APPENDIX C

Meta-Analysis of Cancer Results from Epidemiological Studies

CONTENTS—Appendix C: Meta-Analysis of Cancer Results from Epidemiological Studies

LIST OF TABLES ..... C-iii
LIST OF FIGURES ..... C-iv

APPENDIX C. META-ANALYSIS OF CANCER RESULTS FROM EPIDEMIOLOGICAL STUDIES C-1
C.1. METHODOLOGY C-1
C.2. META-ANALYSIS FOR LYMPHOMA C-4
C.2.1. Overall Effect of TCE Exposure C-4
C.2.1.1. Selection of RR Estimates C-4
C.2.1.2. Results of Meta-Analyses C-10
C.2.2. Lymphoma Effect in the Highest Exposure Groups C-15
C.2.2.1. Selection of RR Estimates C-15
C.2.2.2. Results of Meta-Analyses C-19
C.2.3. Discussion of Lymphoma Meta-Analysis Results C-21
C.3. META-ANALYSIS FOR KIDNEY CANCER C-24
C.3.1. Overall Effect of TCE Exposure C-24
C.3.1.1. Selection of RR Estimates C-24
C.3.1.2. Results of Meta-Analyses C-30
C.3.2. Kidney Cancer Effect in the Highest Exposure Groups C-33
C.3.2.1. Selection of RR Estimates C-33
C.3.2.2. Results of Meta-Analyses C-39
C.3.3. Discussion of Kidney Cancer Meta-Analysis Results C-43
C.4. META-ANALYSIS FOR LIVER CANCER C-46
C.4.1. Overall Effect of TCE Exposure C-46
C.4.1.1. Selection of RR Estimates C-46
C.4.1.2. Results of Meta-Analyses C-49
C.4.2. Liver Cancer Effect in the Highest Exposure Groups C-52
C.4.2.1. Selection of RR Estimates C-52
C.4.2.2. Results of Meta-Analyses C-56
C.4.3. Discussion of Liver Cancer Meta-Analysis Results C-58
C.5. DISCUSSION OF STRENGTHS, LIMITATIONS, AND UNCERTAINTIES IN THE META-ANALYSES C-59
C.6. CONCLUSIONS C-60
C.7. REFERENCES C-62

LIST OF TABLES

C-1. Selected RR estimates for lymphoma associated with TCE exposure (overall effect) from cohort studies C-6

C-2. Selected RR estimates for lymphoma associated with TCE exposure from case-control studies C-8

C-3. Summary of some meta-analysis results for TCE (overall) and lymphoma C-11

C-4. Selected RR estimates for lymphoma risk in highest TCE exposure groups C-17

C-5. Summary of some meta-analysis results for TCE (highest exposure groups) and lymphoma C-20

C-6. Selected RR estimates for kidney cancer associated with TCE exposure (overall effect) from cohort studies C-26

C-7. Selected RR estimates for renal cell carcinoma associated with TCE exposure from case-control studies C-27

C-8. Summary of some meta-analysis results for TCE (overall) and kidney cancer C-31

C-9. Selected RR estimates for kidney cancer risk in highest TCE exposure groups C-35

C-10. Summary of some meta-analysis results for TCE (highest exposure groups) and kidney cancer C-41

C-11. Selected RR estimates for liver cancer associated with TCE exposure (overall effect) from cohort studies C-47

C-12. Summary of some meta-analysis results for TCE and liver cancer C-50

C-13. Selected RR estimates for liver cancer risk in highest TCE exposure groups C-54

LIST OF FIGURES

C-1. Meta-analysis of lymphoma and overall TCE exposure C-12

C-2. Funnel plot of SE by log RR estimate for TCE and lymphoma studies C-14

C-3. Cumulative meta-analysis of TCE and lymphoma studies, progressively including studies with increasing SEs C-15

C-4. Meta-analysis of lymphoma and TCE exposure—highest exposure groups C-21

C-5. Meta-analysis of kidney cancer and overall TCE exposure C-32

C-6. Funnel plot of SE by log RR estimate for TCE and kidney cancer studies C-34

C-7. Meta-analysis of kidney cancer and TCE exposure—highest exposure groups C-42

C-8. Meta-analysis of kidney cancer and TCE exposure—highest exposure groups, with assumed null RR estimates for Anttila, Axelson, and Hansen (see text) C-43

C-9. Meta-analysis of liver cancer and TCE exposure C-51

C-10. Funnel plot of SE by log RR estimate for TCE and liver cancer studies C-53

C-11. Meta-analysis of liver cancer and TCE exposure—highest exposure groups, with assumed null RR estimates for Hansen and Zhao (see text) C-57

APPENDIX C. META-ANALYSIS OF CANCER RESULTS FROM EPIDEMIOLOGICAL STUDIES

C.1. METHODOLOGY

An initial review of the epidemiological studies indicated some evidence for associations between trichloroethylene (TCE) exposure and lymphomas and cancers of the kidney and liver. To investigate further these possible associations, we performed meta-analyses of the epidemiological study results for these three cancer types. Meta-analysis provides a systematic way to combine study results for a given effect across multiple (sufficiently similar) studies. The resulting summary (weighted average) estimate is a quantitatively objective way of reflecting results from multiple studies, rather than relying on a single study, for instance. Combining the results of smaller studies to obtain a summary estimate also increases the statistical power to observe an effect, if one exists. Furthermore, meta-analyses typically are accompanied by other analyses of the epidemiological studies, including analyses of publication bias and investigations of possible factors responsible for any heterogeneity across studies.

Given the diverse nature of the epidemiological studies for TCE, random-effects models were used for the primary analyses, and fixed-effect analyses were conducted for comparison. Both approaches combine study results (in this case, relative risk [RR] estimates) weighted by the inverse invariance; however, they differ in their underlying assumptions about what the study results represent and how the variances are calculated. For a random-effects model, it is assumed that there is true heterogeneity across studies and that both between-study and within-study components of variation need to be taken into account; this was done using the methodology of DerSimonian and Laird (1986). For a fixed-effect model, it is assumed that the studies are all essentially measuring the same thing and all the variance is within-study variance; thus, for the fixed-effect model, the RR estimate from each study is simply weighted by the inverse of the (within-study) variance of the estimate.

Studies for the meta-analyses were selected as described in Appendix B, Section II-9. The general approach for selecting RR estimates was to select the reported RR estimate that best reflected an RR for TCE exposure vs. no TCE exposure (overall effect). When available, RR estimates from internal analyses were selected over standardized incidence or mortality ratios (SIRs, SMRs) and adjusted RR estimates were generally selected over crude estimates. Incidence estimates would normally be preferred to mortality estimates; however, for the two studies providing both incidence and mortality results, incidence ascertainment was for a substantially shorter period of time than mortality follow-up, so the endpoint with the greater number of cases was used to reflect the results that had better case ascertainment. For separate

analyses, an RR estimate for the highest exposure group was selected from studies that presented results for different exposure groups. Exposure groups based on some measure of cumulative exposure were preferred, if available; however, often duration was the sole exposure metric used. Specific selection choices are described in the following subsections detailing the actual analyses.

The meta-analysis calculations are based on (natural) logarithm-transformed values. Thus, each RR estimate was transformed to its natural logarithm (referred to here as "log RR," the conventional terminology in epidemiology), and either an estimate of the standard error (SE) of the log RR was obtained, from which to estimate the variance for the weights, or an estimate of the variance of the log RR was calculated directly. If the reported 95% confidence interval limits were proportionally symmetric about the observed RR estimate (i.e., upper confidence limit/RR ≈ RR/lower confidence limit), then an estimate of the SE of the log RR estimate was obtained using the formula

$$SE = \frac{\left[log\left(UCL\right) - log\left(LCL\right) \right]}{3.92}, \qquad \text{(Eq. C-1)}$$

where UCL is the upper confidence limit and LCL is the lower confidence limit (for 90% confidence intervals [CIs], the divisor is 3.29) (Rothman and Greenland, 1998). In all the TCE cohort studies reporting SMRs or SIRs as the overall RR estimates, reported CIs were calculated assuming the number of deaths (or cases) is approximately Poisson distributed. In such cases, the CIs are not proportionally symmetric about the RR estimate (unless the number of deaths is fairly large), and the SE of the log RR estimate was estimated as the inverse of the square root of the observed number of deaths (or cases) (Breslow and Day, 1987). In some case-control studies, no overall odds ratio (OR) was reported, so a crude OR estimate was calculated as OR = (a/b)/(c/d), where a, b, c, and d are the cell frequencies in a 2 × 2 table of cancer cases vs. TCE exposure, and the variance of the log OR was estimated using the formula

$$Var\left[log\left(OR\right) \right] = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}, \qquad \text{(Eq. C-2)}$$

in accordance with the method proposed by Woolf (1955), as described by Breslow and Day (1980).

The analyses that were performed for this assessment include

- meta-analyses to obtain overall summary estimates of RR
- heterogeneity analyses
- analyses of the influence of single studies on the summary estimates
- analyses of the sensitivity of the summary estimate to alternate study inclusion selections or to alternate selections of RR estimates from a study
- publication bias analyses
- meta-analyses to obtain summary estimates for the highest exposure groups in studies that provide data by exposure group, and
- consideration of some potential sources of heterogeneity across studies.

The analyses were conducted using Excel spreadsheets and the software package Comprehensive Meta-Analysis, Version 2 (© 2006, Biostat, Inc.). Figures were generated using the Comprehensive Meta-Analysis software. Note that for these figures, this software recalculates CIs for the studies based on the SE inputs, and the resulting CIs are not always identical to those reported in the original studies, in particular those based on Poisson distributions. However, the recalculated CIs are merely outputs and are not the basis for any calculations in the software; SEs were obtained as described above, and these SEs and the log RRs constitute the inputs for the meta-analysis calculations.

The heterogeneity (or homogeneity) analysis tests the hypothesis that the study results are homogeneous, i.e., that all the RR estimates are estimating the same population RR and the total variance is no more than would be expected from within-study variance. Heterogeneity was assessed using the statistic Q described by DerSimonian and Laird (1986). The Q-statistic represents the sum of the weighted squared differences between the summary RR estimate (obtained under the null hypothesis, i.e., using a fixed-effect model) and the RR estimate from each study, and, under the null hypothesis, Q approximately follows a χ^2 distribution with degrees of freedom equal to the number of studies minus one. However, this test can be under-powered when the number of studies is small, and it is only a significance test, i.e., it is not very informative about the *extent* of any heterogeneity. Therefore, the I^2 value (Higgins et al., 2003) was also considered. $I^2 = 100\% \times (Q - df)/Q$, where Q is the Q-statistic and df is the degrees of freedom, as described above. This value estimates the percentage of variation that is due to study heterogeneity. Typically, I^2 values of 25%, 50%, and 75% are considered low, moderate, and high amounts of heterogeneity, respectively.

Subgroup analyses were sometimes conducted to examine whether or not the combined RR estimate varied significantly between different types of studies (e.g., case-control vs. cohort studies). In such subgroup analyses of categorical variables (e.g., study design), analysis of variance was used to determine if there was significant heterogeneity between the subgroups. Applying analysis of variance to meta-analyses with two subgroups (df = 1), $Q_{\text{between subgroups}} = Q_{\text{overall}} - (Q_{\text{subgroup1}} + Q_{\text{subgroup2}}) = z\text{-value}^2$, where Q_{overall} is the Q-statistic calculated across all the studies and $Q_{\text{subgroup1}}$ and $Q_{\text{subgroup2}}$ are the Q-statistics calculated within each subgroup.

Publication bias is a systematic error that occurs if statistically significant studies are more likely to be submitted and published than nonsignificant studies. Studies are more likely to be statistically significant if they have large effect sizes (in this case, RR estimates); thus, an upward bias would result in a meta-analysis if the available published studies have higher effect sizes than the full set of studies that was actually conducted. One feature of publication bias is that smaller studies tend to have larger effect sizes than larger studies, since smaller studies need larger effect sizes in order to be statistically significant. Thus, many of the techniques used to analyze publication bias examine whether or not effect size is associated with study size. Methods used to investigate potential publication bias for this assessment included funnel plots, which plot effect size vs. study size (actually, SE vs. log RR here); the "trim and fill" procedure of Duvall and Tweedie (2000), which imputes the "missing" studies in a funnel plot (i.e., the studies needed to counterbalance an asymmetry in the funnel plot resulting from an ostensible publication bias) and recalculates a summary effect size with these studies present; forest plots (arrays of RRs and CIs by study) sorted by precision (i.e., SE) to see if effect size shifts with study size; Begg and Mazumdar rank correlation test (Begg and Mazumdar, 1994), which examines the correlation between effect size estimates and their variances after standardizing the effect sizes to stabilize the variances; Egger's linear regression test (Egger et al., 1997), which tests the significance of the bias reflected in the intercept of a regression of effect size/SE on 1/SE; and cumulative meta-analyses after sorting by precision to assess the impact on the summary effect size estimate of progressively adding the smaller studies.

C.2. META-ANALYSIS FOR LYMPHOMA

C.2.1. Overall Effect of TCE Exposure

C.2.1.1. *Selection of RR Estimates*

The selected RR estimates for lymphoma associated with TCE exposure from the selected epidemiological studies are presented in Table C-1 for cohort studies and in Table C-2 for case-control studies. A few of the more recent case-control studies classified lymphomas along the lines of the recent WHO/REAL classification system (World Health

Organization/Revised European-American Classification of Lymphoid Neoplasms) (Harris et al., 2000); however, most of the available TCE studies reported lymphoma results according to the International Classification of Diseases (ICD), Revisions 7, 8, and 9, and focused on non-Hodgkin lymphoma (NHL; ICD 200 + 202). For consistency of endpoint in the lymphoma meta-analyses, RR estimates for ICD 200 + 202 were selected, wherever possible; otherwise, estimates for the classification(s) best approximating NHL were selected. In addition, many of the studies provided RR estimates only for males and females combined, and we are not aware of any basis for a sex difference in the effects of TCE on lymphoma risk; thus, wherever possible, RR estimates for males and females combined were used. The only study of much size (in terms of number of lymphoma cancer cases) that provided results separately by sex was Raaschou-Nielsen (2003). This study reports an insignificantly higher SIR for females (1.4, 95% CI: 0.73, 2.34) than for males (1.2, 95% CI: 0.98, 1.52).

Beyond selecting adjusted RR estimates for lymphoma classification and both sexes, when multiple estimates were available, the preference was to select the RR estimate that represented the largest population in a study, while trying to minimize the likelihood of TCE exposure misclassification. Sensitivity analyses were generally done to investigate the impact of these alternate selection choices, as well as to estimate the impacts of study findings that were not reported.

Thus, for example, for Axelson et al. (1994), in which a small subcohort of females was studied but only results for the larger male subcohort were reported, the reported male-only results were used in the primary analysis; however, an attempt was made to estimate the female contribution to an overall RR estimate for both sexes and its impact on the meta-analysis. Axelson et al. (1994) reported that there were no cases of lymphoma observed in females, but the expected number was not presented. To estimate the expected number, the expected number for males was multiplied by the ratio of female-to-male person-years in the study and by the ratio of female-to-male age-adjusted incidence rates for NHL.[1] The male results and the estimated female contribution were then combined into an RR estimate for both sexes assuming a Poisson distribution, and this alternate RR estimate for the Axelson et al. (1994) study was used in a sensitivity analysis.

---

[1]Person-years for men and women $\leq$79 years were obtained from Axelson et al. (1994): 23516.5 and 3691.5, respectively. Lifetime age-adjusted incidence rates for NHL for men and women were obtained from the National Cancer Institute's 2000-2004 SEER-17 (Surveillance Epidemiology and End Results from 17 geographical areas) database (http://seer.cancer.gov/statfacts/html/nhl.html): 23.2/100,000 and 16.3/100,000, respectively. The calculation for estimating the expected number of cases in females in the cohort assumes that the males and females have similar TCE exposures and that the relative distributions of age-related incidence risk for the males and females in the cohort are adequately represented by the ratios of person-years and lifetime incidence rates used in the calculation.

Table C-1. Selected RR estimates for lymphoma associated with TCE exposure (overall effect) from cohort studies

Study	RR	95% LCL	95% UCL	RR type	log RR	SE(log RR)	Alternate RR estimates	Comments
Anttila et al., 1995	1.81	0.78	3.56	SIR	0.593	0.354	None	ICD-7 200 + 202.
Axelson et al., 1994	1.52	0.49	3.53	SIR	0.419	0.447	1.36 (0.44, 3.18) with estimated female contribution to SIR added (see text)	ICD-7 200 and 202. Results reported separately; combined assuming Poisson distribution. Results reported for males only, but there was a small female component to the cohort.
Boice et al., 1999	1.19	0.65	1.99	SMR	0.174	0.267	1.19 (0.83, 1.65) for any potential exposure	ICD-9 200 + 202. For potential routine exposure.
Greenland et al., 1994	0.76	0.24	2.42	OR	−0.274	0.590	None	ICD-8 200-202. Nested case-control study.
Hansen et al., 2001	3.1	1.3	6.1	SIR	1.13	0.354	None	ICD-7 200 + 202. Male and female results reported separately; combined assuming Poisson distribution.
Morgan et al., 1998	1.01	0.46	1.92	SMR	0.00995	0.333	1.36 (0.35, 5.21) unpublished RR for ICD 200 (see text)	ICD 200 + 202. Results reported by Mandel et al. (2006). ICD Revision 7, 8, or 9, depending on year of death.
Raaschou-Nielsen et al., 2003	1.24	1.01	1.52	SIR	0.215	0.104	None	ICD-7 200 + 202.
Radican et al., 2008	1.36	0.77	2.39	Mortality HR	0.307	0.289	None	ICD-8,-9 200 + 202; ICD-10 C82-C85. Time variable = age; covariates = sex and race. Referent group is workers with no chemical exposures.

Table C-1. Selected RR estimates for lymphoma associated with TCE exposure (overall effect) from cohort studies (continued)

Study	RR	95% LCL	95% UCL	RR type	log RR	SE(log RR)	Alternate RR estimates	Comments
Zhao et al., 2005	1.44	0.90	2.30	Mortality RR	0.363	0.239	Incidence RR: 0.77 (0.42, 1.39) Boice 2006 SMR for ICD-9 200 + 202: 0.21 (0.01, 1.18)	All lymphohematopoietic cancer (ICD-9 200-208), not just 200 + 202. Males only; adjusted for age, socioeconomic status (SES), time since first employment. Mortality results reflect more exposed cases (33) than do incidence results (17). Overall RR estimated by combining across exposure groups (see text). Boice 2006 cohort overlaps Zhao cohort; just 1 exposed death for ICD 200 + 202; 9 for 200−208 (vs. 33 in Zhao).

Table C-2. Selected RR estimates for lymphoma associated with TCE exposure from case-control studies[a]

Study	RR	95% LCL	95% UCL	log RR	SE(log RR)	Lymphoma type	Comments
Hardell et al., 1994	7.2	1.3	42	1.97	0.887	NHL	Rappaport classification system. Males only; controls matched for age, place of residence, vital status.
Miligi et al., 2006	0.93	--[b]	--[b]	−0.0726	0.168	NHL + CLL	NCI working formulation. Crude OR; overall adjusted OR not presented.
Nordstrom et al., 1998	1.5	0.7	3.3	0.405	0.396	HCL	HCL specifically. Males only; controls matched for age and county; analysis controlled for age.
Persson and Frederikson, 1999	1.2	0.5	2.4	0.182	0.400	NHL	Classification system not specified. Controls selected from same geographic areas; ORs stratified on age and sex.
Seidler et al., 2007	1.0	0.74	1.4	−0.223	0.177	B-cell and T-cell NHL	WHO classification. Overall results for B-cell and T-cell NHL from personal communication (see text). Adjusted for smoking and alcohol consumption. Case-control pairs matched on sex, region, and age.
Siemiatycki, 1991	1.1	0.5	2.5	0.0953	0.424	NHL	ICD-9 200 + 202. SE and 95% CI calculated from reported 90% CIs; males only; adjusted for age, income, and cigarette smoking index.
Wang et al., 2009	1.2	0.9	1.8	0.182	0.177	"NHL"; various lymphoma subtypes + mast cell tumors	ICD-O M-9590-9642, 9690-9701, 9740-9750. Females only; adjusted for age, family history of lymphohematopoietic cancers, alcohol consumption, and race.

[a]The RR estimates are all ORs for incident cases.
[b]Not calculated.

NHL: non-Hodgkin lymphoma; CLL: chronic lymphocytic leukemia; HCL: hairy cell leukemia (a subgroup of NHL).

Most of the selections in Tables C-1 and C-2 should be self-evident, but some are discussed in more detail here, in the order the studies are presented in the tables. For Boice et al. (1999), results for "potential routine exposure" were selected for the primary analysis, because this exposure category was considered to have less exposure misclassification, and results for "any potential exposure" were used in a sensitivity analysis. The Greenland et al. (1994) study is a case-control study nested within a worker cohort, and we treat it here as a cohort study (see Appendix B, Section II-9.1). For Morgan et al. (1998), the reported results did not allow for the combination of ICD 200 and 202, so the SMR estimate for the combined 200 + 202 grouping was taken from the meta-analysis paper of Mandel et al. (2006), who included one of the investigators from the Morgan et al. (1998) study. RR estimates for overall TCE exposure from internal analyses of the Morgan et al. (1998) cohort data were available from an unpublished report (Environmental Health Strategies, 1997; the published paper only presented the internal analyses results for exposure subgroups), but only for ICD 200; from these, the RR estimate from the Cox model which included age and sex was selected, because those are the variables deemed to be important in the published paper (Morgan et al., 1998). Although the results from internal analyses are generally preferred, in this case the SMR estimate was used in the primary analysis and the internal analysis RR estimate was used in a sensitivity analysis because the latter estimate represented an appreciably smaller number of deaths (3, based on ICD 200 only) than the SMR estimate (9, based on ICD 200 + 202). For Radican et al. (2008), the Cox model hazard ratio (HR) from the 2000 follow-up was used. In the Radican et al. (2008) Cox regressions, age was the time variable, and sex and race were covariates. It should also be noted that the referent group is composed of workers with no chemical exposures, not just no exposure to TCE.

For Zhao et al. (2005), RR estimates were only reported for ICD-9 200−208 (all lymphohematopoietic cancers), and not for 200 + 202 alone. Given that other studies have not reported associations between leukemias and TCE exposure, combining all lymphohematopoietic cancers would dilute any lymphoma effect, and the Zhao results are expected to be an underestimate of any TCE effect on lymphoma alone. Another complication with the Zhao et al. (2005) study is that no results for an overall TCE effect are reported. We were unable to obtain any overall estimates from the study authors, so, as a best estimate, the results across the "medium" and "high" exposure groups were combined, under assumptions of group independence, even though the exposure groups are not independent (the "low" exposure group was the referent group in both cases). Zhao et al. (2005) present RR estimates for both incidence and mortality; however, the time frame for the incidence accrual is smaller than the time frame for mortality accrual and fewer exposed incident cases (17) were obtained than deaths (33). Thus, because better case ascertainment occurred for mortality than for incidence, the mortality

results were used for the primary analysis, and the incidence results were used in a sensitivity analysis. A sensitivity analysis was also done using results from Boice et al. (2006) in place of the Zhao et al. (2005) RR estimate. The cohorts for these studies overlap, so they are not independent studies and should not be included in the meta-analysis concurrently. Boice et al. (2006) report an RR estimate for an overall TCE effect for lymphoma alone; however, it is based on far fewer cases (1 death in ICD-9 200 + 202; 9 deaths for 200−208) and is an SMR rather than an internal analysis RR estimate, so the Zhao et al. (2005) estimates are preferred for the primary analysis.

For the case-control studies, the main issue was the lymphoma classifications. Miligi et al. (2006) include chronic lymphocytic leukemias (CLLs) in their NHL results, consistent with the current WHO/REAL classification. Also, Miligi et al. (2006) do not report an overall adjusted RR estimate, so a crude estimate of the OR was calculated for the two TCE exposure categories together vs. no TCE exposure. The Nordstrom et al. (1998) study was a case-control study of hairy cell leukemias (HCLs), which are a subgroup of NHLs, so only results for HCL were reported. For Seidler et al. (2007), an overall adjusted OR for B-cell and T-cell NHL combined was kindly provided by Dr. Seidler (personal communication from Andreas Seidler, Bundesanstalt fur Arbeitsschutz u. Arbeitsmedizin, to Cheryl Scott, U.S. EPA, 13 November 2007). Wang et al. (2009) refer to their cases as "NHL" cases; however, according to the ICD-O classification system that they used, their cases are more specifically various particular subtypes of malignant lymphoma (9590-9642, 9690-9701) and mast cell tumors (9740-9750) (Morton et al., 2003). No alternate RR estimates were considered for any of the case-control studies of lymphoma.

C.2.1.2. *Results of Meta-Analyses*

Results from some of the meta-analyses that were conducted on the epidemiological studies of TCE and lymphoma are summarized in Table C-3. The summary estimate from the primary random effects meta-analysis of the 16 studies was 1.23 (95% CI: 1.04, 1.44) (see Figure C-1). No single study was overly influential; removal of individual studies resulted in summary, or "pooled," RR (RRp) estimates that ranged from 1.16 (with the removal of Hansen) to 1.28 (with the removal of Seidler) and were all statistically significant. Removal of Hardell, whose RR estimate is a relative outlier (see Figure C-1), only decreased the RRp estimate to 1.20 (1.04, 1.39), since this study does not contribute a lot of weight to the meta-analysis. Removal of studies other than Hansen or Hardell resulted in RRp estimates that were all greater than 1.20.

Table C-3. Summary of some meta-analysis results for TCE (overall) and lymphoma

Analysis	# of studies	Model	Summary RR estimate (RRp)	95% LCL	95% UCL	Heterogeneity	Comments
All studies	16	Random	1.23	1.04	1.44	Not significant (p = 0.10)	Statistical significance of RRp not dependent on individual studies.
		Fixed	1.19	1.06	1.34		
Cohort	9	Random	1.35	1.13	1.61	Not significant (p = 0.35)	Not significant difference between CC and cohort studies (p = 0.13).
		Fixed	1.33	1.14	1.54		Significant difference between CC and cohort studies (p = 0.03).
Case-control	7	Random	1.07	0.84	1.37	Not significant (p = 0.17)	
		Fixed	1.03	0.86	1.23		
Alternate RR selections[a]	16	Random	1.19	1.00	1.41	Not significant (p = 0.07)	With estimated Zhao overall RR for incidence rather than mortality.
	16	Random	1.21	1.01	1.45	Not significant (p = 0.053)	With Boice (2006) study rather than Zhao.
	16	Random	1.22	1.04	1.44	Not significant (p = 0.10)	With estimated female contribution to Axelson.
	16	Random	1.22	1.05	1.43	Not significant (p = 0.10)	With Boice (1999) any potential exposure SMR.
	16	Random	1.24	1.05	1.46	Not significant (p = 0.10)	With Morgan et al. (1998) unpublished RR.
Highest exposure groups	12	Random	1.57	1.27	1.94	None observable (fixed = random)	Statistical significance not dependent on single study. See Table C-5 for results with alternate RR selections.
		Fixed	1.57	1.27	1.94		

[a]Changing the primary analysis by one alternate RR each time; more details on alternate RR estimates in text.

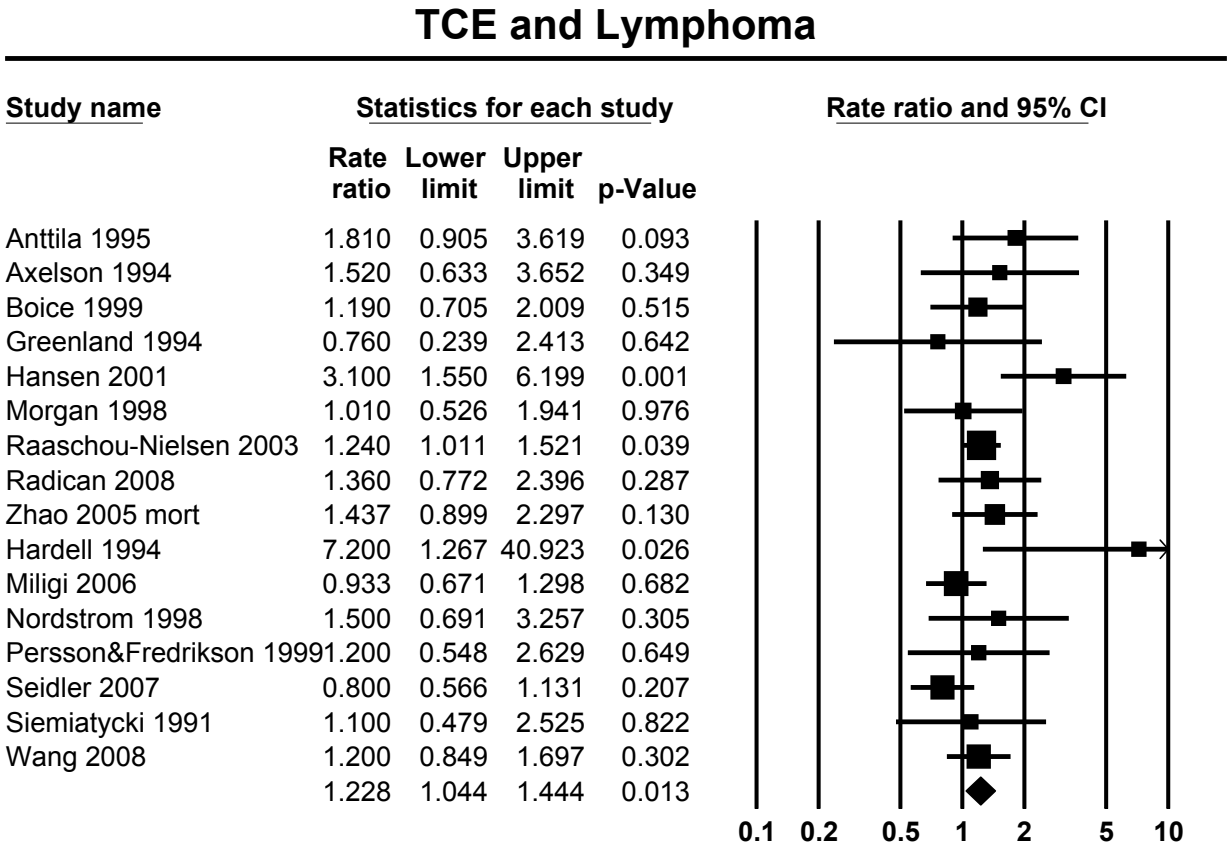

Figure C-1. Meta-analysis of lymphoma and overall TCE exposure. The pooled estimate is in the bottom row. Symbol sizes reflect relative weights of the studies. The horizontal midpoint of the bottom diamond represents the summary RR estimate, and the horizontal extremes depict the 95% CI limits.

Similarly, the RRp estimate was not highly sensitive to alternate RR estimate selections. Use of the five alternate selections, individually, resulted in RRp estimates that ranged from 1.19 to 1.24 (see Table C-3) and were all statistically significant except when the Zhao incidence estimate ($p = 0.050$) was used instead of the Zhao mortality estimate. As discussed above, the Zhao mortality estimate is preferred over the incidence estimate in this instance because it is based on nearly twice as many cases (33 vs. 17).

There was some heterogeneity apparent across the 16 studies, although it was not statistically significant ($p = 0.10$). The I^2 value (see Section C.1) was 33%, suggesting low-to-moderate heterogeneity. Subgroup analyses were done examining the cohort and case-control studies separately. With the random effects model (and tau-squared not pooled across subgroups), the resulting RRp estimates were 1.35 (95% CI: 1.13, 1.61) for the cohort studies and 1.07 (0.84, 1.37) for the case-control studies. There was residual heterogeneity in each of the subgroups, but in neither case was it statistically. I^2 values were 10% for the cohort studies, suggesting low heterogeneity, and 33% for the case-control studies, suggesting low-to-moderate heterogeneity. The difference between the RRp estimates for the cohort and case-control subgroups was not statistically significant under the random effects model, although it was under the fixed effect model (see Table C-3). Some thought was given to further analyses to investigate the source(s) of the heterogeneity, such as qualitative tiering or subgroups based on likelihood for correct exposure classification or on likelihood for higher vs. lower exposures across the studies. Ultimately, these approaches were rejected because in many of the studies it was difficult to judge (and weight) the extent of exposure misclassification or the degree of TCE exposure with any precision. In other words, there was inadequate information to reliably assess either the extent to which each study accurately classified exposure status or the relative TCE exposure levels and prevalences of exposure to different levels across studies. See Section C.2.3 below for a qualitative discussion of some potential sources of heterogeneity.

As discussed in Section C.1, publication bias was examined in several different ways. The funnel plot in Figure C-2 suggests some relationship between RR estimate and study size (if there were no relationship, the studies would be symmetrically distributed around the pooled RR estimate rather than veering towards higher RR estimates with increasing SEs), although the observed asymmetry is highly influenced by the Hardell study, which is a relative outlier and which contributes little weight to the overall meta-analysis, as discussed above. The Begg and Mazumdar rank correlation test and Egger's linear regression test were not statistically significant; it should be noted, however, that both of these tests have low power. Duval and Tweedie's trim-and-fill procedure yielded a pooled RR estimate (under the random effects model) of 1.13 (95% CI: 0.94, 1.35) when the 4 studies deemed missing from the funnel plot were filled into the meta-analysis (these studies are filled in so as to counter-balance the apparent asymmetry of the more extreme values in the funnel plot). Eliminating the Hardell study made little difference to the results of the publication bias analyses. The results of a cumulative meta-analysis, incorporating studies with increasing SE one at a time, are depicted in Figure C-3. This procedure is a transparent way of examining the effects of including studies with increasing SE. The figure shows that the pooled RR estimate is 1.05 after inclusion of the 4 largest (i.e.,

most precise) studies, which constitute about 50% of the weight. The pooled RR estimate increases to 1.12 with inclusion of the 8 most precise studies, which represent ½ of the total number of studies and about 75% of the weight. The pooled RR estimate becomes fairly stable after addition of the next 2 most precise study (RRp = 1.21), which adds another 9% of the weight. Adding in the 6 least precise studies (16% of the weight) barely increases the pooled RR estimate further. In summary, there is some evidence of potential publication bias in this data set. It is uncertain, however, that this reflects actual publication bias rather than an association between effect size and SE resulting for some other reason, e.g., a difference in study populations or protocols in the smaller studies. Furthermore, if there is publication bias in this data set, it does not appear to account completely for the findings of an increased lymphoma risk.

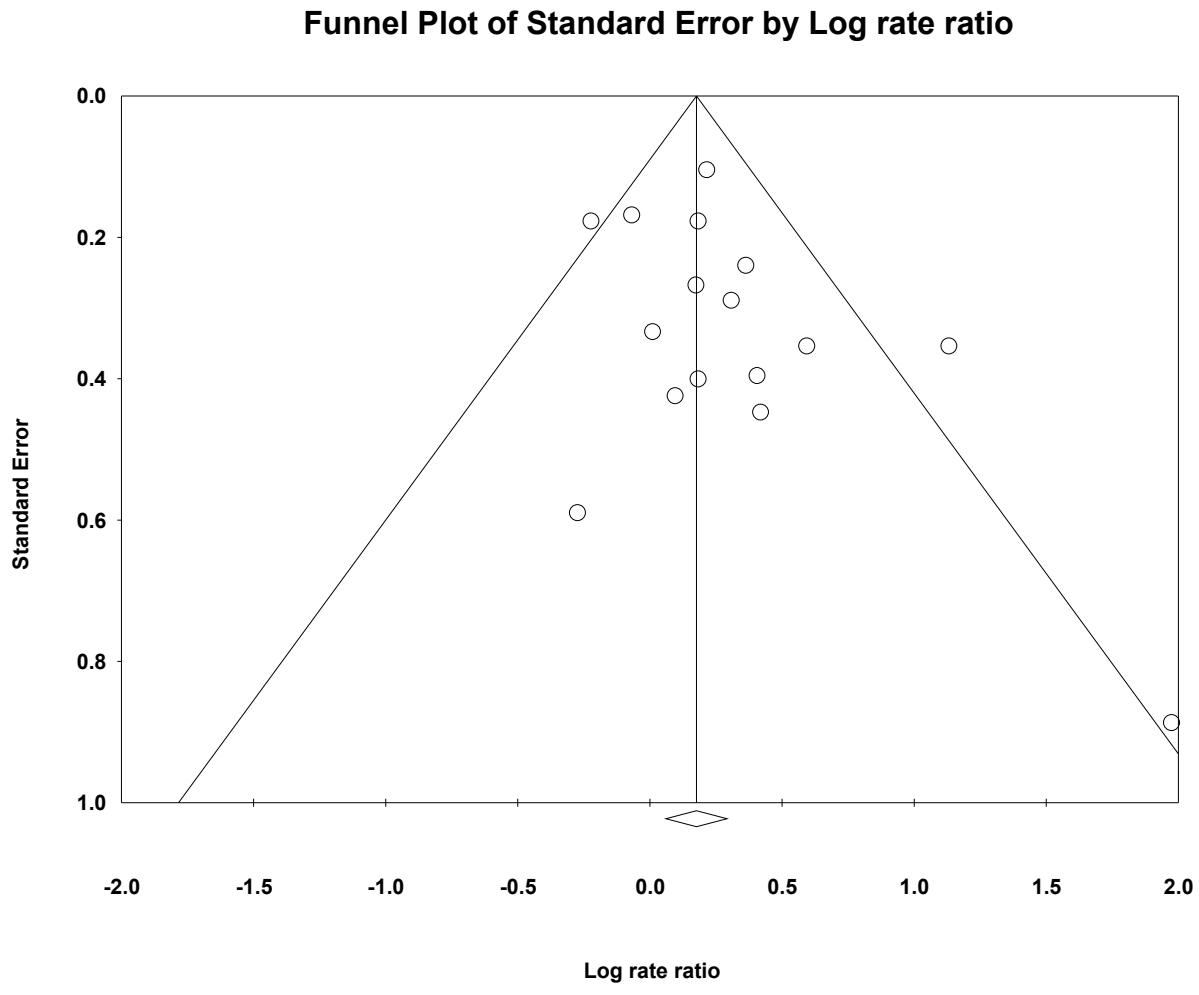

Figure C-2. Funnel plot of SE by log RR estimate for TCE and lymphoma studies.

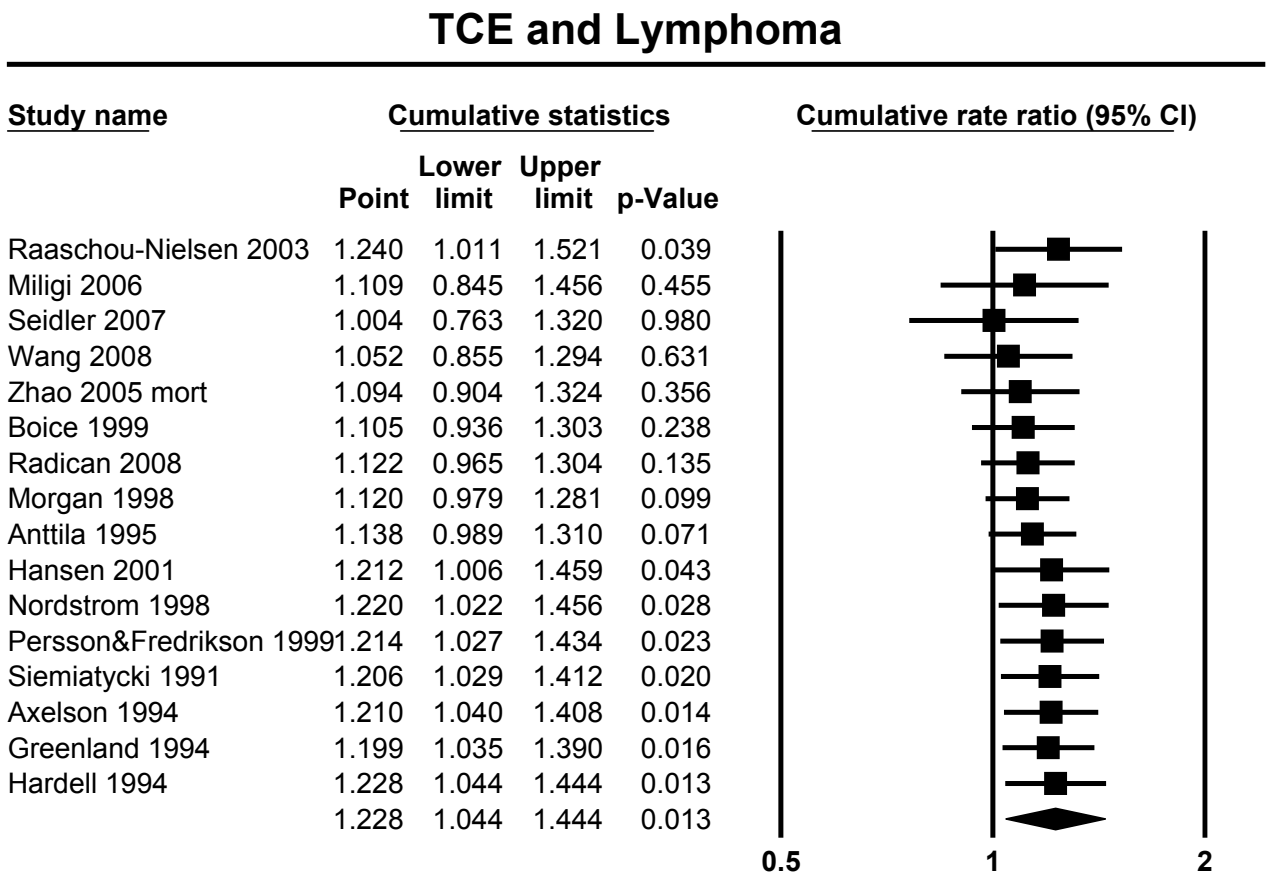

Figure C-3. Cumulative meta-analysis of TCE and lymphoma studies, progressively including studies with increasing SEs.

C.2.2. Lymphoma Effect in the Highest Exposure Groups

C.2.2.1. *Selection of RR Estimates*

The selected RR estimates for lymphoma in the highest TCE exposure categories, for studies that provided such estimates, are presented in Table C-4. All 8 cohort studies (but not the nested case-control study of Greenland et al. [1994]) and 4 of the 7 case-control studies did report lymphoma risk estimates categorized by exposure level. As in Section C.2.1.1 for the overall risk estimates, estimates to best correspond to NHL as represented by ICD-7, -8, and -

9 200 and 202 were selected, and, wherever possible, RR estimates for males and females combined were used.

As above for the overall TCE effect, for Axelson et al. (1994), in which a small subcohort of females was studied but only results for the larger male subcohort were reported, the reported male-only high-exposure group results were used in the primary analysis; however, an attempt was made to estimate the female contribution to a high-exposure group RR estimate for both sexes and its impact on the meta-analysis. To estimate the expected number in the highest exposure group for females, the expected number in the highest exposure group for males was multiplied by the ratio of total female-to-male person-years in the study and by the ratio of female-to-male age-adjusted incidence rates for NHL. The RR estimate for both sexes was used as an alternate RR estimate for the Axelson et al. (1994) study in a sensitivity analysis.

For Boice et al. (1999), only results for workers with "any potential exposure" (rather than "potential routine exposure") were presented by exposure category, and the referent group is workers not exposed to any solvent. For Hansen et al. (2001), exposure group data were presented only for males. To estimate the female contribution to a highest-exposure group RR estimate for both sexes, it was assumed that the expected number of cases in females had the same overall-to-highest-exposure group ratio as in males. The RR estimate for both sexes was then calculated assuming a Poisson distribution, and this estimate was used in the primary analysis. Hansen et al. (2001) present results for three exposure metrics; the cumulative exposure metric was preferred for the primary analysis, and results for the other two metrics were used in sensitivity analyses. For Morgan et al. (1998), results did not allow for the combination of ICD 200 and 202, so the highest-exposure group RR estimate for ICD 200 only was used. The primary analysis used results for the cumulative exposure metric, and a sensitivity analysis was done with the results for the peak exposure metric.

For Radican et al. (2008), it should be noted that the referent group is composed of workers with no chemical exposures, not just no exposure to TCE. In addition, exposure group results were reported separately for males and females and were combined for this assessment using inverse-variance weighting, as in a fixed effect meta-analysis. Radican et al. (2008) present only mortality HR estimates by exposure group; however, in an earlier follow-up of this same cohort, Blair et al. (1998) present both incidence and mortality RR estimates by exposure group. The mortality RR estimate based on the more recent follow-up of Radican et al. (2008) (17 deaths in the highest exposure group) was used in the primary analysis, while the incidence RR estimate based on similarly combined results from Blair et al. (1998) (9 cases) was used as an alternate estimate in a sensitivity analysis.

Table C-4. Selected RR estimates for lymphoma risk in highest TCE exposure groups

Study	RR	95% LCL	95% UCL	Exposure category	log RR	SE(log RR)	Alternate RR estimates	Comments
Anttila et al., 1995	1.4	0.17	5.04	100+ µmol/L U-TCA[a]	0.336	0.707	none	SIR. ICD 200 + 202.
Axelson et al., 1994	6.25	0.16	34.83	≥2-yr exposure and 100+ mg/L U-TCA	1.83	1.00	5.62 (0.14, 31.3) with estimated female contribution added (see text)	SIR. ICD 200 + 202. Results reported for males only, but there was a small female component to the cohort.
Boice et al., 1999	1.62	0.82	3.22	≥5-yr exposure	0.482	0.349	None	Mortality RR. ICD 200 + 202. For potential routine or intermittent exposure. Adjusted for date of birth, dates 1st and last employed, race, and sex. Referent group is workers not exposed to any solvent.
Hansen et al., 2001	2.7	0.56	8.0	≥1080 mos × mg/m^3	0.993	0.577	3.7 (1.0, 9.5) for ≥75 mos exposure duration 2.9 (0.79, 7.5) for ≥19 mg/m^3 mean exposure	SIR. ICD 200 + 202. Exposure-group results presented only for males. Female results estimated and combined with male results assuming Poisson distribution (see text).
Morgan et al., 1998	0.81	0.1	6.49	High cumulative exp. score	−0.211	1.06	1.31 (0.28, 6.08) for med/high peak vs. low/no	Mortality RR. ICD 200 only. Adjusted for age and sex.
Raaschou-Nielsen et al., 2003	1.6	1.1	2.2	≥5 yrs in subcohort with expected higher exp. levels	0.470	0.183	None	SIR. ICD 200 + 202.

Table C-4. Selected RR estimates for lymphoma risk in highest TCE exposure groups (continued)

Study	RR	95% LCL	95% UCL	Exposure category	log RR	SE(log RR)	Alternate RR estimates	Comments
Radican et al., 2008	1.41	0.71	2.81	>25 unit-yr	0.337	0.350	Blair et al. (1998) 0.97 (0.42, 2.2) incidence RR	Mortality HR. ICD 200 + 202. Male and female results presented separately and combined (see text). Cox regression time variable = age; covariate = race. Referent group is workers with no chemical exposures.
Zhao et al., 2005	1.30	0.52	3.23	High exposure score	0.262	0.466	Incidence RR: 0.20 (0.03, 1.46)	Mortality RR. Results for all lymphohematopoietic cancer (ICD-9 200−208), not just 200 + 202. Males only; adjusted for age, SES, time since first employment. Mortality results reflect more exposed cases (6 in high-exposure group) than do incidence results (1 in high-exposure group).
Miligi et al., 2006	1.2	0.7	2.0	Med/high exposure intensity	0.182	0.268	1.0 (0.5, 2.6) for med/high intensity and >15-yr exp.	Incidence OR. NHL + CLL (see Section C.2.1.1).
Seidler et al., 2007	2.3	1.0	5.2	>35 ppm-yr	0.833	0.421	None	Incidence OR. Results for B-cell and T-cell NHL from personal communication (see Section C.2.1.1). Adjusted for smoking and alcohol consumption. Case-control pairs matched on sex, region, and age.
Siemiatycki 1991	0.8	0.2	3.3	Substantial	−0.223	0.719	None	Incidence OR. NHL. SE and 95% CI calculated from reported 90% CIs. Males only; adjusted for age, income, and cigarette smoking index.
Wang et al., 2009	2.2	0.9	5.4	Medium-high intensity	0.788	0.457	None	Incidence OR. "NHL" (various malignant lymphoma subtypes and mast cell tumors). Females only; adjusted for age, family history of lymphohematopoietic cancers, alcohol consumption, and race.

[a]Mean personal trichloroacetic acid in urine. 1 µmol/L = 0.1634 mg/L.

For Zhao et al. (2005), RR estimates were only reported for ICD-9 200−208 (all lymphohematopoietic cancers), and not for 200 + 202 alone. Given that other studies have not reported associations between leukemias and TCE exposure, combining all lymphohematopoietic cancers would dilute any lymphoma effect, and the Zhao results are expected to be an underestimate of any TCE effect on lymphoma alone. Zhao et al. (2005) present RR estimates for both incidence and mortality in the highest exposure group; however, the time frame for the incidence accrual is smaller than the time frame for mortality accrual and fewer incident cases (1) were obtained than deaths (6), so the mortality results were used for the primary analysis to reflect the better case ascertainment in the mortality data, and the incidence results were used in a sensitivity analysis.

Miligi et al. (2006) include CLLs in their NHL results, consistent with the current WHO/REAL classification. Miligi et al. (2006) report RR estimates for medium and high exposure intensity overall and by duration of exposure; however, there was incomplete information for the duration breakdowns (e.g., a case missing), so the RR estimate for med/high exposure intensity overall was used in the primary analysis, and the RR estimate for med/high exposure for >15 years was used in a sensitivity analysis. For Seidler et al. (2007), an adjusted OR for B-cell and T-cell NHL combined for the >35 ppm-years exposure category was kindly provided by Dr. Seidler (personal communication from Andreas Seidler, Bundesanstalt fur Arbeitsschutz u. Arbeitsmedizin, to Cheryl Scott, U.S. EPA, 13 November 2007). Wang et al. (2009) refer to their cases as "NHL" cases; however, according to the ICD-O classification system that they used, their cases are more specifically various particular subtypes of malignant lymphoma (9590-9642, 9690-9701) and mast cell tumors (9740-9750) (Morton et al., 2003).

C.2.2.2. *Results of Meta-Analyses*

Results from the meta-analyses that were conducted for lymphoma in the highest exposure groups are summarized at the bottom of Table C-3 and reported in more detail in Table C-5. The pooled RR estimate from the primary random effects meta-analysis of the 12 studies with results presented for exposure groups was 1.57 (95% CI: 1.27, 1.94) (see Figure C-4). No single study was overly influential; removal of individual studies resulted in RRp estimates that were all statistically significant (all with $p \leq 0.001$) and that ranged from 1.53 (with the removal of Seidler) to 1.65 (with the removal of Miligi). Similarly, the RRp estimate was not highly sensitive to alternate RR estimate selections. Use of the 7 alternate selections, individually, resulted in RRp estimates that were all statistically significant (all with $p < 0.001$) and all in the narrow range from 1.54 to 1.60 (see Table C-5). There was no observable heterogeneity across the 12 studies in either the primary analysis or any of the alternate RR analyses.

Table C-5. Summary of some meta-analysis results for TCE (highest exposure groups) and lymphoma

Analysis	Model	Combined RR estimate	95% LCL	95% UCL	Heterogeneity	Comments
Primary analysis	Random	1.57	1.27	1.94	None obs (fixed = random)	Statistical significance not dependent on single study.
Alternate RR selections[a]	Random	1.54	1.24	1.91	None obs	With Blair et al. (1998) incidence RR instead of Radican mortality HR.
	Random	1.55	1.24	1.92	None obs	With Zhao incidence.
	Random	1.57	1.27	1.94	None obs	With estimated female contribution for Axelson.
	Random	1.57	1.27	1.95	None obs	With Morgan peak.
	Random	1.58	1.28	1.96	None obs	With Hansen mean exposure.
	Random	1.60	1.28	2.00	None obs	With Miligi with >15 yrs.
	Random	1.60	1.30	1.98	None obs	With Hansen duration.

[a]Changing the primary analysis by one alternate RR estimate each time.

obs = observable.

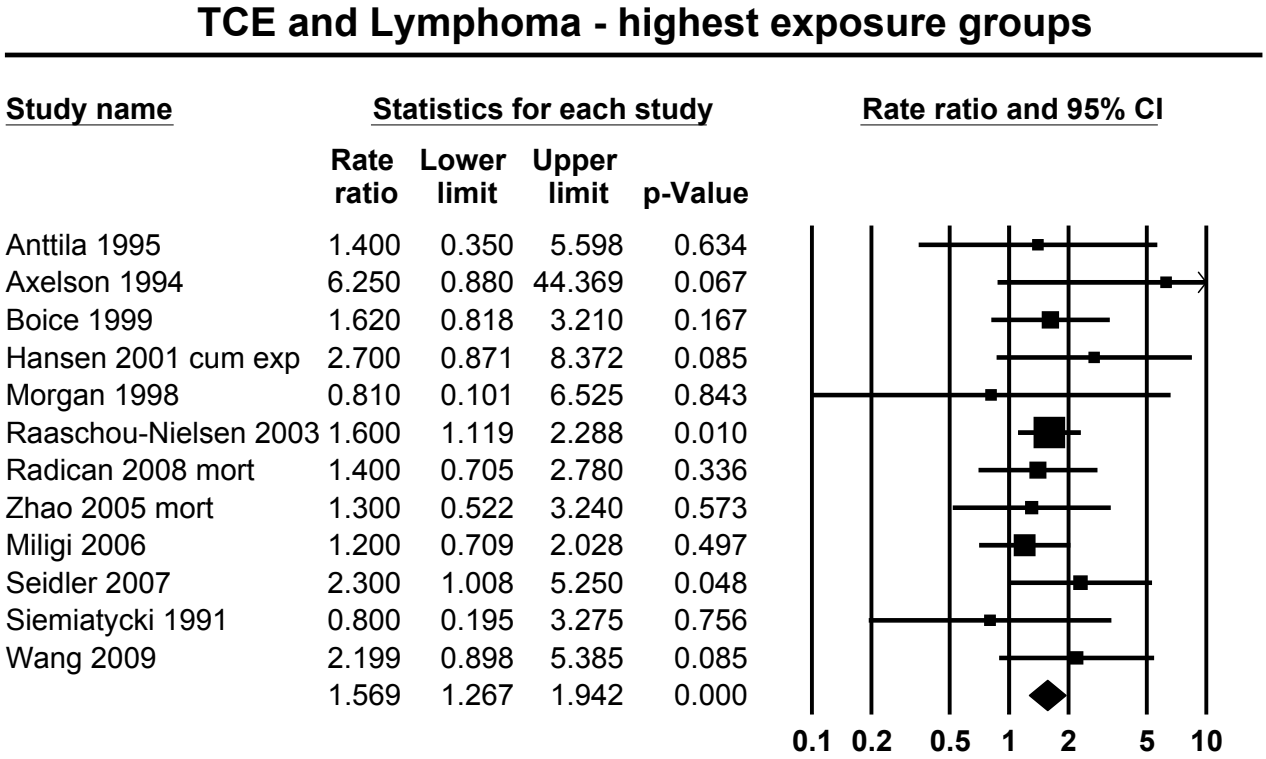

Figure C-4. Meta-analysis of lymphoma and TCE exposure—highest exposure groups. (The pooled estimate is in the bottom row. Symbol sizes reflect relative weights of the studies. The horizontal midpoint of the bottom diamond represents the pooled RR estimate, and the horizontal extremes depict the 95% CI limits.)

C.2.3. Discussion of Lymphoma Meta-Analysis Results

For the most part, the meta-analyses of the overall effect of TCE exposure on lymphoma suggest a small, statistically significant increase in risk. The pooled estimate from the primary random effects meta-analysis of the 16 studies was 1.23 (95% CI: 1.04, 1.44). This result was not overly influenced by any single study, nor was it overly sensitive to individual RR estimate selections. In terms of the statistical significance of the RRp estimate, the only alternate analysis (involving either a study removal or an alternate RR estimate) that did not yield a statistically significant RRp was the analysis in which the Zhao mortality RR estimate was substituted with the incidence estimate, resulting in an RRp estimate of 1.19 (1.00, 1.41); although, as noted

above, this substitution is considered clearly inferior to the Zhao mortality estimate that was used in the primary analysis. Thus, the finding of an increased risk of lymphoma associated with TCE exposure, though the increased risk is not large in magnitude, is fairly robust.

There is some evidence of potential publication bias in this data set; however, it is uncertain that this is actually publication bias rather than an association between SE and effect size resulting for some other reason, e.g., a difference in study populations or protocols in the smaller studies. Furthermore, if there is publication bias in this data set, it does not appear to account completely for the finding of an increased lymphoma risk.

Although there was some heterogeneity across the 16 studies, it was not statistically significant ($p = 0.10$). The I^2 value was 33%, suggesting low-to-moderate heterogeneity. Similarly, when subgroup analyses were done of cohort and case-control studies separately, there was some observable heterogeneity in each of the subgroups, but it was not statistically significant in either case. I^2 values were 10% for the cohort studies, suggesting low heterogeneity, and 33% for the case-control studies, suggesting low-to-moderate heterogeneity. In the subgroup analyses, the increased risk of lymphoma was strengthened in the cohort study analysis and virtually eliminated in the case-control study analysis, although the subgroup RRp estimates were not statistically significantly different under the random effects model. Study design itself is unlikely to be an underlying cause of heterogeneity and, to the extent that it may explain some of the differences across studies, is more probably a surrogate for some other difference(s) across studies that may be associated with study design. Furthermore, other potential sources of heterogeneity may be masked by the broad study design subgroupings. The true source(s) of heterogeneity across these studies is an uncertainty. As discussed above, further quantitative investigations of heterogeneity were ruled out because of database limitations. A qualitative discussion of some potential sources of heterogeneity follows.

Study differences in exposure assessment approach, exposure prevalence, average exposure intensity, and lymphoma classification are possible sources of heterogeneity. Many studies included TCE assignment from information on job and task exposures, e.g., a job-exposure matrix (JEM) (Siemiatycki, 1991; Morgan et al., 1998; Boice et al., 1999, 2006; Zhao et al., 2005; Miligi et al., 2006; Seidler et al., 2007; Radican et al., 2008; Wang et al., 2009), or from an exposure biomarker in either breath or urine (Axelson et al., 1994; Anttila et al., 1995; Hansen et al., 2001). Three case-control studies relied on self-reported exposure to TCE (Hardell et al., 1994; Nordstrom et al., 1998; Persson and Fredrikson, 1999). Misclassification is possible with all exposure assessment approaches. No information is available to judge the degree of possible misclassification bias associated with a particular exposure assessment approach; it is quite possible that in some cohort studies, in which past

exposure is inferred from various data sources, exposure misclassification may be as great as in population-based or hospital-based case-control studies. Approaches based upon JEMs can provide order-of-magnitude estimates that are useful for distinguishing groups of workers with large differences in exposure; however, smaller differences usually cannot be reliably distinguished (NRC, 2006). Biomonitoring can provide information on potential TCE exposure in an individual, but the biomarkers used aren't necessarily specific for TCE and they reflect only recent exposures. The lack of heterogeneity in the analysis of the highest exposure groups provides some evidence of exposure misclassification as a source of heterogeneity in the overall analysis.

General population studies have special problems in evaluating exposure, because the subjects could have worked in any job or setting that is present within the population (Copeland et al., 1977; Nelson et al., 1994; McGuire et al., 1998; 't Mannetje et al., 2002; NRC, 2006). Low exposure prevalence in the four population case-control studies (Siemiatycki, 1991; Miligi et al., 2006; Seidler et al., 2007; Wang et al., 2009) may be another source of heterogeneity. Prevalence of TCE exposure among cases in the case-control studies was low, ranging from 3% in Siemiatycki (1991) to 13% in Seidler et al. (2007) and Wang et al. (2009). However, prevalence of high TCE exposure in these case-control studies was even rarer—3% of all cases in Miligi et al. (2006) and Seidler et al. (2007), 2% in Wang et al. (2009), and less than 1% in Siemiatycki (1991). Low exposure prevalence, especially in the relatively large Miligi et al. (2006) and Seidler et al. (2007) case-control studies (see Figure C-1), may be one of the underlying characteristics differentiating the case-control and cohort studies and explaining some of the heterogeneity across the studies.

Study differences in lymphoma groupings and in lymphoma classification schemes are another potential source of heterogeneity in the meta-analysis. All studies included a broad but sometimes slightly different group of lymphosarcoma, reticulum-cell sarcoma, and other lymphoid tissue neoplasms, with the exception of the Nordstrom et al. (1998) case-control study, which examined hairy cell leukemia, now considered a lymphoma, and the Zhao et al. (2005) cohort study, which reported only results for *all* lymphohematopoietic cancers, including nonlymphoid types. Persson and Fredrikson (1999) do not identify the classification system for defining NHL, and Hardell et al. (1994) define NHL using the Rappaport classification system. Miligi et al. (2006) used an NCI classification system and considered chronic lymphocytic leukemias and NHLs together as lymphomas, while Seidler et al. (2007) used the REAL classification system, which reclassifies lymphocytic leukemias and NHLs as lymphomas of B-cell or T-cell origin. The cohort studies (except for Zhao et al.) and the case-control study of Siemiatycki (1991) have some consistency in coding NHL, with NHL defined as lymphosarcoma

and reticulum-cell sarcoma (ICD code 200) and other lymphoid tissue neoplasms (ICD 202) using the ICD Revisions 7, 8, or 9. Revisions 7 and 8 are essentially the same with respect to NHL; under Revision 9, the definition of NHL was broadened to include some neoplasms previously classified as Hodgkin's lymphomas (Banks, 1992). Wang et al. (2009) refer to their cases as "NHL" cases; however, according to the ICD-O classification system that they used, their cases are more specifically various particular subtypes of malignant lymphoma (9590-9642, 9690-9701) and mast cell tumors (9740-9750) (Morton et al., 2003).

Twelve of the 16 studies categorized results by exposure level. Different exposure metrics were used, and the purpose of combining results across the different highest exposure groups was not to estimate an RRp associated with some level of exposure, but rather to see the impacts of combining RR estimates that should be less affected by exposure misclassification. In other words, the highest exposure category is more likely to represent a greater differential TCE exposure compared to people in the referent group than the exposure differential for the overall (typically any vs. none) exposure comparison. Thus, if TCE exposure increases the risk of lymphoma, the effects should be more apparent in the highest exposure groups. Indeed, the RRp estimate from the primary meta-analysis of the highest exposure group results was 1.57 (95% CI: 1.27, 1.94), which is greater than the RRp estimate of 1.23 (95% CI: 1.04, 1.44) from the overall exposure analysis. This result for the highest exposure groups was not overly influenced by any single study, nor was it overly sensitive to individual RR estimate selections. Heterogeneity was not observed in any of the relevant analyses. The robustness of this finding lends substantial support to a conclusion that TCE exposure increases the risk of lymphoma.

C.3. META-ANALYSIS FOR KIDNEY CANCER

C.3.1. Overall Effect of TCE Exposure

C.3.1.1. *Selection of RR Estimates*

The selected RR estimates for kidney cancer associated with TCE exposure from the epidemiological studies are presented in Table C-6 for cohort studies and in Table C-7 for case-control studies. The majority of the cohort studies reported results for all kidney cancers, including cancers of the renal pelvis and ureter (i.e., ICD-7 180; ICD-8 and -9 189.0−189.2; ICD-10 C64−C66); whereas the majority of the case-control studies focused on renal cell carcinoma (RCC), which comprises roughly 85% of kidney cancers. Where both all kidney cancer and RCC were reported, the primary analysis used the results for RCC, because RCC and the other forms of kidney cancer are very different cancer types and it seemed preferable not to combine them; the results for all kidney cancers were then used in a sensitivity analysis. The preference for the RRC results alone is supported by the results in rodent cancer bioassays,

where TCE-associated rat kidney tumors are observed in the renal tubular cells (Section 4.3.5), and in metabolism studies, where the focus of studies for the GSH conjugation pathway (considered the primary metabolic pathway for kidney toxicity) is in renal cortical and tubular cells (Sections 3.3.3.2 and 4.3.6).

As for lymphoma, many of the studies provided RR estimates only for males and females combined, and we are not aware of any basis for a sex difference in the effects of TCE on kidney cancer risk; thus, wherever possible, RR estimates for males and females combined were used. Of the three larger (in terms of number of cases) studies that did provide results separately by sex, Dosemeci et al. (1999) suggest that there may be a sex difference for TCE exposure and RCC (OR = 1.04 [95% CI: 0.6, 1.7] in males and 1.96 [1.0, 4.0] in females), while Raaschou-Nielsen et al. (2003) report the same SIR (1.2) for both sexes and crude ORs calculated from data from the Pesch et al. (2000) study (provided in a personal communication from Baeta Pesch, Forschungsinstitut für Arbeitsmedizin (BGFA), to Cheryl Scott, U.S. EPA, 21 February 2008) are 1.28 for males and 1.23 for females. Radican et al. (2008) and Hansen et al. (2001) also present some results by sex, but both of these studies have too few cases to be informative about a sex difference for kidney cancer.

Most of the selections in Tables C-6 and C-7 should be self-evident, but some are discussed in more detail here, in the order the studies are presented in the tables. For Axelson et al. (1994), in which a small subcohort of females was studied but only results for the larger male subcohort were reported, the reported male-only results were used in the primary analysis; however, as for lymphoma, an attempt was made to estimate the female contribution to an overall RR estimate for both sexes and its impact on the meta-analysis. Axelson et al. (1994) reported neither the observed nor the expected number of kidney cancer cases for females. It was assumed that none were observed. To estimate the expected number, the expected number for males was multiplied by the ratio of female-to-male person-years in the study and by the ratio of female-to-male age-adjusted incidence rates for kidney cancer.[2] The male results and the estimated female contribution were then combined into an RR estimate for both sexes assuming a Poisson distribution, and this alternate RR estimate for the Axelson et al. (1994) study was used in a sensitivity analysis.

[2]Person-years for men and women $\leq$79 years were obtained from Axelson et al. (1994): 23516.5 and 3691.5, respectively. Lifetime age-adjusted incidence rates for cancer of the kidney and renal pelvis for men and women were obtained from the National Cancer Institute's 2000–2004 SEER-17 (Surveillance Epidemiology and End Results from 17 geographical locations) database (http://seer.cancer.gov/statfacts/html/kidrp.html): 17.8/100,000 and 8.8/100,000, respectively. The calculation for estimating the expected number of cases in females in the cohort assumes that the males and females have similar TCE exposures and that the relative distributions of age-related incidence risk for the males and females in the cohort are adequately represented by the ratios of person-years and lifetime incidence rates used in the calculation.

Table C-6. Selected RR estimates for kidney cancer associated with TCE exposure (overall effect) from cohort studies

Study	RR	95% LCL	95% UCL	RR type	log RR	SE(log RR)	Alternate RR estimates	Comments
Anttila et al., 1995	0.87	0.32	1.89	SIR	−0.139	0.408	none	ICD-7 180.
Axelson et al., 1994	1.16	0.42	2.52	SIR	0.148	0.408	1.07 (0.39, 2.33) with estimated female contribution to SIR added (see text)	ICD-7 180. Results reported for males only, but there was a small female component to the cohort.
Boice et al., 1999	0.99	0.4	2.04	SMR	−0.010	0.378	None	ICD-9 189.0−189.2. For potential routine exposure. Results for any potential exposure not reported.
Greenland et al., 1994	0.99	0.30	3.32	OR	−0.010	0.613	None	Nested case-control study. ICD-8 codes not specified, presumably all of 189.
Hansen et al., 2001	1.1	0.3	2.8	SIR	0.095	0.500	None	ICD-7 180. Male and female results reported separately; combined assuming Poisson distribution.
Morgan et al., 1998	1.14	0.51	2.58	Mortality RR	0.134	0.415	Published SMR 1.32 (0.57, 2.6)	ICD-9 189.0−189.2. Unpublished RR, adjusted for age and sex (see text).
Raaschou-Nielsen et al., 2003	1.20	0.94	1.50	SIR	0.182	0.199	1.20 (0.98, 1.46) for ICD-7 180	RCC.
Radican et al., 2008	1.18	0.47	2.94	Mortality HR	0.166	0.468	None	ICD-8, -9 189.0, ICD-10 C64. Time variable = age; covariates = sex and race. Referent group is workers with no chemical exposures.
Zhao et al., 2005	1.7	0.38	7.9	Mortality RR	0.542	0.775	Incidence RR: 2.0 (0.47, 8.2) Mortality RR no lag: 0.89 (0.22, 3.6) Incidence RR no lag : 2.1 (0.56, 8.1) Boice (2006) SMR: 2.22 (0.89, 4.57)	ICD-9 189. Males only. Adjusted for age, SES, time since first employment, exposure to other carcinogens. 20-yr lag. Mortality results reflect same number exposed cases (10 with no lag) as do incidence results, so no reason to prefer mortality results, but they are used in primary analysis to avoid appearance of "cherry-picking." Overall RR estimated by combining across exposure groups (see text). Boice (2006) cohort overlaps Zhao cohort; just 7 exposed deaths.

Table C-7. Selected RR estimates for renal cell carcinoma associated with TCE exposure from case-control studies[a]

Study	RR estimate	95% LCL	95% UCL	log RR	SE(log RR)	Alternate RR estimates	Comments
Brüning et al., 2003	2.47	1.36	4.49	0.904	0.305	1.80 (1.01, 3.20) for longest job held in industry with TCE exposure	Self-assessed exposure. Adjusted for age, sex, and smoking.
Charbotel et al., 2006	1.88	0.89	3.98	0.631	0.382	1.64 (0.95, 2.84) for full study	Subgroup with good level of confidence about exp assessment. Matched on sex, age. Adjusted for smoking, body mass index.
Dosemeci et al., 1999	1.30	0.9	1.9	0.262	0.191		Adjusted for age, sex, smoking, hypertension and/or use of diuretics and/or anti-hypertension drugs, body mass index.
Pesch et al., 2000	1.24	--[b]	--[b]	0.215	0.094	1.13 with German JEM	With JTEM (job task exposure matrix). Crude OR calculated from data provided in personal communication (see text).
Siemiatycki 1991	0.8	0.3	2.2	−0.223	0.524		"Kidney cancer." SE and 95% CI calculated from reported 90% CIs. Males only; adjusted for age, income, and cigarette smoking index.

[a]The RR estimates are all ORs for incident cases.
[b]Not calculated.

For Boice et al. (1999), only results for "potential routine exposure" were reported for kidney cancer. This is our preferred TCE exposure definition for the Boice study, because it was considered to have less exposure misclassification than "any potential exposure;" however, since the results for the latter definition were not presented, they could not be used in a sensitivity analysis, as was done for lymphoma. Boice et al. (1999) report in general that the SMRs for workers with any potential exposure "were similar to those for workers with daily potential exposure." In their published paper, Morgan et al. (1998) present only SMRs for overall TCE exposure, although the results from internal analyses are presented for exposure subgroups. RR estimates for overall TCE exposure from the internal analyses of the Morgan et al. (1998) cohort data were available from an unpublished report (Environmental Health Strategies, 1997); from these, the RR estimate from the Cox model which included age and sex was selected, because those are the variables deemed to be important in the published paper. The internal analysis RR estimate was preferred for the primary analysis, and the published SMR result was used in a sensitivity analysis. Raaschou-Nielsen et al. (2003) reported results for RCC and renal pelvis/ureter separately. As discussed above, RCC estimates were used in the primary analysis, and the results for both kidney cancer categories were combined (across sexes as well), assuming a Poisson distribution, and used in a sensitivity analysis. For Radican et al. (2008), the Cox model hazard ratio (HR) from the 2000 follow-up was used. In the Radican et al. (2008) Cox regressions, age was the time variable, and sex and race were covariates. It should also be noted that the referent group is composed of workers with no chemical exposures, not just no exposure to TCE.

For Zhao et al. (2005), no results for an overall TCE effect are reported. We were unable to obtain any overall estimates from the study authors, so, as a best estimate, as was done for lymphoma, the results across the "medium" and "high" exposure groups were combined, under assumptions of group independence, even though the exposure groups are not independent (the "low" exposure group was the referent group in both cases). Unlike for lymphoma, adjustment for exposure to other carcinogens made a considerable difference, so Zhao et al. (2005) also present kidney results with this additional adjustment, with and without a 20-year lag. Estimates of RR with this additional adjustment were selected over those without. In addition, a 20-year lag seemed reasonable for kidney cancer, so the lagged estimates were preferred to the unlagged; unlagged estimates were used in sensitivity analyses. Zhao et al. (2005) present RR estimates for both incidence and mortality. Unlike for lymphoma, the number of exposed incident cases (10 with no lag) was identical to the number of deaths, so there was no reason to prefer the mortality results over the incidence results. (In fact, there were more exposed incident cases [10 vs. 7] after lagging.) However, the mortality results, which yield a lower RR estimate, were selected

for the primary analysis to avoid any appearance of "cherry-picking," and incidence RR estimates were used in sensitivity analyses. A sensitivity analysis was also done using results from Boice et al. (2006) in place of the Zhao et al. (2005) RR estimate. The cohorts for these studies overlap, so they are not independent studies and should not be included in the meta-analysis concurrently. Boice et al. (2006) report results for an overall TCE effect for kidney cancer; however, the results are SMR estimates rather than internal comparisons and are based on fewer exposed deaths (7), so either Zhao et al. (2005) estimate is preferred over the Boice et al. (2006) estimate.

Regarding the case-control studies, for Brüning et al. (2003), the results based on self-assessed exposure were preferred because, although TCE exposure was probably under ascertained with this measure, there were greater concerns about the result based on the alternate measure reported—longest-held job in an industry with TCE exposure. Even though this study was conducted in the Arnsberg region of Germany, an area with high prevalence of exposure to TCE, the exposure prevalence in both cases (87%) and controls (79%) seemed inordinately high, and this for not just any job in an industry with TCE exposure, but for the longest-held job. Furthermore, Table V of Brüning et al., which presents this result, states that the result is for longest-held job in industries with TCE *or tetrachloroethylene* exposure. Additionally, some of the industries with exposure to TCE presented in Table V have many jobs that would not entail TCE exposure (e.g., white-collar workers), so the assessment based on industry alone likely has substantial misclassification. Both of these—inclusion of tetrachloroethylene and exposure assessment by industry—could result in overstating TCE exposure prevalence. Results based on the longest-held-job measure were used in a sensitivity analysis.

For Charbotel et al. (2006), results from the analysis that considered "only job periods with a good level of confidence for TCE exposure assessment" (Table 7 of Charbotel et al., 2006) were preferred, as these estimates would presumably be less influenced by exposure misclassification. Estimates from the full study analysis were used in a sensitivity analysis. For Pesch et al. (2000), TCE results were presented for 2 different exposure assessments. Estimates using the job-task-exposure-matrix (JTEM) approach were preferred because they seemed to represent a more comprehensive exposure assessment (see Appendix B, Section II-4); estimates based on the JEM approach were used in a sensitivity analysis. Furthermore, results were presented only by exposure category, with no overall RR estimate reported. Case and control numbers for the different exposure categories were kindly provided by Dr. Pesch (personal communication from Baete Pesch, BGFA, to Cheryl Scott, U.S. EPA, 21 February 2008), and we calculated crude overall ORs for males and females combined for each exposure assessment approach.

C.3.1.2. *Results of Meta-Analyses*

Results from some of the meta-analyses that were conducted on the epidemiological studies of TCE and kidney cancer are summarized in Table C-8. The pooled estimate from the primary random effects meta-analysis of the 14 studies was 1.25 (95% CI: 1.11, 1.41) (see Figure C-5). As shown in Figure C-5, the analysis was dominated by 2 (contributing almost 70% of the weight) or 3 (almost 80% of the weight) large studies. No single study was overly influential; removal of individual studies resulted in RRp estimates that were all statistically significant (all with $p < 0.005$) and that ranged from 1.22 (with the removal of Brüning) to 1.27 (with the removal of Raaschou-Nielsen).

Similarly, the RRp estimate was not highly sensitive to alternate RR estimate selections. Use of the 10 alternate selections, individually, resulted in RRp estimates that were all statistically significant (all with $p < 0.002$) and that ranged from 1.19 to 1.27 (see Table C-8). In fact, as can be seen in Table C-8, all but one of the alternates had negligible impact. The Zhao, Axelson, Brüning, and Charbotel original values and alternate selections were associated with very little weight and, thus, have little influence in the RRp. The Raaschou-Nielsen value carried more weight, but the alternate RR estimate was identical to the original, although with a narrower CI, and so did not alter the RRp. Only the Pesch alternate (with the JEM exposure assessment approach instead of the JTEM approach) had much impact, resulting in an RRp estimate of 1.19 (95% CI: 1.07, 1.32). As noted above, the JTEM approach is preferred. The JEM approach takes jobs into account but not tasks; thus, it is expected to have greater potential for exposure misclassification. Indeed, a comparison of exposure prevalences for the two approaches suggests that the JEM approach is less discriminating about exposure; 42% of cases were defined as TCE-exposed under the JEM approach, but only 18% of cases were exposed under the JTEM approach.

Table C-8. Summary of some meta-analysis results for TCE (overall) and kidney cancer

Analysis	# of studies	Model	Combined RR estimate	95% LCL	95% UCL	Heterogeneity	Comments
All studies	14	Random	1.25	1.11	1.41	None obs	Statistical significance not dependent on single study. No apparent publication bias.
		Fixed	1.25	1.11	1.41		
Cohort	9	Random	1.16	0.96	1.40	None obs	Not significant difference between CC and cohort studies (p = 0.23).
		Fixed	1.16	0.96	1.40		Not significant difference between CC and cohort studies (p = 0.29).
Case-control	5	Random	1.41	1.08	1.83	Not significant (p = 0.17)	
		Fixed	1.32	1.13	1.54		
Alternate RR selections[a]	14	Random	1.25	1.11	1.40−1.41	None obs	With 3 different alternates from Zhao (see Table C-6).
	14	Random	1.27	1.13	1.43	None obs	With Boice (2006) study rather than Zhao
	14	Random	1.25	1.11	1.41	None obs	With estimated female contribution to Axelson.
	14	Random	1.26	1.11	1.41	None obs	With Morgan published SMR.
	14	Random	1.25	1.11	1.40	None obs	With Raaschou-Nielsen all kidney cancer.
	14	Random	1.24	1.10	1.39	None obs	With Brüning longest job held in industry with TCE.
	14	Random	1.25	1.11	1.41	None obs	With Charbotel full study
	14	Random	1.19	1.07	1.32	None obs	With Pesch JEM.
Highest exposure groups	9	Random	1.59	1.26	2.01	None obs	
	12	Random	1.53	1.23	1.91	None obs	Using RR = 1 for Anttila, Axelson, and Hansen (see text). See Table C-10 for alternate RR selection results.

[a]Changing the primary analysis by one alternate RR each time.

obs = observable.

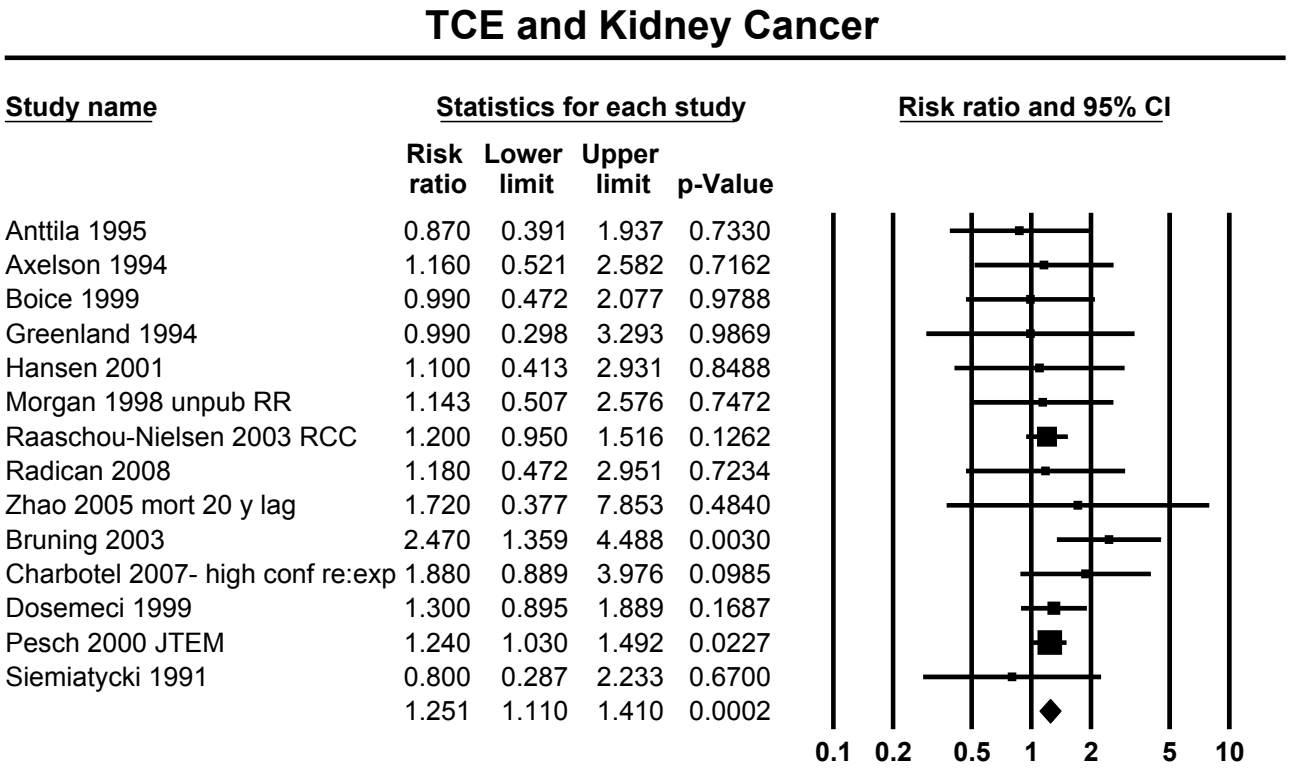

Figure C-5. Meta-analysis of kidney cancer and overall TCE exposure. The pooled estimate is in the bottom row. Symbol sizes reflect relative weights of the studies. The horizontal midpoint of the bottom diamond represents the pooled RR estimate and the horizontal extremes depict the 95% CI limits.

There was no apparent heterogeneity across the 14 studies, i.e., the random effects model and the fixed effect model gave the same results. Nonetheless, subgroup analyses were done examining the cohort and case-control studies separately. With the random effects model (and tau-squared not pooled across subgroups), the resulting RRp estimates were 1.16 (95% CI: 0.96, 1.40) for the cohort studies and 1.41 (1.08, 1.83) for the case-control studies. There was heterogeneity in the case-control subgroup, but it was not statistically significant and the I^2 value

of 38% suggests that the extent of the heterogeneity in this subgroup was low-to-moderate. Nor was the difference between the RRp estimates for the cohort and case-control subgroups statistically significant under either the random effects model or the fixed effect model. Further quantitative investigations of heterogeneity were not pursued because of database limitations and, in any event, there is no evidence for heterogeneity of study results in this database. A qualitative discussion of some potential sources of heterogeneity across studies is nonetheless included in Section C.3.3.

As discussed in Section C.1, publication bias was examined in several different ways. The funnel plot in Figure C-6 shows little relationship between RR estimate and study size, and, indeed, none of the other tests performed found any evidence of publication bias. Duval and Tweedie's trim-and-fill procedure, for example, determined that no studies were missing from the funnel plot, i.e., there was no asymmetry to counterbalance. Similarly, the results of a cumulative meta-analysis, incorporating studies with increasing SE one at a time, shows no evidence of a trend of increasing effect size with addition of the less precise studies. Including the 3 most precise studies, reflecting 78% of the weight, the RRp goes from 1.24 to 1.22 to 1.23. The addition of the Brüning study brings the RRp to 1.32 and the weight to 82%. After the addition of the next 5 studies, the RRp stabilizes at about 1.26, and further addition of the 5 least precise studies has little impact.

C.3.2. Kidney Cancer Effect in the Highest Exposure Groups

C.3.2.1. *Selection of RR Estimates*

The selected RR estimates for kidney cancer in the highest TCE exposure categories, for studies that provided such estimates, are presented in Table C-9. Five of the 9 cohort studies and 4 of the 5 case-control studies reported kidney cancer risk estimates categorized by exposure level. As in Section C.3.1.1 for the overall risk estimates, estimates for RCC were preferentially selected when presented, and, wherever possible, RR estimates for males and females combined were used.

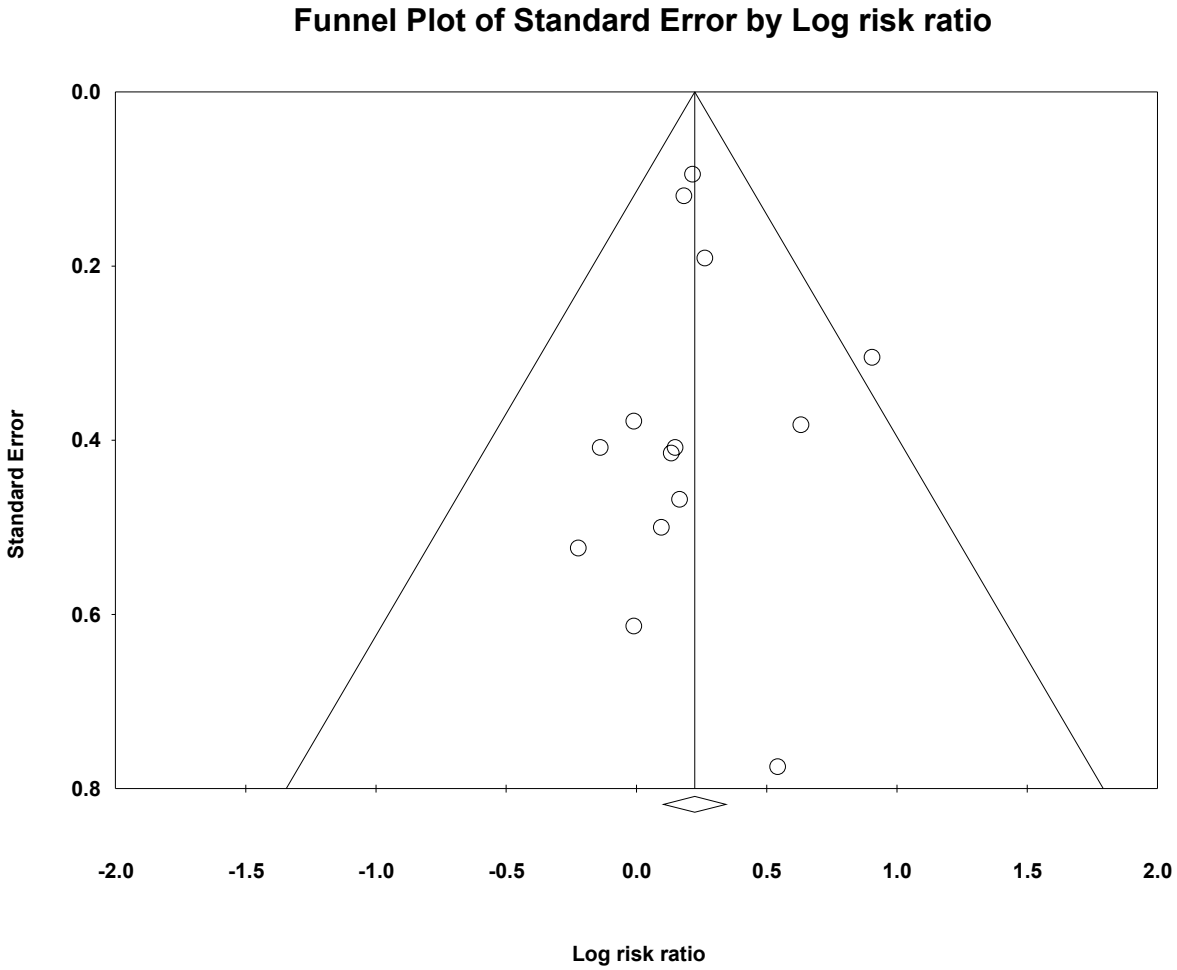

Figure C-6. Funnel plot of SE by log RR estimate for TCE and kidney cancer studies

Table C-9. Selected RR estimates for kidney cancer risk in highest TCE exposure groups

Study	RR	95% LCL	95% UCL	Exposure category	log RR	SE(log RR)	Alternate RR estimates	Comments
Anttila et al., 1995				100+ µmol/L U-TCA [a]			1.0 assumed	Reported high exposure group results for some cancer sites but not kidney.
Axelson et al., 1994				≥2 yr exposure and 100+ mg/L U-TCA			1.0 assumed	Reported high exposure group results for some cancer sites but not kidney.
Boice et al., 1999	0.69	0.22	2.12	≥5 yr exp	−0.371	0.578	None	Mortality RR. ICD-9 189.0−189.2. For potential routine or intermittent exposure. adjusted for date of birth, dates 1st and last employed, race, and sex. Referent group is workers not exposed to any solvent.
Hansen et al., 2001				≥1080 mos × mg/m^3			1.0 assumed	Reported high exposure group results for some cancer sites but not kidney.
Morgan et al., 1998	1.59	0.68	3.71	High cumulative exposure score	0.464	0.433	1.89 (0.85, 4.23) for med/high peak vs. low/no	Mortality RR. ICD-9 189.0−189.2. Adjusted for age and sex.
Raaschou-Nielsen et al., 2003	1.7	1.1	2.4	≥5 yrs in subcohort with expected higher exposure levels	0.531	0.183	1.4 (0.99, 1.9) ICD-7 180 ≥5 yrs in total cohort	SIR. RCC.
Radican et al., 2008	1.11	0.35	3.49	>25 unit-yr	0.104	0.582	Blair et al. (1998) incidence RR 0.9 (0.3, 3.2)	Mortality HR. ICD-8, -9 189.0, ICD-10 C64. Male and female results presented separately and combined (see text). Referent group is workers with no chemical exposures.

Table C-9. Selected RR estimates for kidney cancer risk in highest TCE exposure groups (continued)

Study	RR	95% LCL	95% UCL	Exposure category	log RR	SE(log RR)	Alternate RR estimates	Comments
Zhao et al., 2005	7.40	0.47	116	High exposure score	2.00	1.41	Mortality RR: 1.82 (0.09, 38.6) Incidence RR no lag: 7.71 (0.65, 91.4) Mortality RR no lag: 0.96 (0.09, 9.91) Boice 2006 mortality RR: 2.12 (0.63, 7.11) for ≥5 yrs as test stand mechanic; 3.13 (0.74,13.2) for ≥4 test-yr engine flush	Incidence RR. ICD-9 189. Males only. Adjusted for age, SES, time since first employment, exposure to other carcinogens. 20-yr lag. Incidence results reflect more exposed cases (4 with no lag) than do mortality results (3), so they are used in primary analysis.
Brüning et al., 2003	2.69	0.84	8.66	≥20 yrs self-assessed exposure	0.990	0.595	None	Incidence OR. RCC. Adjusted for age, sex, and smoking.
Charbotel et al., 2006	3.34	1.27	8.74	High cumulative dose	1.21	0.492	3.80 (1.27, 11.40) for high cum + peaks 1.96 (0.71, 5.37) for high cum + peaks in full study 2.63 (0.79, 8.83) for high cum in full study	Incidence OR. RCC. In subgroup with good level of confidence for TCE exposure. Adjusted for smoking and body mass index. Matched on sex and age. Alternate full study estimates were additionally adjusted for exposure to cutting fluids and other petroleum oils.

Table C-9. Selected RR estimates for kidney cancer risk in highest TCE exposure groups (continued)

Study	RR	95% LCL	95% UCL	Exposure category	log RR	SE(log RR)	Alternate RR estimates	Comments
Pesch et al., 2000	1.4	0.9	2.1	Substantial	0.336	0.219	1.2 (0.9, 1.7) for JEM	Incidence OR. RCC. JTEM approach. Adjusted for age, study center, and smoking. Sexes combined.
Siemiatycki 1991	0.8	0.2	3.4	Substantial	−0.233	0.736	none	Incidence OR. Kidney cancer. SE and 95% CI calculated from reported 90% CIs. Males only; adjusted for age, income, and cigarette smoking index.

[a]Mean personal trichloroacetic acid in urine. 1 µmol/L = 0.1634 mg/L.

Three of the 9 cohort studies (Anttila et al., 1995; Axelson et al., 1994; Hansen et al., 2001) did not report kidney cancer risk estimates categorized by exposure level even though these same studies reported such estimates for selected other cancer sites. To address this reporting bias, attempts were made to obtain the results from the primary investigators, and, failing that, an alternate analysis was performed in which null estimates (RR = 1.0) were included for all 3 studies. This alternate analysis was then used as the main analysis, e.g., the basis of comparison for the sensitivity analyses. For the SE (of the logRR) estimates for these null estimates, SE estimates from other sites for which highest-exposure-group results were available were used. For Anttila et al. (1995), the SE estimate for liver cancer in the highest exposure group was used, because liver cancer and kidney cancer had similar numbers of cases in the overall study (5 and 6, respectively). For Axelson et al. (1994), the SE estimate for NHL in the highest exposure group was used, because NHL and kidney cancer had similar numbers of cases in the overall study (5 and 6, respectively). For Hansen et al. (2001), the SE estimate for NHL in the highest exposure group was used, because NHL was the only cancer site of interest in this assessment for which highest-exposure-group results were available.

For Boice et al. (1999), only results for workers with "any potential exposure" (rather than "potential routine exposure") were presented by exposure category, and the referent group is workers not exposed to any solvent. For Morgan et al. (1998), the primary analysis used results for the cumulative exposure metric, and a sensitivity analysis was done with the results for the peak exposure metric.

For Radican et al. (2008), it should be noted that the referent group is workers with no chemical exposures, not just no TCE exposure. In addition, exposure group results were reported separately for males and females and were combined for this assessment using inverse-variance weighting, as in a fixed effect meta-analysis. Radican et al. (2008) present only mortality HR estimates by exposure group; however, in an earlier follow-up of this same cohort, Blair et al. (1998) present both incidence and mortality RR estimates by exposure group. The mortality RR estimate based on the more recent follow-up of Radican et al. (2008) (6 deaths in the highest exposure group) was used in the primary analysis, while the incidence RR estimate based on similarly combined results from Blair et al. (1998) (4 cases) was used as an alternate estimate in a sensitivity analysis.

Zhao et al. (2005) present kidney cancer RR estimates adjusted for exposure to other carcinogens, because, unlike for lymphoma, this adjustment made a considerable difference. Estimates of RR with this additional adjustment were selected over those without. Furthermore, the kidney results were presented with and without a 20-year lag. A 20-year lag seemed reasonable for kidney cancer, so the lagged estimates were preferred to the unlagged; unlagged

estimates were used in sensitivity analyses. In addition, the incidence results reflect more cases (4 with no lag) in the highest exposure group than do the mortality results (3), so the incidence result (with the 20-year lag) was used for the primary analysis, and the unlagged incidence result and the mortality results were used in a sensitivity analysis. Sensitivity analyses were also done using results from Boice et al. (2006) in place of the Zhao et al. (2005) RR estimate. The cohorts for these studies overlap, so they are not independent studies. Boice et al. (2006) report mortality RR estimates for kidney cancer by years worked as a test stand mechanic, a job with potential TCE exposure, and by a measure that weighted years with potential exposure from engine flushing by the number of flushes each year. No results were presented for a third metric, years worked with potential exposure to any TCE, because the Cox proportional hazards model did not converge. The Boice et al. (2006) estimates are adjusted for years of birth and hire and for hydrazine exposure.

For Charbotel et al. (2006), results from the analysis that considered "only job periods with a good level of confidence for TCE exposure assessment" (Table 7 of Charbotel et al., 2006) were preferred, as these estimates would presumably be less influenced by exposure misclassification. Estimates from the full study analysis, additionally adjusted for exposure to cutting fluids and other petroleum oils, were used in a sensitivity analysis. Additionally, the high cumulative dose results were preferred, but the results for high cumulative dose + peaks were included in sensitivity analyses. For Pesch et al. (2000), TCE results were presented for two different exposure assessments. As discussed above, estimates using the JTEM approach were preferred because they seemed to represent a more comprehensive exposure assessment; estimates based on the JEM approach were used in a sensitivity analysis.

C.3.2.2. *Results of Meta-Analyses*

Results from the meta-analyses that were conducted for kidney cancer in the highest exposure groups are summarized at the bottom of Table C-8 and reported in more detail in Table C-10. The pooled RR estimate from the random effects meta-analysis of the 9 studies with results presented for exposure groups was 1.59 (95% CI: 1.26, 2.01) (see Figure C-7). The RRp estimate from the primary random effects meta-analysis with null RR estimates (i.e., 1.0) included for Anttila, Axelson, and Hansen to address reporting bias (see above) was 1.53 (1.23, 1.91) (see Figure C-8). The inclusion of these 3 additional studies contributed just under 8% of the total weight. As with the overall kidney cancer meta-analyses, the meta-analyses of the highest-exposure groups were dominated by 2 studies (Raaschou-Nielsen and Pesch), which provided about 66% of the weight. No single study was overly influential; removal of individual studies resulted in RRp estimates that were all statistically significant (all with $p < 0.02$) and that

ranged from 1.43 (with the removal of Raaschou-Nielsen) to 1.58 (with the removal of Boice [1999] or Pesch).

Similarly, the RRp estimate was not highly sensitive to alternate RR estimate selections. Use of the 12 alternate selections, individually, resulted in RRp estimates that were all statistically significant (all with $p < 0.002$) and that ranged from 1.42 to 1.55, with all but 2 of the alternate selections yielding RRp estimates in the narrow range of 1.49−1.55 (see Table C-10). The lowest RRp estimates, 1.42 in both cases, were obtained when the alternate selections involved the 2 large studies. One of the alternate selections was for Raaschou-Nielsen, with a highest-exposure group estimate for all kidney cancer in the total cohort, rather than RCC in the subcohort expected to have higher exposure levels. The latter value is strongly preferred because, as discussed above, the subcohort is likely to have less exposure misclassification. Furthermore, RCC is very different from other types of kidney cancer, and TCE, if an etiological factor, may not be etiologically associated with all kidney cancers, so using the broad category may dilute a true association with RCC, if one exists. The other alternate selection with a considerable impact on the RRp estimate was for Pesch, with the highest exposure group result based on the JEM exposure assessment approach, rather than the JTEM approach. As discussed above, the JTEM approach is preferred because it seemed to be a more comprehensive and discriminating approach, taking actual job tasks into account, rather than just larger job categories. Thus, although results with these alternate selections are presented for comprehensiveness and transparency, the primary analysis is believed to reflect better the potential association between kidney cancer (in particular, RCC) and TCE exposure.

There was no observable heterogeneity across the studies for any of the meta-analyses conducted with the highest-exposure groups, including those in which RR values for Anttila, Axelson, and Hansen were assumed. No subgroup analyses (e.g., cohort vs. case-control studies) were done with the highest exposure group results.

Table C-10. Summary of some meta-analysis results for TCE (highest exposure groups) and kidney cancer

Analysis	Model	Combined RR estimate	95% LCL	95% UCL	Heterogeneity	Comments
Analysis based on reported results	Random	1.59	1.26	2.01	None obs (fixed = random)	
Primary analysis	Random	**1.53**	**1.23**	**1.91**	None obs	Includes assumed values for Anttila, Axelson, and Hansen (see text). Statistical significance not dependent on single study.
Alternate RR selections[a]	Random	1.52	1.22	1.90	None obs	With Blair et al. (1998) incidence RR instead of Radican mortality HR.
	Random	1.55	1.24	1.94	None obs	With Morgan peak metric.
	Random	1.42	1.15	1.75	None obs	With Raaschou-Nielsen for all kidney cancer ≥5 yrs in total cohort.
	Random	1.51−1.54	1.21−1.23	1.89−1.92	None obs	With Zhao incidence unlagged and mortality with and without lag.
	Random	1.53−1.54	1.23−1.24	1.91−1.92	None obs	With Boice (2006) alternates for Zhao (see text).
	Random	1.49−1.52	1.19−1.22	1.86−1.91	None obs	With Charbotel high cumulative dose + peaks in subgroup; and high cumulative dose and high cumulative dose + peaks in full study additionally adjusted for exposure to cutting fluids and other petroleum oils..
	Random	1.42	1.16	1.74	None obs	With Pesch JEM.

[a]Changing the primary analysis by one alternate RR each time.

obs = observable.

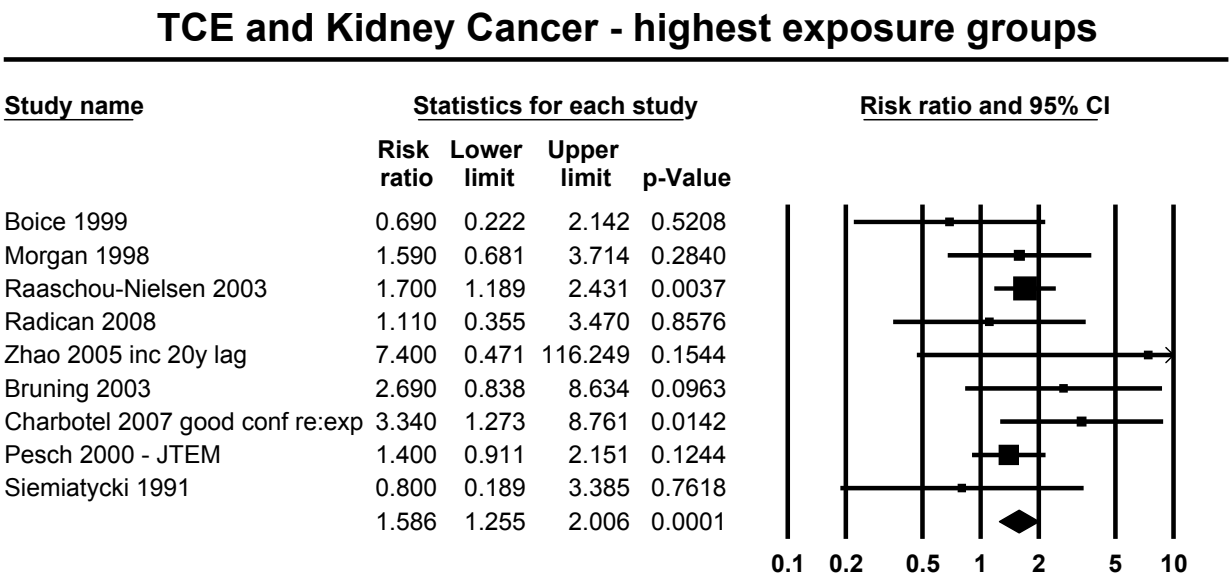

Figure C-7. Meta-analysis of kidney cancer and TCE exposure—highest exposure groups. The pooled estimate is in the bottom row. Symbol sizes reflect relative weights of the studies. The horizontal midpoint of the bottom diamond represents the pooled RR estimate and the horizontal extremes depict the 95% CI limits.

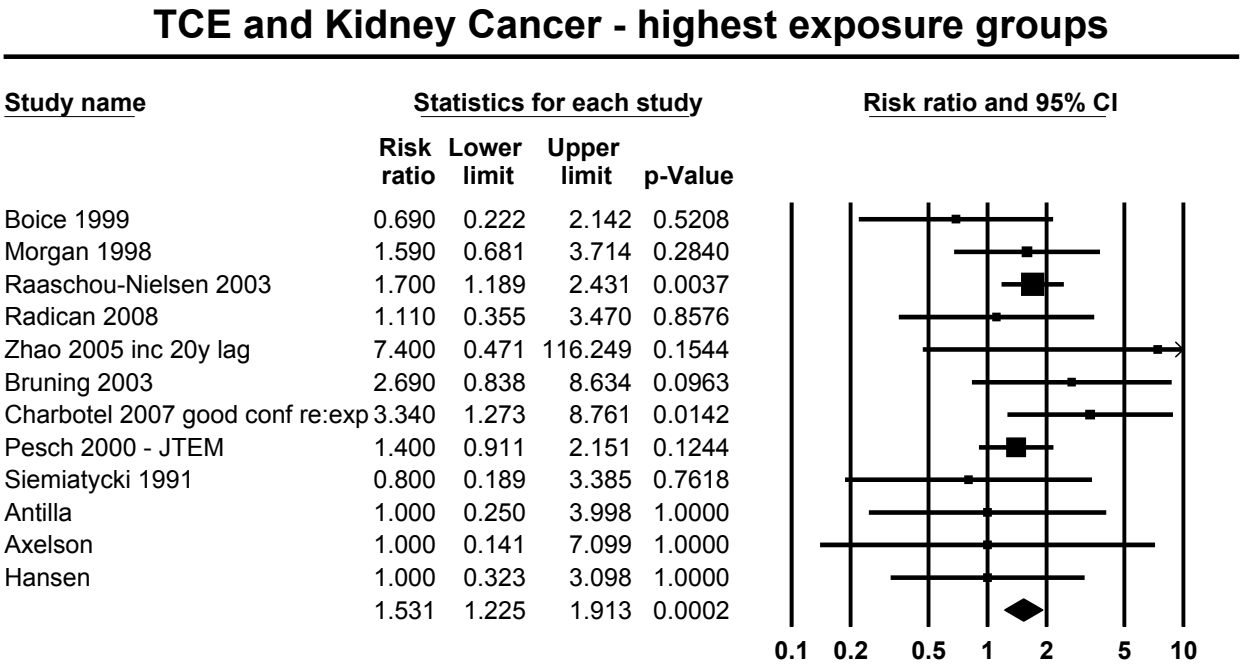

Figure C-8. Meta-analysis of kidney cancer and TCE exposure—highest exposure groups, with assumed null RR estimates for Anttila, Axelson, and Hansen (see text).

C.3.3. Discussion of Kidney Cancer Meta-Analysis Results

For the most part, the meta-analyses of the overall effect of TCE exposure on kidney cancer suggest a small, statistically significant increase in risk. The pooled estimate from the primary random effects meta-analysis of the 14 studies was 1.25 (95% CI: 1.11, 1.41). Although the analysis was dominated by 2−3 large studies that contribute 70−80% of the weight, the pooled estimate was not overly influenced by any single study, nor was it overly sensitive to individual RR estimate selections. The largest downward impacts were from the removal of the Brüning study, resulting in an RRp estimate of 1.22 (95% CI: 1.08, 1.37), and from the substitution of the Pesch JTEM RR estimate with the RR estimate based on the JEM approach, resulting in an RRp estimate of 1.19 (1.07, 1.32). Thus, the finding of an increased risk of

kidney cancer associated with TCE exposure is robust. Furthermore, there is no evidence of publication bias in this data set.

In addition, there was no heterogeneity observed across the results of the 14 studies. When subgroup analyses were done of cohort and case-control studies separately, there was some observable heterogeneity among the case-control studies, but it was not statistically significant ($p = 0.17$) and the I^2 value of 38% suggested the extent of the heterogeneity was low-to-moderate. The increased risk of kidney cancer was strengthened in the case-control study analysis and weakened in the cohort study analysis, but the difference between the 2 RRp estimates was not statistically significant. One difference between the case-control and cohort studies is that the case-control studies were of RCC and almost all of the cohort studies were of all kidney cancers, including renal pelvis. As discussed above, RCC is very different from other types of kidney cancer, and TCE, if an etiological factor, may not be etiologically associated with all kidney cancers, so using the broad category may dilute a true association with RCC, if one exists.

With respect to the nonsignificant heterogeneity in the 5 case-control studies, these studies differ in TCE exposure potential to the underlying population from which case and control subjects were identified, and this may be a source of some heterogeneity. Prevalence of exposure to TCE among cases in these studies was 27% in Charbotel et al. (2006) (for high-level-of-confidence jobs), 18% in Brüning et al. (2003) (for self-assessed exposure), 18% in Pesch et al. (2000), 13% in Dosemeci et al. (1999) and 1% in Siemiatycki (1991). Both Brüning et al. (2003) and Charbotel et al. (2006) are studies designed specifically to assess RCC and TCE exposure. These studies were carried out in geographical areas with both a high prevalence and a high degree of TCE exposure. Some information is provided in these and accompanying papers to describe the nature of exposure, making it possible to estimate the order of magnitude of exposure, even though there were no direct measurements (Cherrie et al., 2001; Brüning et al., 2003; Fevotte et al., 2006). The Charbotel et al. (2006) study was carried out in the Arve Valley region in France, where TCE exposure was through metal-degreasing activity in small shops involved in the manufacturing of screws and precision metal parts (Fevotte et al., 2006). Industrial hygiene data from shops in this area indicated high intensity TCE exposures of 100 ppm or higher, particularly from exposures from hot degreasing processes. Considering exposure only from the jobs with a high level of confidence about exposure, 18% of exposed cases were identified with high cumulative exposure to TCE. The source population in the Brüning et al. (2003) study includes the Arnsberg region in Germany, which also has a high prevalence of TCE exposure. A large number of small companies used TCE in metal degreasing in small workrooms. Subjects in this study also described neurological symptoms previously

associated with higher TCE intensities. While subjects in the Brüning et al. (2003) study had potential high TCE exposure intensity, average TCE exposure in this study is considered lower than that in the Charbotel et al. (2006) study because the base population was enlarged beyond the Arnsberg region to areas which did not have the same focus of industry.

Siemiatycki (1991), Dosemeci et al. (1999), and Pesch et al. (2000) are population-based studies. Pesch et al. (2000) includes the Arnsberg area and 4 other regions. Sources of exposure to TCE and other chlorinated solvents are much less well defined, and most subjects identified with TCE exposure probably had minimal contact; estimated average concentrations to exposed subjects were of about 10 ppm or less (NRC, 2006). Neither Dosemeci et al. (1999) nor Siemiatycki (1991) describe the nature of the TCE exposure. TCE exposure potential in these studies is likely lower than in the three other studies and closer to background. Furthermore, the use of generic job-exposure-matrices for exposure assessment in these studies may result in a greater potential for exposure misclassification bias.

Nine of the 14 studies categorized results by exposure level. Three other studies reported results for other cancer sites by exposure level, but not kidney cancer; thus, to address this reporting bias, null values (i.e., RR estimates of 1.0) were used for these studies. Different exposure metrics were used in the various studies, and the purpose of combining results across the different highest exposure groups was not to estimate an RRp associated with some level of exposure, but rather to see the impacts of combining RR estimates that should be less affected by exposure misclassification. In other words, the highest exposure category is more likely to represent a greater differential TCE exposure compared to people in the referent group than the exposure differential for the overall (typically any vs. none) exposure comparison. Thus, if TCE exposure increases the risk of kidney cancer, the effects should be more apparent in the highest exposure groups. Indeed, the RRp estimate from the primary meta-analysis of the highest exposure group results was 1.53 (95% CI: 1.23, 1.91), which is greater than the RRp estimate of 1.25 (95% CI: 1.11, 1.41) from the overall exposure analysis. This result for the highest exposure groups was not overly influenced by any single study, nor was it overly sensitive to individual RR estimate selections. Heterogeneity was not observed in any of the analyses. The robustness of this finding lends substantial support to a conclusion that TCE exposure increases the risk of kidney cancer.

C.4. META-ANALYSIS FOR LIVER CANCER

C.4.1. Overall Effect of TCE Exposure

C.4.1.1. *Selection of RR Estimates*

The selected RR estimates for liver cancer associated with TCE exposure from the epidemiological studies are presented in Table C-11. There were no case-control studies for liver cancer and TCE exposure that were selected for inclusion in the meta-analysis (see Appendix B, Section II-9), so all of the relevant studies are cohort studies. All of the studies reported results for liver cancers plus cancers of the gall bladder and extrahepatic biliary passages (i.e., ICD-7 155.0 + 155.2; ICD-8 and -9 155 + 156). Three of the studies also report results for liver cancer alone (ICD-7 155.0; ICD-8 and -9 155). For the primary analysis, results for cancers of the liver, gall bladder, and biliary passages combined were selected, for the sake of consistency, since these were reported in all the studies. An alternate analysis was also done using results for liver cancer alone for the 3 studies that reported them and the combined liver cancer results for the remainder of the studies.

As for lymphoma and kidney cancer, many of the studies provided RR estimates only for males and females combined, and we are not aware of any basis for a sex difference in the effects of TCE on liver cancer risk; thus, wherever possible, RR estimates for males and females combined were used. The only study of much size (in terms of number of liver cancer cases) that provided results separately by sex was Raaschou-Nielsen (2003). The results of this study suggest that liver cancer risk in females might be slightly higher than the risk in males, but the number of female cases is small (primary liver cancer SIR: males 1.1 [95% CI: 0.74, 1.64; 27 cases], females 2.8 [1.13, 5.80; 7 cases]; gallbladder and biliary passage cancers SIR: males 1.1 [0.61, 1.87; 14 cases]; females 2.8 [1.28, 5.34; 9 cases]). Radican et al. (2008) report HRs for liver/biliary passage cancers combined of 1.36 (95% CI: 0.59, 3.11; 28 deaths) for males and 0.74 (95% CI: 0.18, 2.97; 3 deaths) for females, but these results are based on fewer cases, especially in females.

Table C-11. Selected RR estimates for liver cancer associated with TCE exposure (overall effect) from cohort studies

Study	RR	95% LCL	95% UCL	RR type	log RR	SE(log RR)	Alternate RR estimates	Comments
Anttila et al., 1995	1.89	0.86	3.59	SIR	0.637	0.333	2.27 (0.74, 5.29) for 155.0 alone	ICD-7 155.0 + 155.1; combined assuming Poisson distribution.
Axelson et al., 1994	1.41	0.38	3.60	SIR	0.344	0.5	1.34 (0.36, 3.42) with estimated female contribution to SIR added (see text)	ICD-7 155. Results reported for males only, but there was a small female component to the cohort.
Boice et al., 1999	0.54	0.15	1.38	SMR	−0.616	0.5	0.81 (0.45, 1.33) for any potential exposure	ICD-9 155 + 156. For potential routine exposure.
Greenland et al., 1994	0.54	0.11	2.63	OR	−0.616	0.810	None	ICD-8 155 + 156. Nested case-control study.
Hansen et al., 2001	2.1	0.7	5.0	SIR	0.742	0.447	None	ICD-7 155. Male and female results reported separately; combined assuming Poisson distribution.
Morgan et al., 1998	1.48	0.56	3.91	SMR	0.393	0.495	Published SMR 0.98 (0.36, 2.13)	ICD-9 155 + 156. Unpublished RR, adjusted for age and sex (see text).
Raaschou-Nielsen et al., 2003	1.35	1.03	1.77	SIR	0.300	0.138	1.28 (0.89, 1.80) for ICD-7 155.0	ICD-7 155.0 + 155.1. Results for males and females and different liver cancer types reported separately; combined assuming Poisson distribution.
Radican et al., 2008	1.12	0.57	2.19	Mortality HR	0.113	0.343	1.25 (0.31, 4.97) for ICD-8, -9 155.0	ICD-8, -9 155 + 156, ICD-10 C22-C24. Time variable = age; covariates = sex, race. Referent group is workers with no chemical exposures.
Boice et al., 2006	1.28	0.35	3.27	SMR	0.247	0.5	1.0 assumed for Zhao et al. (2005)	ICD-9 155 + 156. Boice et al. (2006) used in lieu of Zhao et al. (2005) because Zhao et al. (2005) do not report liver cancer results. Boice (2006) cohort overlaps Zhao cohort.

Most of the selections in Table C-11 should be self-evident, but some are discussed in more detail here, in the order the studies are presented in the table. For Axelson et al. (1994), in which a small subcohort of females was studied but only results for the larger male subcohort were reported, the reported male-only results were used in the primary analysis; however, as for lymphoma and kidney cancer, an attempt was made to estimate the female contribution to an overall RR estimate for both sexes and its impact on the meta-analysis. Axelson et al. (1994) reported that there were no cases of liver cancer observed in females, but the expected number was not presented. To estimate the expected number, the expected number for males was multiplied by the ratio of female-to-male person-years in the study and by the ratio of female-to-male age-adjusted incidence rates for liver cancer. The male results and the estimated female contribution were then combined into an RR estimate for both sexes assuming a Poisson distribution, and this alternate RR estimate for the Axelson et al. (1994) study was used in a sensitivity analysis.

For Boice et al. (1999), results for "potential routine exposure" were selected for the primary analysis, because this exposure category was considered to have less exposure misclassification, and results for "any potential exposure" were used in a sensitivity analysis. To estimate the SE(logRR) for the alternate RR selection, it was assumed that the number of exposed cases (deaths) was 15. The actual number was not presented, but 15 was the number that allowed us to reproduce the reported CIs. The number suggested by exposure level in Boice et al. (1999) Table 9 is 13; however, it may be that exposure level data were not available for all the cases. In their published paper, Morgan et al. (1998) present only SMRs for overall TCE exposure, although the results from internal analyses are presented for exposure subgroups. RR estimates for overall TCE exposure from the internal analyses of the Morgan et al. (1998) cohort data were available from an unpublished report (Environmental Health Strategies, 1997); from these, the RR estimate from the Cox model which included age and sex was selected, because those are the variables deemed to be important in the published paper. The internal analysis RR estimate was preferred for the primary analysis, and the published SMR result was used in a sensitivity analysis.

Raaschou-Nielsen et al. (2003) reported results for primary liver cancer (ICD-7 155.0), gallbladder and biliary passage cancers (ICD-7 155.1), and unspecified liver cancers (ICD-7 156) separately. As discussed above, RR estimates for cancers of the liver, gall bladder, and biliary passages combined were preferred for the primary analysis; thus, the results for primary liver cancer and gallbladder/biliary passage cancers were combined (across sexes as well), assuming a Poisson distribution. The results for primary liver cancer only (similarly combined across sexes) were used in an alternate analysis. The results for unspecified liver cancers (ICD-7 156) were

not included in any analyses because, under the ICD-7 coding, 156 can include secondary liver cancers. For Radican et al. (2008), the Cox model hazard ratio (HR) from the 2000 follow-up was used. In the Radican et al. (2008) Cox regressions, age was the time variable, and sex and race were covariates. It should also be noted that the referent group is composed of workers with no chemical exposures, not just no exposure to TCE.

Zhao et al. (2005) did not present RR estimates for liver cancer; thus, results from Boice et al. (2006) were used in the primary analysis. The cohorts for these studies overlap, so they are not independent studies. Zhao et al. (2005), however, was our preferred study for lymphoma and kidney cancer results; thus, in a sensitivity analysis, a null value (RR = 1.0) was assumed for Zhao et al. (2005) to address the potential reporting bias. The SE estimate for kidney cancer (incidence with 0 lag) was used as the SE for the liver cancer. (It is not certain that there was a reporting bias in this case. In the "Methods" section of their paper, Zhao et al. [2005] list the cancer sites examined in the cohort, and liver was not listed; it is not clear if the list of sites was determined *a priori* or *post hoc*.) Also, on the issue of potential reporting bias, the Siemiatycki (1991) study should be mentioned. This study was a case-control study for multiple cancer sites, but only the more common sites, in order to have greater statistical power. Thus, NHL and kidney cancer results were available, but not liver cancer results. Because no liver results were presented for any of the chemicals, this is not a case of reporting bias.

C.4.1.2. *Results of Meta-Analyses*

Results from some of the meta-analyses that were conducted on the epidemiological studies of TCE and liver cancer are summarized in Table C-12. The pooled estimate from the primary random effects meta-analysis of the 9 studies was 1.33 (95% CI: 1.09, 1.64) (see Figure C-9). As shown in Figure C-9, the analysis was dominated by one large study (contributing about 57% of the weight). That large study was critical in terms of statistical significance of the RRp estimate. Without the large Raaschou-Nielsen study, the RRp estimate does not change noticeably, but it is no longer statistically significant (RRp = 1.31; 95% CI: 0.96, 1.79). No other single study was overly influential; removal of any of the other individual studies resulted in RRp estimates that were all statistically significant and that ranged from 1.29 (with the removal of Anttila) to 1.39 (with the removal of Boice [1999]).

Table C-12. Summary of some meta-analysis results for TCE and liver cancer

Analysis	# of studies	Model	Combined RR estimate	95% LCL	95% UCL	Heterogeneity	Comments
All studies (all cohort studies)	9	Random	1.33	1.09	1.64	None obs (fixed = random)	Statistical significance not dependent on single study, except for Raaschou-Nielsen, without which $p = 0.08$. No apparent publication bias.
		Fixed	1.33	1.09	1.64		
All studies; liver cancer only, when available	9	Random	1.31	1.02	1.67	None obs	Used RR estimates for liver cancer alone for the 3 studies that presented these; remaining RR estimates are for liver and gall bladder/biliary passage cancers.
Alternate RR selections[a]	9	Random	1.33	1.08	1.63	None obs	With 1.0 assumed for Zhao in lieu of Boice (2006) (see text).
	9	Random	1.29	1.06	1.56	None obs	With Boice (1999) any potential exposure rather than potential routine exposure.
	9	Random	1.33	1.09	1.63	None obs	With estimated female contribution to Axelson.
	9	Random	1.30	1.07	1.59	None obs	With Morgan published SMR.
Highest exposure groups	6	Random	1.32	0.93	1.86	None obs	
	8	Random	1.28	0.93	1.77	None obs	Primary analysis. Using RR = 1 for Hansen and Zhao (see text).
	7−8	Random	1.24−1.26	0.88−0.91	1.73−1.82	None obs	Using alternate selections for Morgan and Raaschou-Nielsen and excluding Axelson.[a]

[a]Changing the primary analysis by one alternate RR each time.

obs = observable.

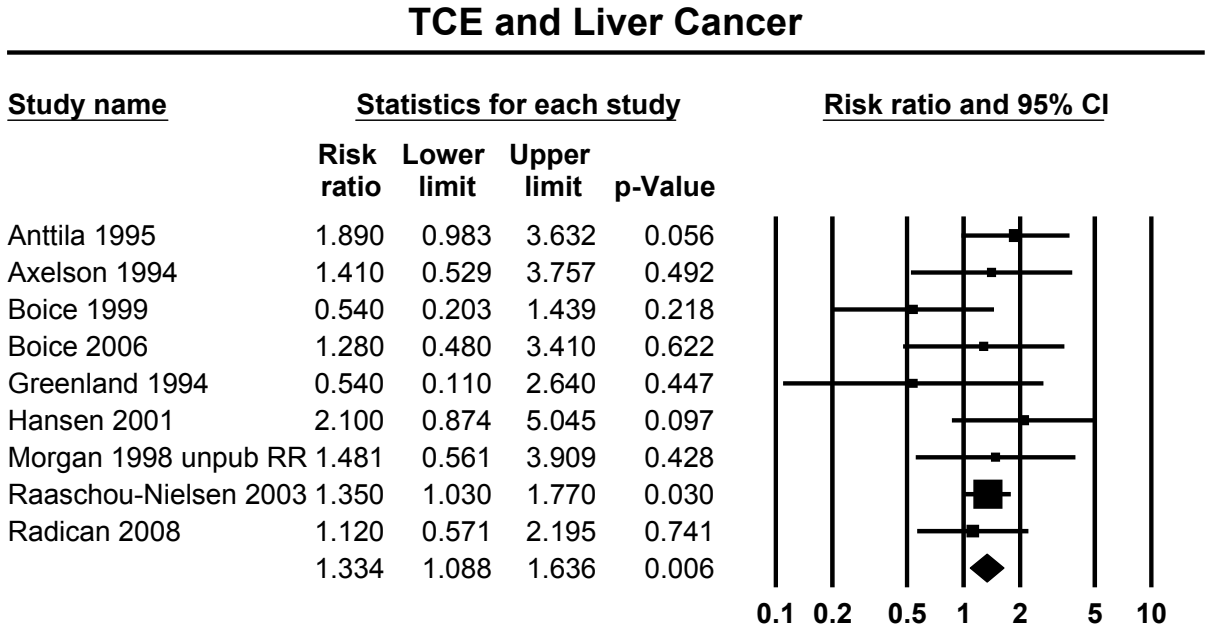

Figure C-9. Meta-analysis of liver cancer and TCE exposure. The pooled estimate is in the bottom row. Symbol sizes reflect relative weights of the studies. The horizontal midpoint of the bottom diamond represents the pooled RR estimate and the horizontal extremes depict the 95% CI limits.

As discussed in Section C.4.1.1, all of the 9 studies presented results for liver and gall bladder/biliary passage cancers combined, and these results were the basis for the primary analysis discussed above. An alternate analysis was performed substituting, simultaneously, results for liver cancer alone for the 3 studies for which these were available. The RRp estimate from this analysis was slightly lower than the one based entirely on results from the combined cancer categories (1.31; 95% CI: 1.02, 1.67). This result was driven by the fact that the RR estimate from the large Raaschou-Nielsen et al. (2003) study decreased from 1.35 for liver and gall bladder/biliary passage cancers combined to 1.28 for liver cancer alone.

Similarly, the RRp estimate was not highly sensitive to other alternate RR estimate selections. Use of the 4 other alternate selections, individually, resulted in RRp estimates that were all statistically significant (all with $p < 0.02$) and that ranged from 1.29 to 1.33 (see

Table C-12). In fact, as can be seen in Table C-12, only one of the alternates had notable impact. The Boice (2006), Zhao, and Axelson original values and alternate selections were associated with very little weight and, thus, have little influence in the RRp. Using the Boice (1999) alternate RR estimate based on any potential exposure rather than potential routine exposure decreased the RRp slightly from 1.33 to 1.29. The alternate Boice (1999) RR estimate is actually larger than the original value (0.81 vs. 0.54); however, use of the less discriminating exposure metric captures more liver cancer deaths, causing the weight of that study to increase from about 4.3% to almost 15%.

There was no apparent heterogeneity across the nine studies, i.e., the random effects model and the fixed effect model gave the same results. Furthermore, all of the liver cancer studies were cohort studies, so no subgroup analyses examining cohort and case-control studies separately, as was done for lymphoma and kidney cancer, were conducted. No alternate quantitative investigations of heterogeneity were pursued because of database limitations and, in any event, there is no evidence for heterogeneity of study results in this database.

As discussed in Section C.1, publication bias was examined in several different ways. The funnel plot in Figure C-10 shows little relationship between RR estimate and study size, and, indeed, none of the other tests performed found any evidence of publication bias. Duval and Tweedie's trim-and-fill procedure, for example, suggested that no studies were missing from the funnel plot, i.e., there was no asymmetry to counterbalance. Similarly, the results of a cumulative meta-analysis, incorporating studies with increasing SE one at a time, shows no evidence of a trend of increasing effect size with addition of the less precise studies. The Raaschou-Nielsen study contributes about 57% of the weight. Including the 2 next most precise studies, the RRp goes from 1.35 to 1.42 to 1.38 and the weight to 76%. With the addition of each of the next 3 most precise studies, the RRp estimate is 1.42. Further addition of the 3 least precise studies gradually brings the RRp back down to 1.33. Thus, if anything, the evidence is somewhat suggestive of an *inverse* relationship between SE and effect size, contrary to what would be expected if publication bias were occurring.

C.4.2. Liver Cancer Effect in the Highest Exposure Groups

C.4.2.1. *Selection of RR Estimates*

The selected RR estimates for liver cancer in the highest TCE exposure categories, for studies that provided such estimates, are presented in Table C-13. Six of the 9 cohort studies reported liver cancer risk estimates categorized by exposure level. As in Section C.4.1.1 for the overall risk estimates, estimates for cancers of the liver and gall bladder/biliary passages combined were preferentially selected, when presented, for the sake of consistency, and, wherever possible, RR estimates for males and females combined were used.

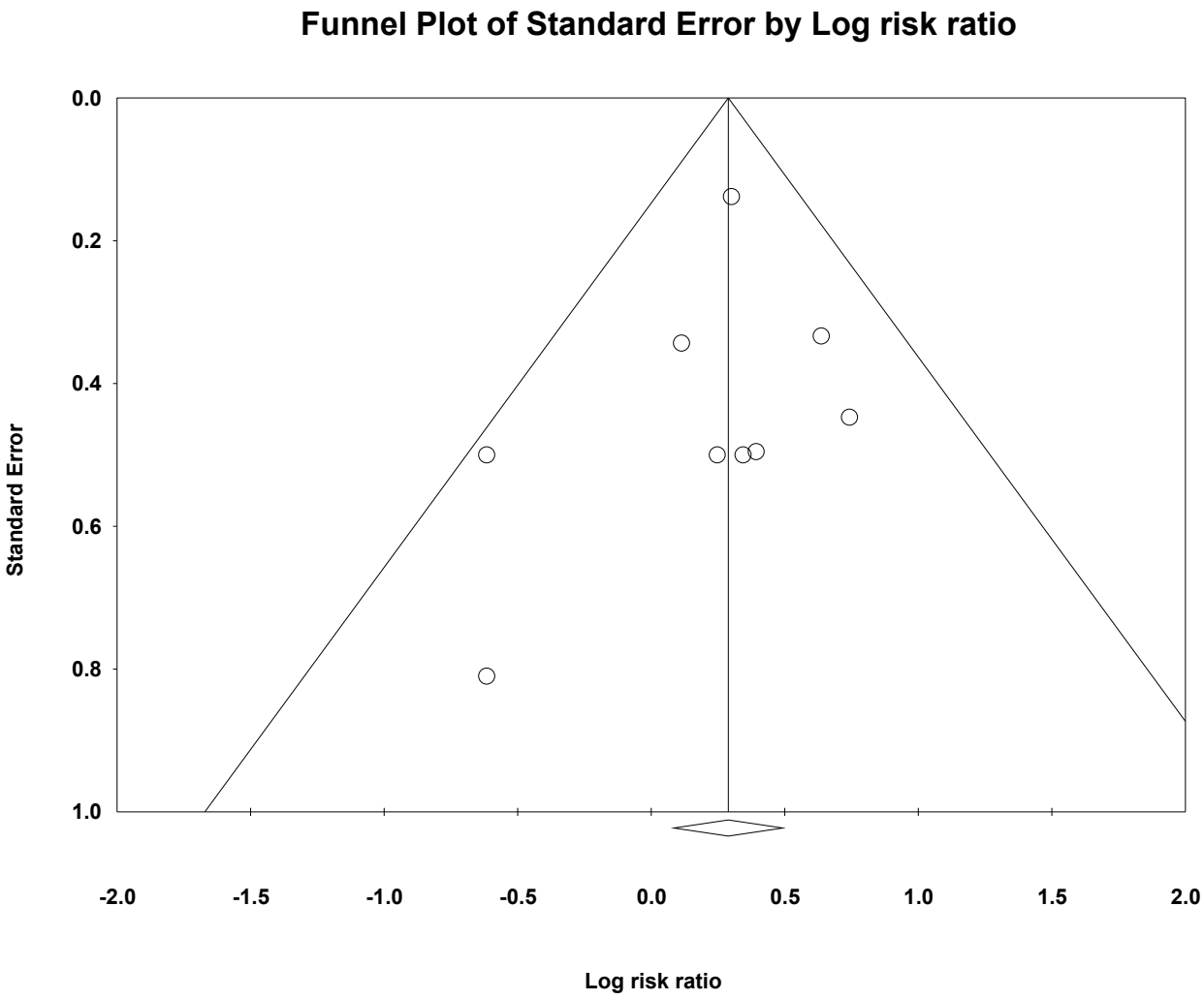

Figure C-10. Funnel plot of SE by log RR estimate for TCE and liver cancer studies.

Table C-13. Selected RR estimates for liver cancer risk in highest TCE exposure groups

Study	RR	95% LCL	95% UCL	Exposure category	log RR	SE(log RR)	Alternate RR estimates	Comments
Anttila et al., 1995	2.74	0.33	9.88	100+ µmol/L U-TCA [a]	1.008	0.707		SIR. ICD-7 155.0 (liver only).
Axelson et al., 1994	3.7	0.09	21	100+ mg/L U-TCA	1.308	1.000	Exclude study	SIR. ICD-7 155. 0 cases observed in highest exposure group (i.e., ≥2 y and 100+ U-TCA), so combined with <2 y and 100+ subgroup and females, estimating the expected numbers (see text).
Boice et al., 1999	0.94	0.36	2.46	≥ 5 yr exposure	−0.062	0.490	None	Mortality RR. ICD-9 155 + 156. For potential routine or intermittent exposure. Adjusted for date of birth, dates 1st and last employed, race, and sex. Referent group is workers not exposed to any solvent.
Hansen et al., 2001				≥ 1080 mos × mg/m^3			1.0 assumed	Reported high exposure group results for some cancer sites but not liver.
Morgan et al., 1998	1.19	0.34	4.16	High cumulative exposure score	0.174	0.639	0.98 (0.29, 3.35) for med/high peak vs. low/no	Mortality RR. ICD-9 155 + 156. Adjusted for age and sex.
Raaschou-Nielsen et al., 2003	1.2	0.7	1.9	≥ 5 yrs	0.182	0.243	1.1 (0.5, 2.1) ICD-7 155.0 (liver only)	SIR. ICD-7 155.0 + 155.1. Male and female results presented separately and combined assuming a Poisson distribution.
Radican et al., 2008	1.49	0.67	3.34	> 25 unit-yr	0.399	0.411	None (see text)	Mortality HR. ICD-8, -9 155 + 156, ICD-10 C22-C24. Male and female results presented separately and combined (see text). Time variable = age, covariate = race. Referent group is workers with no chemical exposures.
Zhao et al., 2005				High exposure score			1.0 assumed	No liver results reported.

[a]Mean personal trichloroacetic acid in urine. 1 µmol/L = 0.1634 mg/L.

Two of the 9 cohort studies (Hansen et al., 2001; Zhao et al., 2005) did not report liver cancer risk estimates categorized by exposure level even though these same studies reported such estimates for selected other cancer sites. To address this reporting bias (as discussed above, Zhao et al. [2005] did not present any liver results, and it is not clear if this was actual reporting bias or an *a priori* decision not to examine liver cancer in the cohort.), attempts were made to obtain the results from the primary investigators, and, failing that, alternate analyses were performed in which null estimates (RR = 1.0) were included for both studies. This alternate analysis was then used as the main analysis, e.g., the basis of comparison for the sensitivity analyses. For the SE (of the logRR) estimates for the null estimates, SE estimates from other sites for which highest-exposure-group results were available were used. For Hansen et al. (2001), the SE estimate for NHL in the highest exposure group was used, because NHL was the only cancer site of interest in this assessment for which highest-exposure-group results were available. For Zhao et al. (2005), the SE estimate for kidney cancer in the highest-exposure group (incidence with 0 lag) was used. (Note that Boice et al. [2006], who studied a cohort that overlapped that of Zhao et al. [2005], also did not present liver cancer results by exposure level.)

For Axelson et al. (1994), there were no liver cancer cases in the highest exposure group (≥2 years and 100+ mean urinary-trichloroacetic acid [U-TCA] level), so no log RR and SE(log RR) estimates were available for the meta-analysis. Instead, the <2 years and ≥2 years results were combined, assuming expected numbers of cases were proportional to person-years, and 100+ U-TCA (with any exposure duration) was used as the highest exposure category. The female contribution to the expected number was also estimated, again assuming proportionality to person-years, and adjusting for the difference between female and male age-adjusted liver cancer incidence rates. The estimated RR and SE values for the combined exposure times and sexes were used in the primary analysis. In an alternate analysis, the Axelson et al. (1994) study was excluded altogether, because we estimated that less than 0.2 cases were expected in the highest-exposure category, suggesting that the study had low power to detect an effect in the highest-exposure group and would contribute little weight to the meta-analysis.

For Boice et al. (1999), only results for workers with "any potential exposure" (rather than "potential routine exposure") were presented by exposure category, and the referent group is workers not exposed to any solvent. For Morgan et al. (1998), the primary analysis used results for the cumulative exposure metric, and a sensitivity analysis was done with the results for the peak exposure metric. For Raaschou-Nielsen et al. (2003), unlike for NHL and RCC, liver cancer results for the subcohort with expected higher exposure levels were not presented, so the only highest-exposure group results were for duration of employment in the total cohort. Results

for cancers of the liver and gall bladder/biliary passages combined were used for the primary analysis and results for liver cancer alone in a sensitivity analysis.

For Radican et al. (2008), it should be noted that the referent group is workers with no chemical exposures, not just no TCE exposure. Furthermore, exposure group results were reported separately for males and females and were combined for this assessment using inverse-variance weighting, as in a fixed effect meta-analysis. In addition to results for biliary passage and liver cancer combined, Radican et al. (2008) present results for liver only by exposure group; however, there were no liver cancer deaths in females and the number expected was not reported, so no alternate analysis for the highest-exposure groups with an RR estimate from Radican et al. (2008) for liver cancer only was conducted. Radican et al. (2008) present only mortality HR estimates by exposure group; however, in an earlier follow-up of this same cohort, Blair et al. (1998) present both incidence and mortality RR estimates by exposure group. As with the Radican et al. (2008) liver cancer only results, however, there were no incident cases for females in the highest-exposure group in Blair et al. (1998) (and the expected number was not reported). Additionally, there were more biliary passage/liver cancer deaths (31) in Radican et al. (2008) than incident cases (13) in Blair et al. (1998) overall and in the highest-exposure group (14 vs. 4). Thus, we elected to use only the Radican et al. (2008) mortality results from this cohort and not to include an alternate analysis based on incidence results from the earlier follow-up.

C.4.2.2. *Results of Meta-Analyses*

Results from the meta-analyses that were conducted for liver cancer in the highest exposure groups are summarized at the bottom of Table C-12. The pooled RR estimate from the random effects meta-analysis of the 6 studies with results presented for exposure groups was 1.32 (95% CI: 0.93, 1.86). As with the overall liver cancer meta-analyses, the meta-analyses of the highest-exposure groups were dominated by one study (Raaschou-Nielsen), which provided about 52% of the weight. The RRp estimate from the primary random effects meta-analysis with null RR estimates (i.e., 1.0) included for Hansen and Zhao to address (potential) reporting bias (see above) was 1.28 (95% CI: 0.93, 1.77) (see Figure C-11). The inclusion of these 2 additional studies contributed about 10% of the total weight. No single study was overly influential (removal of individual studies resulted in RRp estimates that ranged from 1.23 to 1.36) and the RRp estimate was not highly sensitive to alternate RR estimate selections (RRp estimates with alternate selections ranged from 1.24 to 1.26; see Table C-12). In addition, there was no observable heterogeneity across the studies for any of the meta-analyses conducted with the highest-exposure groups. However, none of the RRp estimates was statistically significant.

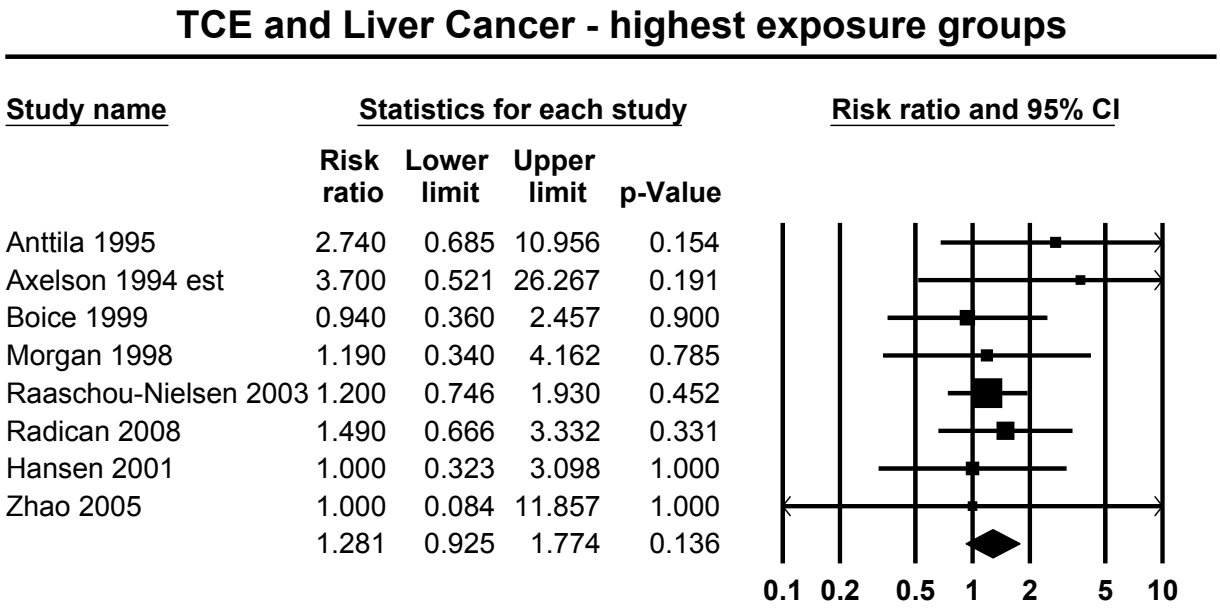

Figure C-11. Meta-analysis of liver cancer and TCE exposure—highest exposure groups, with assumed null RR estimates for Hansen and Zhao (see text).

Furthermore, the RRp estimates for the highest-exposure groups were all less than the significant RRp estimate for an overall effect on liver cancer (1.33; 95% CI: 1.09, 1.64; see Section C.4.2.2 and Table C-12). This contradictory result is driven by the fact that the RR estimate for the highest-exposure group was less than the overall RR estimate for Raaschou-Nielsen, which contributes the majority of the weight to the meta-analyses. The liver cancer results are relatively underpowered with respect to numbers of studies and number of cases, and the Raaschou-Nielsen study, which dominates the analysis, uses duration of employment as an exposure-level surrogate for liver cancer, and duration of employment is a notoriously weak exposure metric. Thus, the contradictory finding that the RRp estimates for the highest-exposure groups were all less than the RRp estimate for an overall effect does not rule out an effect of TCE on liver cancer; however, it certainly does not provide additional support for such an effect.

C.4.3. Discussion of Liver Cancer Meta-Analysis Results

For the most part, the meta-analyses of the overall effect of TCE exposure on liver (and gall bladder/biliary passages) cancer suggest a small, statistically significant increase in risk. The pooled estimate from the primary random effects meta-analysis of the 9 (all cohort) studies was 1.33 (95% CI: 1.09, 1.64). The analysis was dominated by one large study that contributed about 57% of the weight. When this study was removed, the RRp estimate did not change much, but it was no longer statistically significant (RRp = 1.31; 95% CI: 0.96, 1.79). The pooled estimate was not overly influenced by any other single study, nor was it overly sensitive to individual RR estimate selections. The largest downward impacts were from the removal of the Anttila study, resulting in an RRp estimate of 1.29 (95% CI: 1.04, 1.59), and from the substitution of the Boice (1999) RR estimate for potential routine exposure with that for any potential exposure, resulting in an RRp estimate of 1.29 (1.06, 1.56). Substituting the RR estimates for liver/gall bladder/biliary passage cancers with those of liver cancer alone for the 3 studies that provided these results yielded an RRp estimate of 1.31 (1.02, 1.67). There was no evidence of publication bias in this data set, and there was no observable heterogeneity across the study results.

Six of the 9 studies provided liver cancer results by exposure level. Two other studies reported results for other cancer sites by exposure level, but not liver cancer; thus, to address this reporting bias, null values (i.e., RR estimates of 1.0) were used for these studies. Different exposure metrics were used in the various studies, and the purpose of combining results across the different highest exposure groups was not to estimate an RRp associated with some level of exposure, but rather to see the impacts of combining RR estimates that should be less affected by exposure misclassification. In other words, the highest exposure category is more likely to represent a greater differential TCE exposure compared to people in the referent group than the exposure differential for the overall (typically any vs. none) exposure comparison. Thus, if TCE exposure increases the risk of liver cancer, the effects should be more apparent in the highest exposure groups. However, the RRp estimate from the meta-analyses of the highest exposure group results were less than the RRp estimate from the overall exposure analysis. This anomalous result is driven by the fact that, for Raaschou-Nielsen, which contributes the majority of the weight to the meta-analyses, the RR estimate for the highest-exposure group, although greater than 1.0, was less than the overall RR estimate.

Thus, while there is the suggestion of an increased risk for liver cancer associated with TCE exposure, the statistical significance of the pooled estimates is dependent on one study, which provides the majority of the weight in the meta-analyses. Removal of this study does not change the RRp estimate; however, it becomes nonsignificant ($p = 0.08$). Furthermore, meta-

analysis results for the highest-exposure groups yielded *lower* RRp estimates than for an overall effect. These results do not rule out an effect of TCE on liver cancer, because the liver cancer results are relatively underpowered with respect to numbers of studies and number of cases and the overwhelming study in terms of weight uses the weak exposure surrogate of duration of employment for categorizing exposure level; however, at present, there is only modest support for such an effect.

C.5. DISCUSSION OF STRENGTHS, LIMITATIONS, AND UNCERTAINTIES IN THE META-ANALYSES

Meta-analysis provides a systematic way of objectively and quantitatively combining the results of multiple studies to obtain a summary effect estimate. Use of meta-analysis can help risk assessors avoid some of the potential pitfalls in overly relying on a single study or in making more subjective qualitative judgments about the apparent weight of evidence across studies. Combining the results of smaller studies also increases the statistical power to observe an effect, if one exists. In addition, meta-analysis techniques assist in systematically investigating issues such as potential publication bias and heterogeneity in a database.

While meta-analysis can be a useful tool for analyzing a database of epidemiological studies, the analysis is limited by the quality of the input data. If the individual studies are deficient in their abilities to observe an effect (in ways other than low statistical power, which meta-analysis can help ameliorate), the meta-analysis will be similarly deficient. A critical step in the conduct of a meta-analysis is to establish eligibility criteria and clearly and transparently identify all relevant studies for inclusion in the meta-analysis. For the TCE database, a comprehensive qualitative review of available studies was conducted and eligible studies were identified, as described in Appendix B, Section II-9.

Identifying all relevant studies may be hampered if publication bias has occurred. Publication bias is a systematic error that can arise if statistically significant studies are more likely to be published than nonsignificant studies. This can result in an upward bias on the effect size measure, i.e., the relative risk estimate. To address this concern, potential publication bias was investigated for the databases for which meta-analyses were undertaken. For the studies of kidney cancer and liver cancer, there was no evidence of publication bias. For the studies of lymphoma, there was some evidence of potential publication bias. It is uncertain whether this reflects actual publication bias or rather an association between SE and effect size (as discussed in Section C.1, a feature of publication bias is that smaller studies tend to have larger effect sizes) resulting for some other reason, e.g., a difference in study populations or protocols in the smaller studies. Furthermore, if there is publication bias in this data set, it may be creating an

upward bias on the relative risk estimate, but this bias does not appear to account completely for the finding of an increased lymphoma risk (see Section C.2.1.2).

Another concern in meta-analyses is heterogeneity across studies. Random-effects models were used for the primary meta-analyses in this assessment because of the diverse nature of the individual studies. When there is no heterogeneity across the study results, the random-effects model will give the same result as a fixed-effect model. When there is heterogeneity, the random-effects model estimates the between-study variance. Thus, when there is heterogeneity, the random-effects model will generate wider confidence intervals and be more "conservative" than a fixed-effect model. However, if there is substantial heterogeneity, it may be inappropriate to combine the studies at all. In cases of significant heterogeneity, it is important to try to investigate the potential sources of the heterogeneity.

For the studies of kidney cancer and liver cancer, there was no apparent heterogeneity across the study results, i.e., random- and fixed-effects models gave identical summary estimates. For the lymphoma studies, there was heterogeneity, but it was not statistically significant ($p = 0.10$). The I^2 value was 33%, suggesting low-to-moderate heterogeneity. When subgroup analyses were done for the cohort and case-control studies separately, there was some heterogeneity in both groups, but in neither case was it statistically significant. Further attempts to quantitatively investigate the heterogeneity were not pursued because of limitations in the database. The sources of heterogeneity are an uncertainty in the database of studies of TCE and lymphoma. Some potential sources of heterogeneity, which are discussed qualitatively in Section C.2.3, include differences in exposure assessment or in the intensity or prevalence of TCE exposures in the study population and differences in lymphoma classification.

The joint occurrence of heterogeneity and potential publication bias in the database of studies of TCE and lymphoma raises special concerns. Because of the heterogeneity, a random-effects model should be used if these studies are to be combined; yet, the random-effects model gives relatively large weight to small studies, which could exacerbate the potential impacts of publication bias. For the lymphoma studies, the summary relative risk estimates from the random-effects and fixed-effect models are not very different (RRp = 1.23 [95% CI: 1.04, 1.44] and 1.19 [1.06, 1.34], respectively); however, the confidence interval for the fixed-effect estimate does not reflect the between-study variance and is, thus, overly narrow.

C.6. CONCLUSIONS

The strongest finding from the meta-analyses was for TCE and kidney cancer. The summary estimate from the primary random-effects meta-analysis of the 14 studies was RRp = 1.25 (95% CI: 1.11, 1.41). There was no apparent heterogeneity across the study results

(i.e., fixed-effect model gave same summary estimate), and there was no evidence of potential publication bias. The summary estimate was robust across influence and sensitivity analyses; the estimate was not markedly influenced by any single study, not was it overly sensitive to individual RR estimate selections. The findings from the meta-analyses of the highest exposure groups for the studies that provided results categorized by exposure level were similarly robust. The summary estimate was RRp = 1.53 (95% CI: 1.23, 1.91) for the 12 studies included in the analysis. There was no apparent heterogeneity in the highest-exposure group results, and the estimate was not markedly influenced by any single study, nor was it overly sensitive to individual RR estimate selections. In sum, these robust results support a conclusion that TCE exposure increases the risk of kidney cancer.

For the most part, the meta-analyses of the overall effect of TCE exposure on lymphoma also suggest a small, statistically significant increase in risk. The summary estimate from the primary random-effects meta-analysis of the 16 studies was 1.23 (95% CI: 1.04, 1.44). This result was not overly influenced by any single study, nor was it overly sensitive to individual RR estimate selections, although use of one alternate RR estimate considered clearly inferior narrowly eliminated statistical significance of the summary estimate ($p = 0.050$). There is some evidence of potential publication bias in the lymphoma study data set; however, it is uncertain that this is actually publication bias rather than an association between SE and effect size resulting for some other reason, e.g., a difference in study populations or protocols in the smaller studies. Furthermore, if there is publication bias, it does not appear to account completely for the findings of an increased lymphoma risk. There was some heterogeneity across the results of the 16 studies, but it was not statistically significant ($p = 0.10$). The I^2 value was 33%, suggesting low-to-moderate heterogeneity. The source(s) of this heterogeneity remains an uncertainty. The summary estimate from the meta-analysis of the highest exposure groups for the 12 studies which provided results categorized by exposure level was RRp = 1.57 (95% CI: 1.27, 1.94). This result for the highest exposure groups was not overly influenced by any single study, nor was it overly sensitive to individual RR estimate selections, and heterogeneity was not observed in any of the relevant analyses. The robustness of the finding of an increased lymphoma risk for the highest exposure groups strengthens the more moderate evidence from the meta-analyses for overall effect.

The meta-analyses of the overall effect of TCE exposure on liver (and gall bladder/biliary passages) cancer also suggest a small, statistically significant increase in risk, but the study database is more limited. The pooled estimate from the primary random-effects meta-analysis of the 9 (all cohort) studies was 1.33 (95% CI: 1.09, 1.64). The analysis was dominated by one large study that contributed about 57% of the weight. When this study was removed, the RRp

estimate did not change much, but it was less precise (RRp = 1.31; 95% CI: 0.96, 1.79). The pooled estimate was not overly influenced by any other single study, nor was it overly sensitive to individual RR estimate selections. There was no evidence of publication bias in this data set, and there was no observable heterogeneity across the study results. However, the findings from the meta-analyses of the highest-exposure groups for the studies that provided results categorized by exposure level do not add support to the overall effect findings. The summary estimate was RRp = 1.28 (95% CI: 0.93, 1.77) for the 8 studies included in the analysis, which is *lower* than the summary estimate for the overall effect. This contradictory result is driven by the fact that the RR estimate for the highest-exposure group in the individual study which contributes the majority of the weight to the meta-analyses, although greater than 1.0, was less than the overall RR estimate for the same study. In sum, these results do not rule out an effect of TCE on liver cancer, because the liver cancer results are relatively underpowered with respect to numbers of studies and number of cases and the overwhelming study in terms of weight uses the weak exposure surrogate of duration of employment for categorizing exposure level; however, at present, there is only modest support for an increased risk of liver cancer.

C.7. REFERENCES

Anttila, A; Pukkala, E; Sallmén, M; et al. (1995) Cancer incidence among Finnish workers exposed to halogenated hydrocarbons. J Occup Environ Med 37:797−806.

Axelson, O; Selden, A; Andersson, K; et al. (1994) Updated and expanded 1 Swedish cohort study on trichloroethylene and cancer risk. J Occup Med 36:556−562.

Banks, PM. (1992) Changes in diagnosis of non-Hodgkin's lymphomas over time. Cancer Res 52:5453s−5455s.

Begg, CB; Mazumdar, M. (1994) Operating characteristics of a rank correlation test for publication bias. Biometrics 50:1088−1101.

Blair, A; Hartge, P; Stewart, PA; et al. (1998) Mortality and cancer incidence of aircraft maintenance workers exposed to trichloroethylene and other organic solvents and chemicals: extended follow up. Occup Environ Med 55:161−171.

Boice, JD; Marano, DE; Fryzek, JP; et al. (1999) Mortality among aircraft manufacturing workers. Occup Environ Med 56:581−597.

Boice, JD; Marano, DE; Cohen, SS; et al. (2006) Mortality among Rocketdyne worker who tested rocket engines, 1948-1999. J Occup Environ Med 48:1070−1092.

Breslow, NE; Day, NE. (1980) Statistical methods in cancer research. Vol. I. The Analysis of Case-Control Studies. IARC Scientific Publications No. 32. Lyon, France: International Agency for Research on Cancer.

Breslow, NE; Day, NE. (1987) Statistical methods in cancer research, Vol. II. The Design and Analysis of Cohort Studies. IARC Scientific Publications No. 82. Lyon, France: International Agency for Research on Cancer.

Brüning, T; Pesch, B; Wiesenhütter, B; et al. (2003) Renal cell cancer risk and occupational exposure to trichloroethylene: results of a consecutive case-control study in Arnsberg, Germany. Am J Ind Med 43:274−285.

Charbotel, B; Fevotte, J; Hours, M; et al. (2006) Case-control study on renal cell cancer and occupational exposure to trichloroethylene. Part II: Epidemiological aspects. Ann Occup Hyg 50(8):777−787.

Cherrie, JW; Kromhout; H; Semple, S. (2001). The importance of reliable exposure estimates in deciding whether trichloroethylene can cause kidney cancer [letter]. J Cancer Res Clin 127:400−402.

Copeland, KT; Checkoway, H; McMichael, AJ; et al. (1977) Bias due to misclassification in the estimation of relative risk. Am J Epidemiol 105:488−495.

DerSimonian, R; Laird, N. (1986) Meta-analysis in clinical trials. Controlled Clinical Trials 7:177−188.

Dosemeci, M; Cocco, P; Chow, WH. (1999) Gender differences in risk of renal cell carcinoma and occupational exposures to chlorinated aliphatic hydrocarbons. Am J Ind Med 36(1):54−59.

Duval, S; Tweedie, R. (2000) A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. J Am Stat Assoc 95:89−98.

Egger, M; Smith, GD; Schneider, M; et al. (1997) Bias in meta-analysis detected by a simple, graphical test. Brit Med J 315:629−634.

Environmental Health Strategies. (1997) Final Report. Written correspondence from Paul A. Cammer, Ph.D., Trichloroethylene Issues Group, to Cheryl Siegel Scott, US Environmental Protection Agency dated December 22, 1997.

Fevotte, J; Charbotel, B; Muller-Beaute, P; et al. (2006) Case-control study on renal cell cancer and occupational exposure to trichloroethylene. Part I: Exposure assessment. Ann Occup Hyg 50:765−775.

Greenland, S; Salvan, A; Wegman, DH; et al. (1994) A case-control study of cancer mortality at the transformer-assembly facility. Int Arch Occ Env Heal 66:49−54.

Hansen, J; Raaschou-Nielsen, O; Christensen, JM; et al. (2001) Cancer incidence among Danish workers exposed to trichloroethylene. J Occup Environ Med 43:133−139.

Hardell, L; Eriksson, M; Degerman, A. (1994) Exposure to phenoxyacetic acids, chlorophenols, or organic solvents in relation to histopathology, stage, and anatomical localization of non-Hodgkin's lymphoma. Cancer Res 54:2386−2389.

Harris, NL; Jaffe, ES; Diebold, J; et al. (2000) Lymphoma classification – from controversy to consensus: the R.E.A.L. and WHO classification of lymphoid neoplasms. Ann Oncol 11 Suppl 1:3−10.

Higgins JPT, Thompson SG, Deeks JJ, Altman DG. 2003. Measuring inconsistency in meta-analyses. Brit Med J 327:557-560.

Mandel, JH; Kelsh, MA; Mink, PJ; et al. (2006) Occupational trichloroethylene exposure and non-Hodgkin's lymphoma: a meta-analysis and review. Occup Environ Med 63:597−607.

Mannetje, A.; Steenland, K; Attfield, M; et al. (2002) Exposure-response analysis and risk assessment for silica and silicosis mortality in a pooled analysis of six cohorts. Occup. Environ. Med. 59:723−728.

McGuire, V; Nelson, LM; Koepsell, TD; et al. (1998) Assessment of occupational exposures in community-based case-control studies. Annu. Rev. Publ Health 19:35−53.

Miligi, L; Costantini, AS; Benvenuti, A; et al. (2006) Occupational exposure to solvents and the risk of lymphomas. Epidemiology 17:552−561.

Morgan, RW; Kelsh, MA; Zhao, K; et al. (1998) Mortality of aerospace workers exposure to trichloroethylene. Epidemiology 9:424−431

Morton LM, Holford TR, Leaderer B, Zhang Y, Zahm SH, et al. 2003. Alcohol use and risk of non-Hodgkin's lymphoma among Connecticut somen (United States). Cancer Causes Control 14:687-694.

Nelson, LM; Longstreth, WT, Jr,; Koepsell, TD; et al. (1994) Completeness and accuracy of interview data from proxy respondents: Demographic, medical, and life-style factors. Epidemiology 5:204−217.

Nordstrom, M; Hardell, L; Hagberg, H; et al. (1998) Occupational exposures, animal exposure and smoking as risk factors for hairy cell leukaemia evaluated in a case-control study. Brit J Cancer 77:2048−2052.

NRC (National Research Council). (2006) Assessing the Human Health Risks of Trichloroethylene. Key Scientific Issues. Washington, DC: National Academy Press.

Persson B, Fredrikson M. 1999. Some risk factors for non-Hodgkin's lymphoma. Int J Occup Med Environ Health 12:135−142.

Pesch, B; Haerting, J; Ranft, U; et al. (2000) Occupational risk factors for renal cell carcinoma: Agent-specific results from a case-control study in Germany. Int J Epidemiol 29:1014−1024.

Raaschou-Nielsen, O; Hansen, J; McLaughlin, JK; et al. (2003) Cancer risk among workers at Danish companies using trichloroethylene: A cohort study. Am J Epidemiol 158:1182−1192.

Radican L, Blair A, Stewart P, Wartenberg D. 2008. Mortality of aircraft maintenance workers exposed to trichloroethylene and other hydrocarbons and chemicals: extended follow-up. J Occup Environ Med 50:1306-1319.

Rothman, KJ; Greenland, S. (1998) Modern Epidemiology, 2nd Edition. Philadelphia, PA: Lippincott Williams & Wilkins.

Seidler, A; Mohner, M; Berger, J; et al. (2007) Solvent exposure and malignant lymphoma: A population-based case-control study in Germany. J Occup Med Toxicol 2:2.

Siemiatycki, J. (1991) Risk Factors for Cancer in the Workplace. J Siemiatycki, ed. Baca Raton, FL: CRC Press.

Wang R, Zhang Y, Lan Q, Holford TR, Beaderer B, Zahm SH, Boyle P, Dosemeci M, Rothman N, Zhu Y, Qin Q, Zheng T. 2009. Occupational exposure to solvents and risk of non-Hodgkin lymphoma in Connecticut women. Am J Epidemiol 169:176-185.

Woolf, B. (1955) On estimating the relationship between blood group and disease. Ann Hum Genet 19:251−253.

Zhao, Y; Krishnadasan, A; Kennedy, N; et al. (2005) Estimated effects of solvents and mineral oils on cancer incidence and mortality in a cohort of aerospace workers. Am J Ind Med 48:249−258.