the average exposures in these "any exposure" groups are lower, RR increases in these groups are also likely to be low (i.e., closer to 1.0) and thus more susceptible to issues relating to low power, bias, and confounding. Some evidence for this can be seen in our results where the summary RRs in our analyses of high-exposure groups resulted in an elevated and statistically significant summary RR of 1.73 (95% CI, 1.10–2.72), whereas our analyses of "ever-exposed" groups resulted in a summary RR that was lower (RR = 1.38, 95% CI, 1.07–1.77).

The metrics we used to define higher exposure groups varied across studies (e.g., cumulative exposure vs. duration vs. no PPE use). Although this could potentially introduce some heterogeneity, it would not necessarily cause false positive associations. The reason for this is that some of the metrics we used are likely better for identifying those who truly are highly exposed than others. Because all the studies included in our meta-analysis assessed exposure similarly in NHL cases and controls, most exposure misclassification was likely non-differential and using studies with these less accurate or less thorough assessments of exposure most likely biased results toward finding no association. In fact, all the studies included in our meta-analysis likely had at least some degree of non-differential exposure misclassification, which again, most likely biased results toward the null. Overall, because of these issues, our meta-analysis results may underestimate the true effect size of the association between NHL and 2,4-D exposure.

Differential recall bias is a potential concern in cancer case-control studies because cancer cases may recall past exposures with greater or less accuracy than controls. However, there are several factors suggesting that this bias does not explain the results presented here. If this type of recall bias had played a role, then RR estimates derived from cohort studies [38], case-control studies that were nested within a cohort [31,34], or case-control studies where exposure was assessed by an agronomist or hygienist [28,33] (where recall bias is less likely to have had a major impact), should have been substantially lower than those found in our analyses of all studies combined. However, this was not the case: after removing case-control studies that assessed exposures based solely on subject's self-reports, where subjects were asked to recall past exposures, the RR estimate increased (from 1.47 to 2.17).

Another difference between the previous meta-analysis and ours was that the previous meta-analysis included the case-control study by Hartge et al. [19], which found no association between 2,4-D concentrations in home carpet dust and NHL (RR = 0.89, 95% CI, 0.49–1.59). However, these 2,4-D concentrations were measured after cancer diagnosis. Because the half-life of 2,4-D is relatively short (6.2 days) [41] and the latency between first exposure and cancer diagnosis is many years for most known environmental carcinogens, and because people's behaviors can change following cancer diagnosis, we did not include these types of cross-sectional studies in our meta-analysis. As previously mentioned, the objective of this study was to examine the risk of NHL in high exposure groups. Because of this, we used the results presented in McDuffie et al. [32] rather than those from Hohenadel et al. [40], which assessed the same study subjects and was used in the previous meta-analysis. Despite being published earlier, the McDuffie et al. [32] study provides RR estimates for four different 2,4-D exposure levels (from >0 to ≤2 days 2,4-D use/year to >7 days 2,4-D use/year). In contrast, the Hohenadel et al. [40] study only provides a RR estimate for exposed versus unexposed, and thus combines higher and lower exposure groups. Other differences are that we included the Nordstrom et al. study of hairy cell leukemia, a type of low grade NHL [35], and the Cocco et al. study of B cell lymphoma, the most common type of NHL [28]. Excluding these two studies had only a small impact on our summary RR (RR = 1.88; 95% CI, 1.14–3.09, *P*-value = .01), indicating their inclusion is not the reason for the statistically significant results we report.

An assessment of outcome misclassification was done by conducting a subgroup analysis on studies where NHL cases had been independently reviewed by pathologist(s) to confirm the diagnosis [27–30,32,37]. The magnitude of the summary RR in this subgroup analysis (RR = 1.73; 95% CI, 0.92–3.25) was the same as that in our analysis including all studies suggesting that outcome misclassification was minimal. In addition, it appears that in all studies included here, case status was determined independent of 2,4-D exposure. As such, most errors in assessing outcome status were

**Table 4**
Comparison of meta-analyses for 2,4-D and NHL

| Author (year) | Main analysis | | | | Ever versus never | | | | Goodman *et al.* [13] | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Used | RR (95% CI) | Weight | Exposure group | Used | RR (95% CI) | Weight | Exposure group | Used | RR (95% CI) | Weight | Exposure group |
| Burns *et al.* [38] (2011) | Yes | 2.16 (0.45−6.31) | 7.40% | Cumulative exposure >5 unit-years | Yes | 1.36 (0.74−2.29) | 10.06% | Never versus ever (all workers) | Yes | 1.36 (0.74−2.29) | 11.90% | Never versus ever (all production workers) |
| Cantor *et al.* [27] (1992) | Yes | 1.2 (0.9−1.7) | 18.04% | Ever handled without PPE | Yes | 1.2 (0.9−1.6) | 16.16% | Ever handled | No | — | — | Substituted DeRoos *et al.* [39] |
| Cocco *et al.* [28] (2013) | Yes | 0.6 (0.1−3.5) | 4.85% | Cumulative exposure | Yes | 0.6 (0.1−3.5) | 1.81% | Cumulative exposure | No | — | — | — |
| DeRoos *et al.* [39] (2003) | No | — | — | — | No | — | — | — | Yes | 0.8 (0.60−1.10) | 24.38% | Pooled estimate [27,30,37] |
| Hardell *et al.* [29] (1994) | Yes | 13 (1.2−360) | 2.21% | 1 week or more continuously or 1 month total exposed | Yes | 13 (1.2−360) | 0.74% | ≥1 week continuously or 1 month total exposed | Yes | 13 (1.2−360) | 0.64% | 1 week or more continuously or 1 month total exposed |
| Hartge [19] (2005) | No | — | — | — | No | — | — | — | Yes | 0.89 (0.49−1.59) | 11.21% | 2,4-D ≥ 1000 ng/g |
| Hoar *et al.* [30] (1986) | Yes | 7.6 (1.8−32.3) | 6.51% | Days/year ≥ 21 | Yes | 2.6 (1.4−5.0) | 9.04% | All 2,4-D users (no 2,4,5-T use) | No | — | — | Substituted DeRoos *et al.* [39] |
| Hohenadel *et al.* [40] (2011) | No | — | — | — | No | — | — | — | Yes | 0.94 (0.67−1.33) | 21.79% | 10 hours or more of lifetime use |
| Kogevinas *et al.* [31] (1994) | Yes | 0.69 (0.11−4.55) | 4.54% | Cumulative exposure ≥ 10 | Yes | 1.11 (0.46−2.65) | 5.83% | Cumulative exposure ≥ 0.005 | Yes | 1.11 (0.46−2.65) | 5.92% | Cumulative exposure ≥ 0.005 |
| McDuffie *et al.*[32] (2001) | Yes | 1.22 (0.60−2.49) | 13.34% | Days/year ≥ 7 | Yes | 1.26 (0.97−1.64) | 17.10% | Ever exposed (10 hours or more of lifetime use) | No | — | — | Substituted Hohenadel *et al.* [40] |
| Miligi *et al.*[33] (2006) | Yes | 4.4 (1.1−29.1) | 5.45% | > Low, no PPE | Yes | 0.9 (0.5−1.8) | 8.73% | > Low | Yes | 0.9 (0.5−1.8) | 9.87% | > Low |
| Mills *et al.*[34] (2005) | Yes | 3.80 (1.85−7.81) | 13.10% | Pounds applied (dichotomous) | Yes | 3.80 (1.85−7.81) | 7.60% | Pounds applied (dichotomous) | Yes | 3.58 (1.02−12.58) | 3.10% | Pounds applied (dichotomous, adjusted) |
| Nordstrom *et al.*[35] (1998) | Yes | 1.6 (0.3−8.3) | 5.36% | 8 hours or more of exposure | Yes | 1.6 (0.3−8.3) | 2.05% | ≥8 hours exposed | No | — | — | — |
| Woods [36] (1989) | Yes | 0.73 (0.4-1.3) | 14.81% | 1 or 2 days or more/ year | Yes | 0.73 (0.4−1.3) | 9.71% | 1 or 2 days or more/ year | Yes | 0.73 (0.4−1.3) | 11.19% | 1 or 2 days or more/ year |
| Zahm *et al.*[37] (1990) | Yes | 3.3 (0.5−22.1) | 4.39% | 21 days or more/year | Yes | 1.5 (0.9−2.5) | 11.19% | Ever handled | No | — | — | Substituted DeRoos *et al.*[39] |

2,4,5,-T = 2,4,5-trichlorophenoxyacetic acid; 2,4-D = 2,4-dichlorophenoxyacetic acid; CI = confidence interval; NHL = non-Hodgkin's lymphoma; PPE = personal protective equipment; RR = relative risk.

most likely non-differential and thus most likely caused bias toward the null. NHL is a heterogeneous mixture of cancers, however, there were too few studies that provided results on 2,4-D exposure and specific subtypes of NHL to obtain meaningful estimates from subgroup analyses. Importantly though, if 2,4-D exposure is more strongly associated with some subtypes than others, our inclusion of studies examining all subtypes together (and therefore possibly including subtypes less strongly associated with 2,4-D) would have biased our meta-analysis results to the null, not toward the positive association we identified. Overall, misclassification of NHL status appears to be an unlikely cause of the association we identified.

Monotonic dose-response relationships are argued to be supportive of a causal relationship between exposures and outcomes [14], and their assessment has been well outlined in meta-analyses [42]. Because one of the aims of this study was causal inference, we examined dose-response relationships by several methods. First, as discussed above, we performed one meta-analysis preferentially selecting results from higher exposure groups and a separate meta-analysis preferentially selecting subjects with any exposure to 2,4-D ("ever" exposed). The observation that the summary RR increased as the likely average exposure in each group increased (from 1.38 for ever exposed to 1.73 for likely higher exposed) is consistent with a positive dose-response pattern. Second, we performed a subgroup analysis that only included occupational exposures to 2,4-D under the assumption that typical occupational exposures may generally be higher than typical residential exposures [27–31,33,34,36–38]. The higher summary risk estimate we identified in studies that only considered occupational exposures (RR = 1.91, 95% CI, 1.09–3.34) is further evidence that exposure to 2,4-D is associated with NHL in a dose-dependent fashion. Dose-response meta-regressions are another method for assessing dose-response relationships, but this was not possible here because of the heterogeneity in the different methods each study used to assess and categorize 2,4-D exposure.

The association of 2,4-D with NHL is further supported by considerations of biological plausibility based on mechanistic data. According to the IARC working group for Monograph 113, "mechanistic studies provided strong evidence that 2,4-D induces oxidative stress that can operate in humans and moderate evidence that 2,4-D causes immunosuppression, based on *in vivo* and *in vitro* studies" [11]. Both of these key characteristics of human carcinogens [43] may contribute mechanistically to the development of NHL. For example, immunosuppression is strongly associated with the development of NHL [44,45]. Furthermore, increased risk of NHL has been associated with common genetic variants in the oxidative stress pathway, including NADPH oxidase, which plays an important role in signaling for the proliferation of lymphocytes and tumor cells [46].

Some studies in rats and mice have shown immunosuppressive effects of 2,4-D, but others have found little or no effect. A well-designed study by Salazar et al. [47] using pure 2,4-D at non-toxic doses reported that 2,4-D caused significant immunosuppressive effects. The number of phosphorylcholine IgM and IgG antibody-secreting B cells (plasma cells) in bone marrow cells of C57BL/6 mice exposed to 2,4-D was decreased by 2- to 3-fold, indicating substantial immunosuppressive effects on humoral immunity. Human studies to demonstrate that this mechanism can operate in humans are lacking, however, and would be important to perform.

We have conducted a meta-analysis on 2,4-D exposure and NHL that focuses on higher exposure groups. In our analyses, we examined dose-response relationships, heterogeneity, and the role of several forms of bias. Evidence of a dose-response relationship was seen when comparing the results for higher exposure groups to ever versus never exposure groups, and no major exposure or outcome misclassifications, or publication bias were detected. Overall, our review of the current epidemiologic literature suggests that 2,4-D exposure is associated with increased risks of NHL. Given the widespread use of this agent, these findings may have important public health implications.

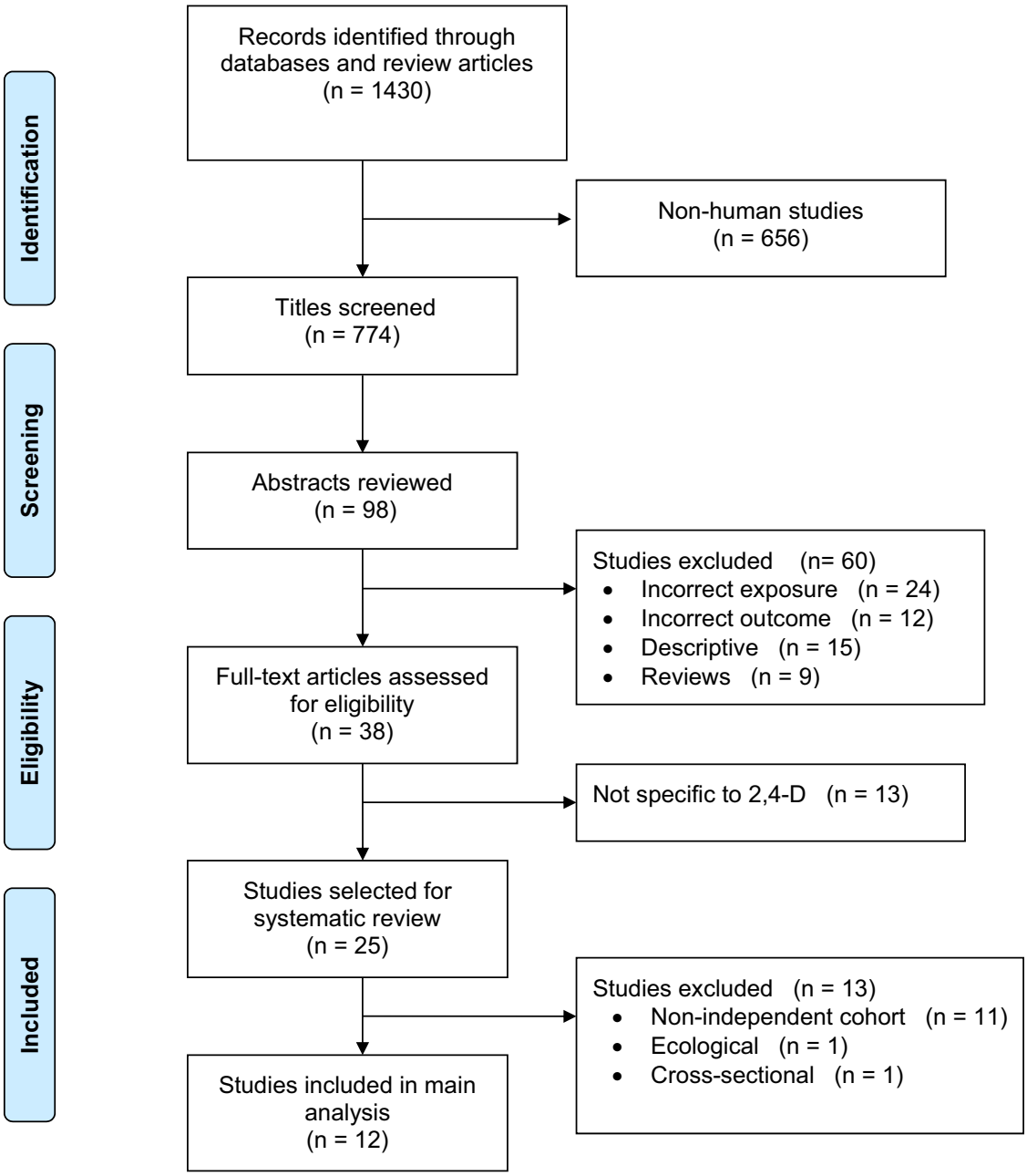
## Acknowledgments

This research was supported by NIFA grant CA-D-ETX-2233-H.

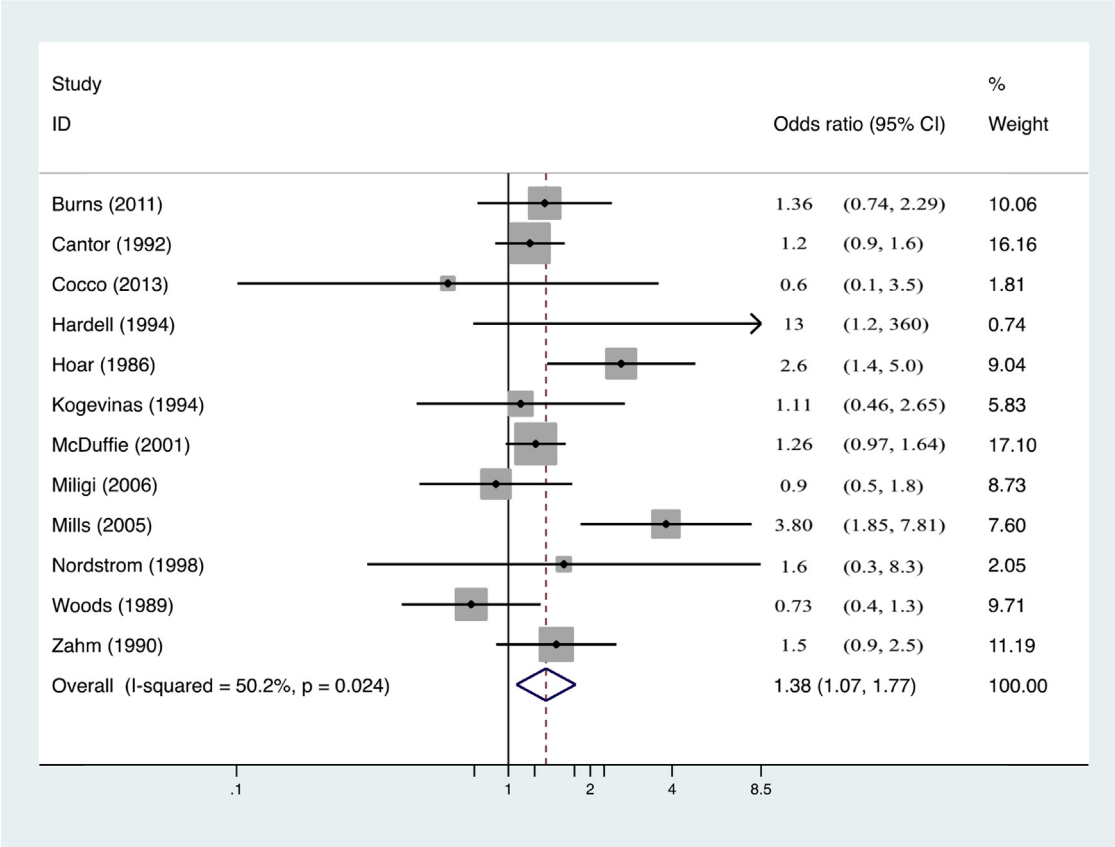
## References

[1] Quastel JH. 2,4-Dichlorophenoxyacetic Acid (2,4-D) as a selective herbicide. Agricultural Control Chemicals. Advances in Chemistry. 1. Washington, D.C.: American Chemical Society; 1950. p. 244–9.
[2] Song Y. Insight into the mode of action of 2,4-dichlorophenoxyacetic acid (2,4-D) as an herbicide. J Integr Plant Biol 2014;56(2):106–13.
[3] Final review report for the active substance 2,4-D. SANCO/11961/2015 Rev 3 final Ed. European Commission Directorate General for Health and Food Safety, 2016. Standing Committee on Plants, Animals, Food and Feed.
[4] Health Canada. Information note; Health Canada Releases Final Re-Evaluation Decision on 2,4-D, REV2013–02. 2008. Available at: http://www.hc-sc.gc.ca/cps-spc/alt_formats/pdf/pubs/pest/decisions/rev2013-02/rev2013-02-eng.pdf [accessed 03.09.2016].
[5] US EPA. Pesticides industry sales and usage: 2006 and 2007 Market Estimates. Office of Chemical Safety and Pollution Prevention, EPA 733-R-11-001. 2011. Available at: https://www.epa.gov/sites/production/files/2015-10/documents/market_estimates2007.pdf [accessed 04.25.2016].
[6] Common list of pesticide products that contain 2,4-D, Carbaryl, Diazinon, Diuron, Malathion, Triclopyr Bee or Trifluralin. Cincinnati, Ohio: National Service Center for Environmental Publications; 2010. Available at: http://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=2000U3UD.txt [accessed 04.18.2016].
[7] Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer; 2013. Available at: http://globocan.iarc.fr [accessed 03.03.2016].
[8] Devesa SS, Fears T. Non-Hodgkin's lymphoma time trends: United States and international data. Cancer Res 1992;52(19 Suppl):5432s–40s.
[9] Hartge P, Smith MT. Environmental and behavioral factors and the risk of non-Hodgkin lymphoma. Cancer Epidemiol Biomarkers Prev 2007;16(3):367–8.
[10] Gulliford JB. 2,4-D, 2,4-DP, and 2,4-DB; decision not to initiate special review. Fed Regist 2007;152:44510–1.
[11] Loomis D, Guyton K, Grosse Y, El Ghissasi F, Bouvard V, Benbrahim-Tallaa L, et al. Carcinogenicity of lindane, DDT, and 2,4-dichlorophenoxyacetic acid. Lancet Oncol 2015;16(8):891–2.
[12] Overall evaluations of carcinogenicity: an updating of IARC Monographs volumes 1 to 42. IARC Monogr Eval Carcinog Risks Hum Suppl 1987;7:1–440.
[13] Goodman JE, Loftus CT, Zu K. 2,4-Dichlorophenoxyacetic acid and non-Hodgkin's lymphoma, gastric cancer, and prostate cancer: meta-analyses of the published literature. Ann Epidemiol 2015;25(8):626–636.e4.
[14] Greenland S. In: Rothman K, Greenland S, editors. Modern epidemiology. 2nd ed. Philadelphia: Lippincott Raven; 1998. p. 7–28.
[15] Axelson O. Aspects on confounding in occupational health epidemiology. Scand J Work Environ Health 1978;4:85–9.
[16] Hill AB. The environment and disease: association or causation? Proc R Sci Med 1965;58(5):295–300.
[17] Schlesselman J. Assessing effects of confounding variables. Am J Epidemiol 1978;108:3–8.
[18] Mills PK. Correlation analysis of pesticide use data and cancer incidence rates in California counties. Arch Environ Health 1998;53(6):410–3.
[19] Hartge P, Colt JS, Severson RK, Cerhan JR, Cozen W, Camann D, et al. Residential herbicide use and risk of non-Hodgkin lymphoma. Cancer Epidemiol Biomarkers Prev 2005;14(4):934–7.
[20] Rodrigues L, Kirkwood BR. Case-control designs in the study of common diseases: updates on the demise of the rare disease assumption and the choice of sampling scheme for controls. Int J Epidemiol 1990;19(1):205–13.
[21] Greenland S. Meta-analysis. In: Rothman K, Greenland S, editors. Modern epidemiology. 2nd ed. Philadelphia: Lippincott Raven; 1998. p. 643–73.
[22] DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177–88.
[23] Petitti D. Meta-analysis. In: Rothman K, Greenland S, editors. Modern epidemiology. 2nd ed. Philadelphia: Lippincott Raven; 1998. p. 643–73.
[24] Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557–60.
[25] Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315(7109):629–34.
[26] Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994;50(4):1088–101.
[27] Cantor KP, Blair A, Everett G, Gibson R, Burmeister LF, Brown LM, et al. Pesticides and other agricultural risk factors for non-Hodgkin's lymphoma among men in Iowa and Minnesota. Cancer Res 1992;52(9):2447–55.

[28] Cocco P, Satta G, Dubois S, Pili C, Pilleri M, Zucca M, et al. Lymphoma risk and occupational exposure to pesticides: results of the Epilymph study. Occup Environ Med 2013;70(2):91–8.
[29] Hardell L, Eriksson M, Degerman A. Exposure to phenoxyacetic acids, chlorophenols, or organic solvents in relation to histopathology, stage, and anatomical localization of non-Hodgkin's lymphoma. Cancer Res 1994;54(9):2386–9.
[30] Hoar SK, Blair A, Holmes FF, Boysen CD, Robel RJ, Hoover R, et al. Agricultural herbicide use and risk of lymphoma and soft-tissue sarcoma. JAMA 1986;256(9):1141–7.
[31] Kogevinas M, Kauppinen T, Winkelmann R, Becher H, Bertazzi PA, Bueno-de-Mesquita HB, et al. Soft tissue sarcoma and non-Hodgkin's lymphoma in workers exposed to phenoxy herbicides, chlorophenols, and dioxins: two nested case-control studies. Epidemiology 1995;6(4):396–402.
[32] McDuffie HH, Pahwa P, McLaughlin JR, Spinelli JJ, Fincham S, Dosman JA, et al. Non-Hodgkin's lymphoma and specific pesticide exposures in men: cross-Canada study of pesticides and health. Cancer Epidemiol Biomarkers Prev 2001;10(11):1155–63.
[33] Miligi L, Costantini AS, Veraldi A, Benvenuti A, Vineis P. WILL. Cancer and pesticides: an overview and some results of the Italian multicenter case-control study on hematolymphopoietic malignancies. Ann N Y Acad Sci 2006;1076:366–77.
[34] Mills PK, Yang R, Riordan D. Lymphohematopoietic cancers in the United Farm Workers of America (UFW), 1988-2001. Cancer Causes Control 2005;16(7):823–30.
[35] Nordstrom M, Hardell L, Magnuson A, Hagberg H, Rask-Andersen A. Occupational exposures, animal exposure and smoking as risk factors for hairy cell leukaemia evaluated in a case-control study. Br J Cancer 1998;77(11):2048–52.
[36] Woods JS, Polissar L. Non-Hodgkin's lymphoma among herbicide-exposed farm workers in western Washington state. Chemosphere 1989;18(1-6):401–6.
[37] Zahm SH, Weisenburger DD, Babbitt PA, Saal RC, Vaught JB, Cantor KP, et al. A case-control study of non-Hodgkin's lymphoma and the herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) in eastern Nebraska. Epidemiology 1990;1(5):349–56.
[38] Burns C, Bodner K, Swaen G, Collins J, Beard K, Lee M. Cancer incidence of 2,4-D production workers. Int J Environ Res Public Health 2011;8(9):3579–90.
[39] De Roos AJ, Zahm SH, Cantor KP, Weisenburger DD, Holmes FF, Burmeister LF, et al. Integrative assessment of multiple pesticides as risk factors for non-Hodgkin's lymphoma among men. Occup Environ Med 2003;60(9):E11.
[40] Hohenadel K, Harris SA, McLaughlin JR, Spinelli JJ, Pahwa P, Dosman JA, et al. Exposure to multiple pesticides and risk of non-Hodgkin lymphoma in men from six Canadian provinces. Int J Environ Res Public Health 2011;8(6):2320–30.
[41] US EPA. Reregistration Eligibility Decision for 2,4-D. Office of Prevention, Pesticides, and Toxic Substances, EPA 738-R-05-002. 2005. p. 25. Available at: https://archive.epa.gov/pesticides/reregistration/web/pdf/24d_red.pdf [accessed 01.05.2017].
[42] Berlin JA, Longnecker MP, Greenland S. Meta-analysis of epidemiologic dose-response data. Epidemiology 1993;4(3):218–28.
[43] Smith MT, Guyton KZ, Gibbons CF, Fritz JM, Portier CJ, Rusyn I, et al. Key characteristics of carcinogens as a basis for organizing data on mechanisms of carcinogenesis. Environ Health Perspect 2016;126(6):713–21.
[44] Tran H, Nourse J, Hall S, Green M, Griffiths L, Gandhi MK. Immunodeficiency-associated lymphomas. Blood Rev 2008;22(5):261–81.
[45] Na R, Laaksonen MA, Grulich AE, Meagher NS, McCaughan GW, Keogh AM, et al. Iatrogenic immunosuppression and risk of non-Hodgkin lymphoma in solid organ transplantation: a population-based cohort study in Australia. Br J Haematol 2016;174:550–62.
[46] Lan Q, Zheng T, Shen M, Zhang Y, Wang SS, Zahm SH, et al. Genetic polymorphisms in the oxidative stress pathway and susceptibility to non-Hodgkin lymphoma. Hum Genet 2007;121(2):161–8.
[47] Salazar KD, de la Rosa P, Barnett JB, Schafer R. The polysaccharide antibody response after Streptococcus pneumoniae vaccination is differentially enhanced or suppressed by 3,4-dichloropropionanilide and 2,4-dichlorophenoxyacetic acid. Toxicol Sci 2005;87(1):123–33.




**Supplement Fig. 1.** PRISMA diagram of literature search results for databases and review articles indexed up to April 25th, 2016.




**Supplement Fig. 2.** Forest plot of the random effects results from the "ever versus never" analysis.

**Supplement Table 1**
Exclusion and inclusion criteria for studies used in quantitative analysis

| Inclusion criteria | Exclusion criteria |
|---|---|
| • Epidemiologic study in a human population on 2,4 D exposure and NHL<br>• Case-control and cohort study design<br>• Studies of NHL incidence or mortality<br>• Assessed occupational or residential 2,4 D use or exposure<br>• Published in a peer-reviewed journal<br>• Provided a relative risk estimate and variance estimate for NHL incidence or the data to estimate them<br>• If relative risks are provided for multiple exposure levels (e.g., high, medium, and low) the relative risk for the highest exposure group was used in the main analysis<br>• For studies with overlapping populations, one result was selected based on the factors in Supplement Table 2<br>• Adjusted, matched, or otherwise controlled for age and gender, or no evidence of major differences between age and gender between the 2,4 D exposed and unexposed groups. | • Cross-sectional or ecologic study design<br>• Did not report relative risk or variance estimates or the data to calculate them<br>• Published only in a government or industry report or as an abstract<br>• Studies that only reported results for the use of multiple pesticides combined<br>• Studies only providing results for broad occupational categories (e.g., pesticide applicators). |

2,4 D = 2,4-dichlorophenoxyacetic acid; NHL = non-Hodgkin lymphoma.

**Supplement Table 2**
Relative risk selection criteria

| Main analysis: high exposures* | Ever versus never exposed analysis* |
|---|---|
| 1. Cumulative exposure<br>2. Exposure intensity<br>3. Duration of exposure<br>4. Any exposure<br>5. If two publications from the same cohort both gave the same exposure metric, the metric calculated with the most cases or the most recent study was included in the analysis. | 1. Any exposure<br>2. If two publications from the same cohort both gave the same exposure metric, the metric calculated with the most cases or the most recent study was included in the analysis. |

PPE = personal protective equipment; RR = relative risk.
* If RR's are given for 2,4-D exposure with and without PPE, the RR for without PPE was used in this meta-analysis.

**Supplement Table 3**
Newcastle–Ottawa quality assessment

Burns (2011)—cohort
Selection
1.) C → no stars
2.) B → no stars
3.) A → 1 star
4.) A → 1 star
Comparability
1.) 2 stars (standardized for age, restricted to males)
Outcome
1.) B → 1 star
2.) A → 1 star (20 years)
3.) A → 1 star
Total: 7/9

Cantor (1992)—case-controlSelection
1.) A → 1 star
2.) A → 1 star
3.) A → 1 star
4.) A → 1 star
Comparability
1.) 2 stars (matched on age, restricted to males)
Exposure
1.) C → no stars
2.) A → 1 star
3.) A → 1 star
Total: 8/9

Cocco (2013)—case-controlSelection
1.) A → 1 star (all cases histologically confirmed, 20% of cases independently validated by pathologists)
2.) A → 1 star
3.) B → no stars (a little over half of controls were from hospital controls, the rest were population controls)
4.) A → 1 star
Comparability
1.) 2 stars (matched on age and sex)
Exposure
1.) C → no stars
2.) A → 1 star
3.) C → no stars (response rates were 88% in cases, 81% in hospital controls, and 52% in population controls)
Total: 6/9

Hardell (1994)—case-controlSelection
1.) A → 1 star
2.) A → 1 star
3.) A → 1 star
4.) A → 1 star
Comparability
1.) Two stars (matched on age and sex)
Exposure
1.) D → no stars
2.) A → 1 star
3.) C → 1 star
Total: 7/9

Hoar (1986)—case-controlSelection
1.) A → 1 star
2.) A → 1 star
3.) A → 1 star
4.) A → 1 star
Comparability
1.) Two stars (matched on age, restricted to males)
Exposure
1.) C → no stars
2.) A → 1 star
3.) A → 1 star (96% for cases and 94% for controls)
Total: 8/9

Kogevinas (1995)—nested case-controlSelection
1.) B → no stars (some cases identified by cause of death on death certificates)
2.) A → 1 star (all cases sought in death certificates and from registers where they existed)
3.) A → 1 star
4.) A → 1 star
Comparability
1.) 2 stars (matched on age and sex)
Exposure
1.) A → 1 star
2.) A → 1 star
3.) C → no stars
Total: 7/9

McDuffie (2001)—case-controlSelection
1.) A → 1 star (pathological material reviewed and classified by reference pathologist)
2.) A → 1 star
3.) A → 1 star (random sample of controls from voter records, health insurances, or telephone listings)
4.) A → 1 star
Comparability
1.) 2 stars → (matched on age, restricted to males)
Exposure
1.) D → no stars
2.) A → 1 star
3.) C → no stars (67% of cases and 48% of controls used in analysis)
Total: 7/9

Miligi (2006)—case-controlSelection
1.) A → 1 star
2.) A → 1 star
3.) A → 1 star
4.) A → 1 star
Comparability
1.) Two stars → (matched on age and sex)
Exposure
1.) B → 1 star (agronomists reviewed questionnaires and assessed exposure blindly)
2.) A → 1 star
3.) C → no stars (did not specify response rates)
Total: 8/9

Mills (2005)—Nested case-controlSelection
1.) B → no stars
2.) A → 1 star
3.) A → 1 star
4.) A → 1 star
Comparability
1.) Two stars (matched on age and sex)
Exposure
1.) E → no stars
2.) A → 1 star
3.) A → 1 star
Total: 7/9

Nordstrom (1998)—case-controlSelection
1.) B → no stars
2.) A → 1 star
3.) A → 1 star
4.) A → 1 star
Comparability
1.) Two stars (matched on age, restricted to males)
Exposure
1.) D → no stars (questionnaires mailed to participants)
2.) A → 1 star
3.) A → 1 star (91% of cases and 83% of controls)
Total: 7/9

Woods (1987)—Case-controlSelection
1.) B → no stars
2.) A → 1 star
3.) A → 1 star
4.) A → 1 star
Comparability
1.) Two stars → (matched on age, restricted to males)
Exposure
1.) C → no stars
2.) A → 1 star
3.) A → 1 star (both response rates >70% in the underlying study)
Total: 7/9

Zahm (1990)—Case-controlSelection
1.) A → 1 star
2.) A → 1 star
3.) A → 1 star
4.) A → 1 star

(continued)

(continued on next page)

**Supplement Table 3** (*continued*)

| |
|---|
| Comparability |
| 1.) Two stars → (matched on age, restricted to males) |
| Exposure |
| 1.) B → 1 star (telephone interviews, interviewers were blinded) |
| 2.) A → 1 star |
| 3.) A → 1 star (cases were 91% and controls were 87%) |
| Total: 9/9 |