

1 **C. Meta-analysis of Cancer Results from Epidemiological Studies**

2
3 **C.1. Methodology**

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

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

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

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

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

$$SE = \frac{(\log(UCL) - \log(LCL))}{3.92},$$

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

$$Var(\log(OR)) = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d},$$

29 in accordance with the method proposed by Woolf (1955), as described by Breslow and Day
 30 (1980).

1 The analyses that were performed for this assessment include:

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

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

19 The heterogeneity (or homogeneity) analysis tests the hypothesis that the study results are
20 homogeneous, i.e., that all the RR estimates are estimating the same population RR and the total
21 variance is no more than would be expected from within-study variance. Heterogeneity was
22 assessed using the statistic Q described by DerSimonian and Laird (1986). The Q -statistic
23 represents the sum of the weighted squared differences between the summary RR estimate
24 (obtained under the null hypothesis, i.e., using a fixed-effect model) and the RR estimate from
25 each study, and, under the null hypothesis, Q approximately follows a χ^2 distribution with
26 degrees of freedom equal to the number of studies minus one.

27 Publication bias is a systematic error that occurs if statistically significant studies are
28 more likely to be submitted and published than nonsignificant studies. Studies are more likely to
29 be statistically significant if they have large effect sizes (in this case, RR estimates); thus, an
30 upward bias would result in a meta-analysis if the available published studies have higher effect
31 sizes than the full set of studies that was actually conducted. One feature of publication bias is
32 that smaller studies tend to have larger effect sizes than larger studies, since smaller studies need
33 larger effect sizes in order to be statistically significant. Thus, many of the techniques used to
34 analyze publication bias examine whether or not effect size is associated with study size.
35 Methods used to investigate potential publication bias for this assessment included funnel plots,
36 which plot effect size versus study size (actually, SE versus log RR here); the “trim and fill”

1 procedure of Duvall and Tweedie (2000), which imputes the “missing” studies in a funnel plot
2 (i.e., the studies needed to counterbalance an asymmetry in the funnel plot resulting from an
3 ostensible publication bias) and recalculates a summary effect size with these studies present;
4 forest plots (arrays of RRs and CIs by study) sorted by precision (i.e., SE) to see if effect size
5 shifts with study size; Begg and Mazumdar rank correlation test (Begg and Mazumdar, 1994),
6 which examines the correlation between effect size estimates and their variances after
7 standardizing the effect sizes to stabilize the variances; Egger’s linear regression test (Egger et
8 al., 1997), which tests the significance of the bias reflected in the intercept of a regression of
9 effect size/SE on 1/SE; and cumulative meta-analyses after sorting by precision to assess the
10 impact on the summary effect size estimate of progressively adding the smaller studies.
11

12 **C.2. Meta-analysis for Lymphoma**

13 **C.2.1. Overall effect of TCE exposure**

14 **C.2.1.1. Selection of RR estimates**

15 The selected RR estimates for lymphoma associated with TCE exposure from the
16 selected epidemiological studies are presented in Table C.1 for cohort studies and in Table C.2
17 for case-control studies. A few of the more recent case-control studies classified lymphomas
18 along the lines of the recent WHO/REAL classification system (World Health
19 Organization/Revised European-American Classification of Lymphoid Neoplasms) (Harris et al.,
20 2000); however, most of the available TCE studies reported lymphoma results according to the
21 International Disease Classification (ICD), Revisions 7, 8, and 9, and focused on non-Hodgkin
22 lymphoma (NHL; ICD 200 + 202). For consistency of endpoint in the lymphoma meta-analyses,
23 we selected RR estimates for ICD 200 + 202, wherever possible; otherwise, we selected
24 estimates for the classification(s) best approximating NHL. In addition, many of the studies
25 provided RR estimates only for males and females combined, and we are not aware of any basis
26 for a sex difference in the effects of TCE on lymphoma risk; thus, wherever possible, we used
27 RR estimates for males and females combined.

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

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

12 Most of the selections in Tables C.1 and C.2 should be self-evident, but some are
13 discussed in more detail here, in the order the studies are presented in the Tables. For Blair et al.
14 (1998), it should be noted that the referent group is composed of workers with no chemical
15 exposures, not just no exposure to TCE. For Boice et al. (1999), results for “potential routine
16 exposure” were selected for the primary analysis, because this exposure category was considered
17 to have less exposure misclassification, and results for “any potential exposure” were used in a
18 sensitivity analysis. The Greenland et al. (1994) study is a case-control study nested within a
19 worker cohort, and we treat it here as a cohort study (see Appendix B, Section II-9.1). For
20 Morgan et al. (1998), the reported results did not allow for the combination of ICD 200 and 202,
21 so the SMR estimate for the combined 200 + 202 grouping was taken from the meta-analysis
22 paper of Mandel et al. (2006), who included one of the investigators from the Morgan et al.
23 study. RR estimates for overall TCE exposure from internal analyses of the Morgan et al. cohort
24 data were available from an unpublished report (Environmental Health Strategies, 1997; the
25 published paper only presented the internal analyses results for exposure subgroups), but only for
26 ICD 200; from these, the RR estimate from the Cox model which included age and sex was
27 selected, because those are the variables deemed to be important in the published paper (Morgan
28 et al., 1998). Although the results from internal analyses are generally preferred, in this case the

¹ person-years for men and women ≤ 79 y 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.

1 SMR estimate was used in the primary analysis and the internal analysis RR estimate was used in
2 a sensitivity analysis because the latter estimate represented an appreciably smaller number of
3 deaths (3, based on ICD 200 only) than the SMR estimate (9, based on ICD 200 + 202).

4 For Zhao et al. (2005), RR estimates were only reported for ICD-9 200–208 (all
5 lymphohematopoietic cancers), and not for 200 + 202 alone. Given that other studies have not
6 reported associations between leukemias and TCE exposure, combining all lymphohematopoietic
7 cancers would dilute any lymphoma effect, and the Zhao results are expected to be an
8 underestimate of any TCE effect on lymphoma alone. Another complication with the Zhao et al.
9 study is that no results for an overall TCE effect are reported. We were unable to obtain any
10 overall estimates from the study authors, so, as a best estimate, we combined the results across
11 the “medium” and “high” exposure groups, under assumptions of group independence, even
12 though the exposure groups are not independent (the “low” exposure group was the referent
13 group in both cases). Zhao et al. present RR estimates for both incidence and mortality;
14 however, the time frame for the incidence accrual is smaller than the time frame for mortality
15 accrual and fewer exposed incident cases (17) were obtained than deaths (33), so the mortality
16 results were used for the primary analysis, and the incidence results were used in a sensitivity
17 analysis. A sensitivity analysis was also done using results from Boice et al. (2006) in place of
18 the Zhao et al. RR estimate. The cohorts for these studies overlap, so they are not independent
19 studies and should not be included in the meta-analysis concurrently. Boice et al. (2006) report
20 results for an overall TCE effect for lymphoma alone; however, the results are based on far fewer
21 cases (1 death in ICD-9 200 + 202, 9 deaths for 200–208), so the Zhao et al. estimates are
22 preferred for the primary analysis.

23 For the case-control studies, the main issue was the lymphoma classifications. Miligi et
24 al. (2006) include chronic lymphocytic leukemias (CLLs) in their NHL results, consistent with
25 the current WHO/REAL classification. Also, Miligi et al. do not report an overall adjusted RR
26 estimate, so a crude estimate of the OR was calculated for the two TCE exposure categories
27 together versus no TCE exposure. The Nordstrom et al. (1998) study was a case-control study of
28 hairy cell leukemias (HCLs), which are a subgroup of NHLs, so only results for HCL were
29 reported. For Seidler et al. (2007), an overall adjusted OR for B-cell and T-cell NHL combined
30 was kindly provided by Dr. Seidler (personal communication from Andreas Seidler,
31 Bundesanstalt für Arbeitsschutz u. Arbeitsmedizin, to Cheryl Scott, U.S. EPA, 13 November
32 2007). No alternate RR estimates were considered for any of the case-control studies of
33 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 15 studies was 1.27 (95% CI 1.04, 1.53) (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 1.17 (with the removal of Hansen) to 1.33 (with the removal of Seidler) and were all statistically significant with the exception of the RRp estimate obtained by excluding Hansen, which just missed the standard cut-off for significance ($p = 0.051$). Removal of Hardell, whose RR estimate is a relative outlier (see Figure C.1), only decreased the RRp estimate to 1.23 (1.03, 1.47), 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 1.23.

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 1.22 to 1.29 (Table C.3) and were all statistically significant except when the Zhao incidence estimate ($p = 0.059$) or the Boice et al. (2006) SMR ($p = 0.050$) 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 versus 17). Furthermore, even though the Zhao mortality estimate is based on ICD 200–208 rather than 200 + 202 (which would likely underestimate any NHL risk) and was estimated by combining across exposure groups, as discussed above, it is preferred over the Boice (2006) estimate because the latter is an SMR rather than an internal analysis RR estimate and, in particular, because the Boice (2006) estimate is based on only *one* death for ICD 200 + 202 (the cohort had 9 deaths for ICD 200–208 compared to 33 in the Zhao cohort).

Heterogeneity across the 15 studies was statistically significant but just marginally ($p = 0.048$). 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 1.40 (95% CI 1.14, 1.71) for the cohort studies and 1.05 (0.77, 1.43) for the case-control studies. There was residual heterogeneity in each of the subgroups, but in neither case was it statistically significant. Nor was the difference between the RRp estimates for the cohort and case-control subgroups statistically significant under the random effects model, although it was under the fixed effect model (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 quality of exposure information or on likelihood for higher versus lower exposures across the studies. Ultimately, these approaches were rejected because in many of the studies it was difficult to judge (and weight) the quality of the information or the degree of TCE exposure with

1 any precision. See Section C.2.3 below for a qualitative discussion of some potential sources of
2 heterogeneity.

3 As discussed in Section C.1, publication bias was examined in several different ways.
4 The funnel plot in Figure C.2 suggests some relationship between RR estimate and study size (if
5 there were no relationship, the studies would be symmetrically distributed around the pooled RR
6 estimate rather than veering towards higher RR estimates with increasing SEs), although the
7 observed asymmetry is highly influenced by the Hardell study, which is a relative outlier and
8 which contributes little weight to the overall meta-analysis, as discussed above. The Begg and
9 Mazumdar rank correlation test and Egger's linear regression test were not statistically
10 significant; it should be noted, however, that both of these tests have low power. Duval and
11 Tweedie's trim-and-fill procedure yielded a pooled RR estimate (under the random effects
12 model) of 1.12 (95% CI 0.90, 1.38) when the 4 studies deemed missing from the funnel plot
13 were filled into the meta-analysis (these studies are filled in so as to counter-balance the apparent
14 asymmetry of the more extreme values in the funnel plot). Eliminating the Hardell study made
15 little difference to the results of the publication bias analyses. The results of a cumulative meta-
16 analysis, incorporating studies with increasing SE one at a time, are depicted in Figure C.3. This
17 procedure is a transparent way of examining the effects of including studies with increasing SE.
18 The figure shows that the pooled RR estimate is 1.07 after inclusion of the 4 largest (i.e., most
19 precise) studies, which constitute about 50% of the weight. The pooled RR estimate increases to
20 1.11 with inclusion of the 7 most precise studies, which represent about 70% of the weight. The
21 pooled RR estimate becomes fairly stable after addition of the next most precise study ($RR_p =$
22 1.22), which adds another 5% of the weight. Adding in the 7 least precise studies barely
23 increases the pooled RR estimate further, with the exception of the addition of Hardell, the least
24 precise study, which increases the pooled RR estimate from 1.23 to 1.27. In summary, there is
25 some evidence of potential publication bias in this dataset. It is uncertain, however, that this
26 reflects actual publication bias rather than an association between effect size and SE resulting for
27 some other reason, e.g., a difference in study populations or protocols in the smaller studies.
28 Furthermore, if there is publication bias in this dataset, it does not appear to account completely
29 for the findings of an increased lymphoma risk.

30 **C.2.2. Lymphoma Effect in the Highest Exposure Groups**

31 **C.2.2.1. *Selection of RR estimates***

32 The selected RR estimates for lymphoma in the highest TCE exposure categories, for
33 studies that provided such estimates, are presented in Table C.4. All 8 cohort studies (but not the
34 nested case-control study of Greenland et al. [1994]) and 3 of the 6 case-control studies did

1 obtain lymphoma risk estimates categorized by exposure level. As in section C.2.1.1 for the
2 overall risk estimates, we selected estimates to best correspond to NHL as represented by ICD-7,
3 -8, and -9 200 and 202, and, wherever possible, we used RR estimates for males and females
4 combined.

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

14 For Blair et al. (1998), it should be noted that the referent group is composed of workers
15 with no chemical exposures, not just no exposure to TCE. In addition, exposure group results
16 were reported separately for males and females and were combined for this assessment using
17 inverse-variance weighting, as in a fixed effect meta-analysis. Blair et al. present both incidence
18 and mortality RR estimates by exposure group, and there was the same number of incident cases
19 as deaths in the highest exposure category. The incidence estimate was used in the primary
20 analysis, because incidence estimates are generally preferred. In addition, the incidence estimate
21 was lower, so selecting this estimate avoids potential charges of “cherry-picking.” A mortality
22 RR estimate was used as an alternate estimate in a sensitivity analysis.

23 For Boice et al. (1999), only results for workers with “any potential exposure” (rather
24 than “potential routine exposure”) were presented by exposure category, and the referent group is
25 workers not exposed to any solvent. For Hansen et al. (2001), exposure group data were
26 presented only for males. To estimate the female contribution to a highest-exposure group RR
27 estimate for both sexes, it was assumed that the expected number of cases in females had the
28 same overall-to-highest-exposure group ratio as in males. The RR estimate for both sexes was
29 then calculated assuming a Poisson distribution, and this estimate was used in the primary
30 analysis. Hansen et al. present results for three exposure metrics; the cumulative exposure metric
31 was preferred for the primary analysis, and results for the other two metrics were used in
32 sensitivity analyses.

33 For Morgan et al. (1998), results did not allow for the combination of ICD 200 and 202,
34 so the highest-exposure group RR estimate for ICD 200 only was used. The primary analysis
35 used results for the cumulative exposure metric, and a sensitivity analysis was done with the
36 results for the peak exposure metric. For Zhao et al. (2005), RR estimates were only reported for

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

9 Miligi et al. (2006) include chronic lymphocytic leukemias (CLLs) in their NHL results,
10 consistent with the current WHO/REAL classification. Miligi et al. report RR estimates for
11 medium and high exposure intensity overall and by duration of exposure; however, there was
12 incomplete information for the duration breakdowns (e.g., a case missing), so the RR estimate
13 for med/high exposure intensity overall was used in the primary analysis, and the RR estimate
14 for med/high exposure for > 15 years was used in a sensitivity analysis. For Seidler et al. (2007),
15 an adjusted OR for B-cell and T-cell NHL combined for the >35 ppm-years exposure category
16 was kindly provided by Dr. Seidler (personal communication from Andreas Seidler,
17 Bundesanstalt für Arbeitsschutz u. Arbeitsmedizin, to Cheryl Scott, U.S. EPA, 13 November
18 2007).

19 **C.2.2.2. Results of meta-analyses**

20 Results from the meta-analyses that were conducted for lymphoma in the highest
21 exposure groups are summarized at the bottom of Table C.3 and reported in more detail in Table
22 C.5. The pooled RR estimate from the primary random effects meta-analysis of the 11 studies
23 with results presented for exposure groups was 1.50 (95% CI 1.20, 1.88) (see Figure C.4). No
24 single study was overly influential; removal of individual studies resulted in RRp estimates that
25 were all statistically significant and that ranged from 1.44 (with the removal of Raaschou-
26 Nielsen) to 1.58 (with the removal of Miligi).

27 Similarly, the RRp estimate was not highly sensitive to alternate RR estimate selections.
28 Use of the 7 alternate selections, individually, resulted in RRp estimates that were all statistically
29 significant and that ranged from 1.46 to 1.57, with all but one of the alternate selections yielding
30 an RRp estimate greater than or equal to the estimate from the primary analysis (Table C.5). The
31 lowest estimate, 1.46, was obtained using the Zhao incidence RR estimate rather than the
32 mortality estimate, and, in this case, there is a strong preference for using the mortality results
33 because of the underascertainment of incident cases discussed above.

34 There was no observable heterogeneity across the 11 studies in the primary analysis and
35 in all but one of the alternate RR analyses. When the Zhao incidence RR estimate was used,

1 there was some heterogeneity, but it was not statistically significant ($p = 0.34$). No subgroup
2 analyses were done with the highest exposure group results.

3 **C.2.3. Discussion of lymphoma meta-analysis results**

4 For the most part, the meta-analyses of the overall effect of TCE exposure on lymphoma
5 suggest a small, statistically significant increase in risk. The pooled estimate from the primary
6 random effects meta-analysis of the 15 studies was 1.27 (95% CI 1.04, 1.53). This result was not
7 overly influenced by any single study, nor was it overly sensitive to individual RR estimate
8 selections. In terms of the statistical significance of the RRp estimate, the largest impacts were
9 from the removal of the Hansen study, resulting in a RRp estimate of 1.17 (95% CI 1.00, 1.38),
10 and from the substitution of the Zhao mortality RR estimate with either the incidence estimate,
11 resulting in a RRp estimate of 1.22 (0.99, 1.49), or the Boice (2006) estimate, resulting in a RRp
12 estimate of 1.24 (1.00, 1.54); although, as noted above, these substitutions are considered clearly
13 inferior to the Zhao mortality estimate that was used in the primary analysis. Thus, the finding
14 of an increased risk of lymphoma associated with TCE exposure is somewhat robust, but of just
15 marginal statistical significance.

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

21 Heterogeneity across the 15 studies was marginally statistically significant ($p = 0.048$).
22 When subgroup analyses were done of cohort and case-control studies separately, there was
23 some observable heterogeneity in each of the subgroups, but it was not statistically significant in
24 either case. Thus, the differences between cohort and case-control studies explained much of the
25 heterogeneity, with the increased risk of lymphoma strengthened in the cohort study analysis and
26 virtually eliminated in the case-control study analysis. However, study design itself is unlikely
27 to be an underlying cause of heterogeneity and is more probably a surrogate for some other
28 difference(s) across studies that may be associated with study design. Furthermore, other
29 potential sources of heterogeneity may be masked by the broad study design subgroupings. The
30 true source(s) of heterogeneity across these studies is an uncertainty. As discussed above, further
31 quantitative investigations of heterogeneity were ruled out because of database limitations. A
32 qualitative discussion of some potential sources of heterogeneity follows.

33 Study differences in exposure assessment approach, exposure prevalence, average
34 exposure intensity, and lymphoma classification are possible sources of heterogeneity. Many

1 studies included TCE assignment from information on job and task exposures, e.g., a job-
2 exposure matrix (JEM) (Siemiatycki, 1991; Blair et al., 1998; Morgan et al., 1998; Boice et al.,
3 1999, 2006; Zhao et al., 2005; Miligi et al., 2006; Seidler et al., 2007), or from an exposure
4 biomarker in either breath or urine (Axelson et al., 1994; Anttila et al., 1995; Hansen et al.,
5 2001). Three case-control studies relied on self-reported exposure to TCE (Hardell et al., 1994;
6 Nordstrom et al., 1998; Persson and Fredrikson, 1999). Misclassification is possible with all
7 exposure assessment approaches. No information is available to judge the degree of possible
8 misclassification bias associated with a particular exposure assessment approach; it is quite
9 possible that in some cohort studies, in which past exposure is inferred from various data
10 sources, exposure misclassification may be as great as in population-based or hospital-based
11 case-control studies. Approaches based upon JEMs can provide order-of-magnitude estimates
12 that are useful for distinguishing groups of workers with large differences in exposure; however,
13 smaller differences usually cannot be reliably distinguished (NRC, 2006). The lack of
14 heterogeneity in the analysis of the highest exposure groups provides some evidence of exposure
15 misclassification as a source of heterogeneity in the overall analysis.

16 General population studies have special problems in evaluating exposure, because the
17 subjects could have worked in any job or setting that is present within the population (Copeland
18 et al., 1977; Nelson et al., 1994; McGuire et al., 1998; 't Mannetje et al., 2002; NRC, 2006).
19 Low exposure prevalence in the three population case-control studies (Siemiatycki, 1991; Miligi
20 et al., 2006; Seidler et al., 2007) may be another source of heterogeneity. Prevalence of TCE
21 exposure among cases in the case-control studies was low, ranging from 3% in Siemiatycki
22 (1991) to 13% in Seidler et al. (2007). However, prevalence of high TCE exposure in these case-
23 control studies was even rarer — 3% of all cases in Miligi et al. (2006) and Seidler et al. (2007),
24 and less than 1% in Siemiatycki (1991). Low exposure prevalence, especially in the relatively
25 large Miligi et al. (2006) and Seidler et al. (2007) case-control studies (see Figure C.1), may be
26 one of the underlying characteristics differentiating the case-control and cohort studies that is
27 reflected in the finding of study design as an apparent explanation of much of the heterogeneity
28 across the studies.

29 Study differences in lymphoma groupings and in lymphoma classification schemes are
30 another potential source of heterogeneity in the meta-analysis. All studies included a broad but
31 sometimes slightly different group of lymphosarcoma, reticulum-cell sarcoma, and other
32 lymphoid tissue neoplasms, with the exception of the Nordstrom et al. (1998) case-control study,
33 which examined hairy cell leukemia, now considered a lymphoma, and the Zhao et al. (2005)
34 cohort study, which reported only results for *all* lymphohematopoietic cancers, including non-
35 lymphoid types. Persson and Fredrikson (1999) do not identify the classification system for
36 defining NHL, and Hardell et al. (1999) define NHL using the Rappaport classification system.

1 Miligi et al. (2006) used an NCI classification system and considered chronic lymphocytic
2 leukemias and NHLs together as lymphomas, while Seidler et al. used the REAL classification
3 system, which reclassifies lymphocytic leukemias and NHLs as lymphomas of B-cell or T-cell
4 origin. The cohort studies (except for Zhao et al.) and the case-control study of Siemiatycki
5 (1991) have some consistency in coding NHL, with NHL defined as lymphosarcoma and
6 reticulum-cell sarcoma (ICD code 200) and other lymphoid tissue neoplasms (ICD 202) using
7 the ICD Revisions 7, 8, or 9. Revisions 7 and 8 are essentially the same with respect to NHL;
8 under Revision 9, the definition of NHL was broadened to include some neoplasms previously
9 classified as Hodgkin’s lymphomas (Banks, 1992).

10 Eleven of the 15 studies categorized results by exposure level. Different exposure
11 metrics were used, and the purpose of combining results across the different highest exposure
12 groups was not to estimate an RRp associated with some level of exposure, but rather to see the
13 impacts of combining RR estimates that should be less affected by exposure misclassification.
14 In other words, the highest exposure category is more likely to represent a greater differential
15 TCE exposure compared to people in the referent group than the exposure differential for the
16 overall (typically any versus none) exposure comparison. Thus, if TCE exposure increases the
17 risk of lymphoma, the effects should be more apparent in the highest exposure groups. Indeed,
18 the RRp estimate from the primary meta-analysis of the highest exposure group results was 1.50
19 (95% CI 1.20, 1.88), which is greater than the RRp estimate of 1.27 (95% CI 1.04, 1.53) from the
20 overall exposure analysis. This result for the highest exposure groups was not overly influenced
21 by any single study, nor was it overly sensitive to individual RR estimate selections.
22 Heterogeneity was not observed in any of the relevant analyses. The robustness of this finding
23 lends substantial support to a conclusion that TCE exposure increases the risk of lymphoma.

24 **C.3. Meta-analysis for Kidney Cancer**

25 **C.3.1. Overall effect of TCE exposure**

26 **C.3.1.1. *Selection of RR estimates***

27 The selected RR estimates for kidney cancer associated with TCE exposure from the
28 epidemiological studies are presented in Table C.6 for cohort studies and in Table C.7 for case-
29 control studies. The majority of the cohort studies reported results for all kidney cancers,
30 including cancers of the renal pelvis and ureter (i.e., ICD-7 180; ICD-8 and -9 189.0–189.2;
31 ICD-10 C64–C66); whereas the majority of the case-control studies focused on renal cell
32 carcinoma (RCC), which comprises roughly 85% of kidney cancers. Where both all kidney
33 cancer and RCC were reported, the primary analysis used the results for RCC, because RCC and

1 the other forms of kidney cancer are very different cancer types and it seemed preferable not to
2 combine them; the results for all kidney cancers were then used in a sensitivity analysis.

3 As for lymphoma, many of the studies provided RR estimates only for males and females
4 combined, and we are not aware of any basis for a sex difference in the effects of TCE on kidney
5 cancer risk; thus, wherever possible, we used RR estimates for males and females combined. Of
6 the three larger (in terms of number of cases) studies that did provide results separately by sex,
7 Dosemeci et al. (1999) suggest that there may be a sex difference for TCE exposure and RCC
8 (OR = 1.04 [95% CI 0.6, 1.7] in males and 1.96 [1.0, 4.0] in females), while Raaschou-Nielsen et
9 al. (2003) report the same SIR (1.2) for both sexes and crude ORs calculated from data from the
10 Pesch et al. (2000) study (provided in a personal communication from Baeta Pesch,
11 Forschungsinstitut für Arbeitsmedizin (BGFA), to Cheryl Scott, U.S. EPA, 21 February 2008)
12 are 1.28 for males and 1.23 for females. Blair et al. (1998) and Hansen et al. (2001) also present
13 some results by sex, but both of these studies have too few cases to be informative about a sex
14 difference for kidney cancer.

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

² person-years for men and women ≤ 79 y 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.

1 For Blair et al. (1998), it should be noted that the referent group is composed of workers
2 with no chemical exposures, not just no exposure to TCE. For Boice et al. (1999), only results
3 for “potential routine exposure” were reported for kidney cancer. This is our preferred TCE
4 exposure definition for the Boice study, because it was considered to have less exposure
5 misclassification than “any potential exposure;” however, since the results for the latter
6 definition were not presented, they could not be used in a sensitivity analysis, as was done for
7 lymphoma. Boice et al. report in general that the SMRs for workers with any potential exposure
8 “were similar to those for workers with daily potential exposure”. In their published paper,
9 Morgan et al. (1998) present only SMRs for overall TCE exposure, although the results from
10 internal analyses are presented for exposure subgroups. RR estimates for overall TCE exposure
11 from the internal analyses of the Morgan et al. cohort data were available from an unpublished
12 report (Environmental Health Strategies, 1997); from these, the RR estimate from the Cox
13 model, which included age, and sex was selected, because those are the variables deemed to be
14 important in the published paper. The internal analysis RR estimate was preferred for the
15 primary analysis, and the published SMR result was used in a sensitivity analysis. Raaschou-
16 Nielsen et al. (2003), reported results for RCC and renal pelvis/ureter separately. As discussed
17 above, RCC estimates were used in the primary analysis, and the results for both kidney cancer
18 categories were combined (across sexes as well), assuming a Poisson distribution, and used in a
19 sensitivity analysis.

20 For Zhao et al. (2005), no results for an overall TCE effect are reported. We were unable
21 to obtain any overall estimates from the study authors, so, as a best estimate, as was done for
22 lymphoma, we combined the results across the “medium” and “high” exposure groups, under
23 assumptions of group independence, even though the exposure groups are not independent (the
24 “low” exposure group was the referent group in both cases). Unlike for lymphoma, adjustment
25 for exposure to other carcinogens made a considerable difference, so kidney results were also
26 presented with this additional adjustment, with and without a 20-year lag. Estimates of RR with
27 this additional adjustment were selected over those without. In addition, a 20-year lag seemed
28 reasonable for kidney cancer, so the lagged estimates were preferred to the unlagged; unlagged
29 estimates were used in sensitivity analyses. Zhao et al. (2005) present RR estimates for both
30 incidence and mortality. Unlike for lymphoma, the number of exposed incident cases (10 with
31 no lag) was identical to the number of deaths, so there was no reason to prefer the mortality
32 results over the incidence results. (In fact, there were more exposed incident cases [10 versus 7]
33 after lagging.) However, the mortality results, which yield a lower RR estimate, were selected
34 for the primary analysis to avoid any appearance of “cherry-picking”, and incidence RR
35 estimates were used in sensitivity analyses. A sensitivity analysis was also done using results
36 from Boice et al. (2006) in place of the Zhao et al. RR estimate. The cohorts for these studies

1 overlap, so they are not independent studies and should not be included in the meta-analysis
2 concurrently. Boice et al. (2006) report results for an overall TCE effect for kidney cancer;
3 however, the results are SMR estimates rather than internal comparisons and are based on fewer
4 exposed deaths (7), so either Zhao et al. estimate is preferred over the Boice et al. estimate.

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

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

33 **C.3.1.2. *Results of meta-analyses***

34 Results from some of the meta-analyses that were conducted on the epidemiological
35 studies of TCE and kidney cancer are summarized in Table C.8. The pooled estimate from the

1 primary random effects meta-analysis of the 14 studies was 1.26 (95% CI 1.11, 1.42) (see Figure
2 C.5). As shown in Figure C.5, the analysis was dominated by 2 (contributing almost 70% of the
3 weight) or 3 (almost 80% of the weight) large studies. No single study was overly influential;
4 removal of individual studies resulted in pooled RR (RRp) estimates that were all statistically
5 significant and that ranged from 1.22 (with the removal of Brüning) to 1.28 (with the removal of
6 Raaschou-Nielsen).

7 Similarly, the RRp estimate was not highly sensitive to alternate RR estimate selections.
8 Use of the 10 alternate selections, individually, resulted in RRp estimates that were all
9 statistically significant and that ranged from 1.19 to 1.27 (Table C.8). In fact, as can be seen in
10 Table C.8, all but one of the alternates had negligible impact. The Zhao, Axelson, Brüning, and
11 Charbotel original values and alternate selections were associated with very little weight and,
12 thus, have little influence in the RRp. The Raaschou-Nielsen value carried more weight, but the
13 alternate RR estimate was identical to the original, although with a narrower CI, and so did not
14 alter the RRp. Only the Pesch alternate (with the JEM exposure assessment approach instead of
15 the JTEM approach) had much impact, resulting in an RRp estimate of 1.19 (95% CI 1.07, 1.32).
16 As noted above, the JTEM approach is preferred. The JEM approach takes jobs into account but
17 not tasks; thus, it is expected to have greater potential for exposure misclassification. Indeed, a
18 comparison of exposure prevalences for the 2 approaches suggests that the JEM approach is less
19 discriminating about exposure; 42% of cases were defined as TCE-exposed under the JEM
20 approach, but only 18% of cases were exposed under the JTEM approach.

21 There was no apparent heterogeneity across the 14 studies, i.e., the random effects model
22 and the fixed effect model gave the same results. Nonetheless, subgroup analyses were done
23 examining the cohort and case-control studies separately. With the random effects model studies
24 (and tau-squared not pooled across subgroups), the resulting RRp estimates were 1.16 (95% CI
25 0.96, 1.41) for the cohort studies and 1.41 (1.08, 1.83) for the case-control studies. There was
26 heterogeneity in the case-control subgroup, but it was not statistically significant ($p = 0.17$). Nor
27 was the difference between the RRp estimates for the cohort and case-control subgroups
28 statistically significant under either the random effects model or the fixed effect model. Further
29 quantitative investigations of heterogeneity were not pursued because of database limitations
30 and, in any event, there is no evidence for heterogeneity of study results in this database. A
31 qualitative discussion of some potential sources of heterogeneity across studies is nonetheless
32 included in Section C.3.3.

33 As discussed in Section C.1, publication bias was examined in several different ways.
34 The funnel plot in Figure C.6 shows little relationship between RR estimate and study size, and,
35 indeed, none of the other tests performed found any evidence of publication bias. Duval and
36 Tweedie's trim-and-fill procedure, for example, determined that no studies were missing from

1 the funnel plot, i.e., there was no asymmetry to counterbalance. Similarly, the results of a
2 cumulative meta-analysis, incorporating studies with increasing SE one at a time, shows no
3 evidence of a trend of increasing effect size with addition of the less precise studies. Including
4 the 3 most precise studies, reflecting 78.6% of the weight, the RRp goes from 1.24 to 1.22 to
5 1.23. The addition of the Brüning study brings the RRp to 1.32 and the weight to 82.6%. After
6 the addition of the next 5 studies, the RRp stabilizes at about 1.26, and further addition of the 5
7 least precise studies has little impact.

8 **C.3.2. Kidney cancer effect in the highest exposure groups**

9 **C.3.2.1. Selection of RR estimates**

10 The selected RR estimates for kidney cancer in the highest TCE exposure categories, for
11 studies that provided such estimates, are presented in Table C.9. Five of the 9 cohort studies and
12 4 of the 5 case-control studies reported kidney cancer risk estimates categorized by exposure
13 level. As in Section C.3.1.1 for the overall risk estimates, we preferentially selected estimates
14 for RCC when presented, and, wherever possible, we used RR estimates for males and females
15 combined.

16 Three of the 9 cohort studies (Anttila et al., 1995; Axelson et al., 1994; Hansen et al.,
17 2001) did not report kidney cancer risk estimates categorized by exposure level even though
18 these same studies reported such estimates for selected other cancer sites. To address this
19 reporting bias, we attempted to obtain the results from the primary investigators, and, failing that,
20 we performed alternate analyses in which we inserted null estimates (RR = 1.0) for all 3 studies.
21 We then used this alternate analysis as the main analysis, e.g. the basis of comparison for the
22 sensitivity analyses. For the SE (of the logRR) estimates, we used SE estimates from other sites
23 for which highest-exposure-group results were available. For Anttila et al. (1995), we used the
24 SE estimate for liver cancer in the highest exposure group, because liver cancer and kidney
25 cancer had similar numbers of cases in the overall study (5 and 6, respectively). For Axelson et
26 al. (1994), we used the SE estimate for NHL in the highest exposure group, because NHL and
27 kidney cancer had similar numbers of cases in the overall study (5 and 6, respectively). For
28 Hansen et al. (2001), we used the SE estimate for NHL in the highest exposure group, because
29 NHL was the only cancer site of interest in this assessment for which highest-exposure-group
30 results were available.

31 For Blair et al. (1998), it should be noted that the referent group is workers with no
32 chemical exposures, not just no TCE exposure. In addition, exposure group results were
33 reported separately for males and females and were combined for this assessment using inverse-
34 variance weighting, as in a fixed effect meta-analysis. Blair et al. present both incidence and

1 mortality RR estimates by exposure group, but there were more deaths (5) than incident cases (4)
2 in the highest exposure group, so the mortality RR estimate was used in the primary analysis.
3 The incidence RR estimate was used as an alternate estimate in a sensitivity analysis.

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

9 Zhao et al. (2005) present kidney cancer RR estimates adjusted for exposure to other
10 carcinogens, because, unlike for lymphoma, this adjustment made a considerable difference.
11 Estimates of RR with this additional adjustment were selected over those without. Furthermore,
12 the kidney results were presented with and without a 20-year lag. A 20-year lag seemed
13 reasonable for kidney cancer, so the lagged estimates were preferred to the unlagged; unlagged
14 estimates were used in sensitivity analyses. In addition, the incidence results reflect more cases
15 (4 with no lag) in the highest exposure group than do the mortality results (3), so the incidence
16 result (with the 20-year lag) was used for the primary analysis, and the unlagged incidence result
17 and the mortality results were used in a sensitivity analysis. Sensitivity analyses were also done
18 using results from Boice et al. (2006) in place of the Zhao et al. RR estimate. The cohorts for
19 these studies overlap, so they are not independent studies. Boice et al. (2006) report mortality
20 RR estimates for kidney cancer by years worked as a test stand mechanic, a job with potential
21 TCE exposure, and by a measure that weighted years with potential exposure from engine
22 flushing by the number of flushes each year. No results were presented for a third metric, years
23 worked with potential exposure to any TCE, because the Cox proportional hazards model did not
24 converge. The Boice et al. estimates are adjusted for years of birth and hire and for hydrazine
25 exposure.

26 For Charbotel et al. (2006), results from the analysis that considered “only job periods
27 with a good level of confidence for TCE exposure assessment” (Charbotel et al., Table 7) were
28 preferred, as these estimates would presumably be less influenced by exposure misclassification.
29 Estimates from the full study analysis were used in a sensitivity analysis. Additionally, the high
30 cumulative dose results were preferred, but the results for high cumulative dose + peaks were
31 included in sensitivity analyses. For Pesch et al. (2000), TCE results were presented for 2
32 different exposure assessments. As discussed above, we preferred the estimates using the JTEM
33 approach because they seemed to represent a more comprehensive exposure assessment;
34 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 1.61 (95% CI 1.27, 2.03) (see Figure C.7). 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 70% of the weight. 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 1.55 (1.24, 1.94) (see Figure C.8). The inclusion of these 3 additional studies contributed just under 8% of the total weight. No single study was overly influential; removal of individual studies resulted in RRp estimates that were all statistically significant and that ranged from 1.46 (with the removal of Raaschou-Nielsen) to 1.61 (with the removal of 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 and that ranged from 1.43 to 1.57, with all but 2 of the alternate selections yielding RRp estimates in the range of 1.52–1.57 (Table C.10). The lowest RRp estimates, 1.43 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 versus case-control studies) were done with the highest exposure group results.

1 C.3.3. Discussion of kidney cancer meta-analysis results

2 For the most part, the meta-analyses of the overall effect of TCE exposure on kidney
3 cancer suggest a small, statistically significant increase in risk. The pooled estimate from the
4 primary random effects meta-analysis of the 14 studies was 1.26 (95% CI 1.11, 1.42). Although
5 the analysis was dominated by 2–3 large studies that contribute 70–80% of the weight, the
6 pooled estimate was not overly influenced by any single study, nor was it overly sensitive to
7 individual RR estimate selections. The largest downward impacts were from the removal of the
8 Brüning study, resulting in a RRp estimate of 1.22 (95% CI 1.08, 1.38), and from the substitution
9 of the Pesch JTEM RR estimate with the RR estimate based on the JEM approach, resulting in a
10 RRp estimate of 1.19 (1.07, 1.32). Thus, the finding of an increased risk of kidney cancer
11 associated with TCE exposure is robust. Furthermore, there is no evidence of publication bias in
12 this dataset.

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

24 With respect to the nonsignificant heterogeneity in the 5 case-control studies, these
25 studies differ in TCE exposure potential to the underlying population from which case and
26 control subjects were identified, and this may be a source of some heterogeneity. Prevalence of
27 exposure to TCE among cases in these studies was 27% in Charbotel et al. (2006) (for high-
28 level-of-confidence jobs), 18% in Brüning et al. (2003) (for self-assessed exposure), 18% in
29 Pesch et al. (2000), 13% in Dosemeci et al. (1999) and 1% in Siemiatycki (1991). Both Brüning
30 et al. (2003) and Charbotel et al. (2006) are studies designed specifically to assess RCC and TCE
31 exposure. These studies were carried out in geographical areas with both a high prevalence and
32 a high degree of TCE exposure. Some information is provided in these and accompanying
33 papers to describe the nature of exposure, making it possible to estimate the order of magnitude
34 of exposure, even though there were no direct measurements (Cherrie et al., 2001; Brüning et al.,
35 2003; Fevotte et al., 2006). The Charbotel et al. (2006) study was carried out in the Arve Valley

1 region in France, where TCE exposure was through metal-degreasing activity in small shops
2 involved in the manufacturing of screws and precision metal parts (Fevotte et al., 2006).
3 Industrial hygiene data from shops in this area indicated high intensity TCE exposures of 100
4 ppm or higher, particularly from exposures from hot degreasing processes. Considering
5 exposure only from the jobs with a high level of confidence about exposure, 18% of exposed
6 cases were identified with high cumulative exposure to TCE. The source population in the
7 Brüning et al. (2003) study includes the Arnsberg region in Germany, which also has a high
8 prevalence of TCE exposure. A large number of small companies used TCE in metal degreasing
9 in small workrooms. Subjects in this study also described neurological symptoms previously
10 associated with higher TCE intensities. While subjects in the Brüning et al. (2003) study had
11 potential high TCE exposure intensity, average TCE exposure in this study is considered lower
12 than that in the Charbotel et al. (2006) study because the base population was enlarged beyond
13 the Arnsberg region to areas which did not have the same focus of industry.

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

23 Nine of the 14 studies categorized results by exposure level. Three other studies reported
24 results for other cancer sites by exposure level, but not kidney cancer; thus, to address this
25 reporting bias, null values (i.e., RR estimates of 1.0) were used for these studies. Different
26 exposure metrics were used in the various studies, and the purpose of combining results across
27 the different highest exposure groups was not to estimate an RRp associated with some level of
28 exposure, but rather to see the impacts of combining RR estimates that should be less affected by
29 exposure misclassification. In other words, the highest exposure category is more likely to
30 represent a greater differential TCE exposure compared to people in the referent group than the
31 exposure differential for the overall (typically any versus none) exposure comparison. Thus, if
32 TCE exposure increases the risk of kidney cancer, the effects should be more apparent in the
33 highest exposure groups. Indeed, the RRp estimate from the primary meta-analysis of the
34 highest exposure group results was 1.55 (95% CI 1.24, 1.94), which is greater than the RRp
35 estimate of 1.26 (95% CI 1.11, 1.42) from the overall exposure analysis. This result for the
36 highest exposure groups was not overly influenced by any single study, nor was it overly

1 sensitive to individual RR estimate selections. Heterogeneity was not observed in any of the
2 analyses. The robustness of this finding lends substantial support to a conclusion that TCE
3 exposure increases the risk of kidney cancer.

4 **C.4. Meta-analysis for Liver Cancer**

5 **C.4.1. Overall effect of TCE exposure**

6 **C.4.1.1. Selection of RR estimates**

7 The selected RR estimates for liver cancer associated with TCE exposure from the
8 epidemiological studies are presented in Table C.11. There were no case-control studies for liver
9 cancer and TCE exposure that were selected for inclusion in the meta-analysis (Appendix B,
10 Section II-9), so all of the relevant studies are cohort studies. All of the studies reported results
11 for liver cancers plus cancers of the gall bladder and extrahepatic biliary passages (i.e., ICD-7
12 155.0 + 155.2; ICD-8 and -9 155 + 156). Three of the studies also report results for liver cancer
13 alone (ICD-7 155.0; ICD-8 and -9 155). For the primary analysis, we selected results for cancers
14 of the liver, gall bladder, and biliary passages combined, for the sake of consistency, since these
15 were reported in all the studies. We also did an alternate analysis using results for liver cancer
16 alone for the 3 studies that reported them and the combined liver cancer results for the remainder
17 of the studies.

18 As for lymphoma and kidney cancer, many of the studies provided RR estimates only for
19 males and females combined, and we are not aware of any basis for a sex difference in the
20 effects of TCE on liver cancer risk; thus, wherever possible, we used RR estimates for males and
21 females combined. The only study of much size (in terms of number of liver cancer cases) that
22 provided results separately by sex was Raaschou-Nielsen (2003). The results of this study
23 suggest that liver cancer risk in females might be slightly higher than the risk in males, but the
24 number of female cases is small (liver cancer SIR: males 1.1 [95 % CI 0.74, 1.64; 27 cases],
25 females 2.8 [1.13, 5.80; 7 cases]; gallbladder and biliary passages cancer SIR: males 1.1 [0.61,
26 1.87; 14 cases]; females 2.8 [1.28, 5.34; 9 cases]).

27 Most of the selections in Table C.11 should be self-evident, but some are discussed in
28 more detail here, in the order the studies are presented in the Table. For Axelson et al. (1994), in
29 which a small subcohort of females was studied but only results for the larger male subcohort
30 were reported, the reported male-only results were used in the primary analysis; however, as for
31 lymphoma and kidney cancer, an attempt was made to estimate the female contribution to an
32 overall RR estimate for both sexes and its impact on the meta-analysis. Axelson et al. reported
33 that there were no cases of liver cancer observed in females, but the expected number was not

1 presented. To estimate the expected number, the expected number for males was multiplied by
2 the ratio of female-to-male person-years in the study and by the ratio of female-to-male age-
3 adjusted incidence rates for liver cancer. The male results and the estimated female contribution
4 were then combined into an RR estimate for both sexes assuming a Poisson distribution, and this
5 alternate RR estimate for the Axelson et al. study was used in a sensitivity analysis.

6 For Blair et al. (1998), it should be noted that the referent group is workers with no
7 chemical exposures, not just no TCE exposure. For Boice et al. (1999), results for “potential
8 routine exposure” were selected for the primary analysis, because this exposure category was
9 considered to have less exposure misclassification, and results for “any potential exposure” were
10 used in a sensitivity analysis. To estimate the SE(logRR) for the alternate RR selection, it was
11 assumed that the number of exposed cases (deaths) was 15. The actual number was not
12 presented, but 15 was the number that allowed us to reproduce the reported CIs. The number
13 suggested by exposure level in Boice et al. (1999) Table 9 is 13; however, it may be that
14 exposure level data were not available for all the cases. In their published paper, Morgan et al.
15 (1998) present only SMRs for overall TCE exposure, although the results from internal analyses
16 are presented for exposure subgroups. RR estimates for overall TCE exposure from the internal
17 analyses of the Morgan et al. cohort data were available from an unpublished report
18 (Environmental Health Strategies, 1997); from these, the RR estimate from the Cox model,
19 which included age, and sex was selected, because those are the variables deemed to be
20 important in the published paper. The internal analysis RR estimate was preferred for the
21 primary analysis, and the published SMR result was used in a sensitivity analysis.

22 Zhao et al. (2005) did not present RR estimates for liver cancer; thus, results from Boice
23 et al. (2006) were used in the primary analysis. The cohorts for these studies overlap, so they are
24 not independent studies. Zhao et al., however, was our preferred study for lymphoma and kidney
25 cancer results; thus, in a sensitivity analysis, a null value (RR = 1.0) was assumed for Zhao et al.
26 to address the potential reporting bias. The SE estimate for kidney cancer (incidence with 0 lag)
27 was used as the SE for the liver cancer. (It is not certain that there was a reporting bias in this
28 case. In the “Methods” section of their paper, Zhao et al. list the cancer sites examined in the
29 cohort, and liver was not listed; it is not clear if the list of sites was determined a priori or post
30 hoc). Also, on the issue of potential reporting bias, the Siemiatycki (1991) study should be
31 mentioned. This study was a case-control study for multiple cancer sites, but only the more
32 common sites, in order to have greater statistical power. Thus, NHL and kidney cancer results
33 were available, but not liver cancer results. Because no liver results were presented for any of
34 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 1.36 (95% CI 1.10, 1.67) (see Figure C.9). As shown in Figure C.9, the analysis was dominated by one large study (contributing almost 60% 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 = 1.36; 95% CI 0.98, 1.89). 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 1.31 (with the removal of Anttila) to 1.42 (with the removal of Boice [1999]).

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 (1.32; 95% CI 1.02, 1.70). 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 and that ranged from 1.30 to 1.35 (Table C.12). In fact, as can be seen in Table C.12, only one of the alternates had notable impact. The Boice et al. (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 et al. (1999) alternate RR estimate based on any potential exposure rather than potential routine exposure decreased the RRp slightly from 1.36 to 1.30. The alternate Boice et al. (1999) RR estimate is actually larger than the original value (0.81 versus 0.54); however, use of the less discriminating exposure metric captures more liver cancer cases (deaths), causing the weight of that study to increase from about 4.5% to about 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.

1 As discussed in Section C.1, publication bias was examined in several different ways.
2 The funnel plot in Figure C.10 shows little relationship between RR estimate and study size, and,
3 indeed, none of the other tests performed found any evidence of publication bias. Duval and
4 Tweedie’s trim-and-fill procedure, for example, suggested that no studies were missing from the
5 funnel plot, i.e., there was no asymmetry to counterbalance. Similarly, the results of a
6 cumulative meta-analysis, incorporating studies with increasing SE one at a time, shows no
7 evidence of a trend of increasing effect size with addition of the less precise studies. The
8 Raaschou–Nielsen study contributes 59.4% of the weight. Including the 2 next most precise
9 studies, the RRp goes from 1.35 to 1.42 to 1.46 and the weight to 75.3%. Further addition of the
10 6 least precise studies gradually brings the RRp back down to 1.36. Thus, if anything, the
11 evidence is somewhat suggestive of an *inverse* relationship between SE and effect size, contrary
12 to what would be expected if publication bias were occurring.

13 **C.4.2. Liver cancer effect in the highest exposure groups**

14 **C.4.2.1. *Selection of RR estimates***

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

21 Two of the 9 cohort studies (Hansen et al., 2001; Zhao et al., 2005) did not report liver
22 cancer risk estimates categorized by exposure level even though these same studies reported such
23 estimates for selected other cancer sites. (As discussed above, Zhao et al. did not present any
24 liver results, and it is not clear if this was reporting bias or an a priori decision not to examine
25 liver cancer in the cohort.) To address this reporting bias, we attempted to obtain the results
26 from the primary investigators, and, failing that, we performed alternate analyses in which we
27 inserted null estimates ($RR = 1.0$) for both studies. We then used this alternate analysis as the
28 main analysis, e.g. the basis of comparison for the sensitivity analyses. For the SE (of the
29 logRR) estimates, we used SE estimates from other sites for which highest-exposure-group
30 results were available. For Hansen et al. (2001), we used the SE estimate for NHL in the highest
31 exposure group, because NHL was the only cancer site of interest in this assessment for which
32 highest-exposure-group results were available. For Zhao et al. (2005), the SE estimate for
33 kidney cancer in the highest-exposure group (incidence with 0 lag) was used. (Note that Boice et

1 al. [2006], who studied a cohort that overlapped that of Zhao et al., also did not present liver
2 cancer results by exposure level.)

3 For Axelson et al. (1994), there were no liver cancer cases in the highest exposure group
4 (≥ 2 years and 100+ mean U-TCA level), so no RR estimate was available for the meta-analysis.
5 We decided to combine the results for the <2 years and ≥ 2 years results, assuming expected
6 numbers of cases were proportional to person-years, and use all exposure times with 100+ U-
7 TCA as the highest exposure category. We also estimated the female contribution to the
8 expected number, again assuming proportionality to person-years, and adjusting for the
9 difference between female and male age-adjusted liver cancer incidence rates. We used the
10 estimated RR and SE values for the combined exposure times and sexes in the primary analysis.
11 In an alternate analysis, we excluded the Axelson et al. study altogether, because we estimated
12 that less than 0.2 cases were expected in the highest-exposure category, suggesting that the study
13 had low power to detect an effect in the highest-exposure group and would contribute little
14 weight to the meta-analysis.

15 For Blair et al. (1998), it should be noted that the referent group is workers with no
16 chemical exposures, not just no TCE exposure. In addition, exposure group results were
17 reported separately for males and females and were combined for this assessment using inverse-
18 variance weighting, as in a fixed effect meta-analysis. Blair et al. present both incidence and
19 mortality RR estimates by exposure group; however, there were no incident cases for females in
20 the highest-exposure group (and the expected number was not reported), and there were more
21 liver cancer deaths (21) than incident cases (13) overall and in the highest-exposure group (5
22 versus 4). Thus, we elected to use only the mortality results from this study.

23 For Boice et al. (1999), only results for workers with “any potential exposure” (rather
24 than “potential routine exposure”) were presented by exposure category, and the referent group is
25 workers not exposed to any solvent. For Morgan et al. (1998), the primary analysis used results
26 for the cumulative exposure metric, and a sensitivity analysis was done with the results for the
27 peak exposure metric. For Raaschou-Nielsen et al. (2003), unlike for NHL and RCC, liver
28 cancer results for the subcohort with expected higher exposure levels were not presented, so the
29 only highest-exposure group results were for duration of employment in the total cohort. We
30 used results for cancers of the liver and gall bladder/biliary passages combined for the primary
31 analysis and results for liver cancer alone in a sensitivity analysis.

32 **C.4.2.2. *Results of meta-analyses***

33 Results from the meta-analyses that were conducted for liver cancer in the highest
34 exposure groups are summarized at the bottom of Table C.12. The pooled RR estimate from the
35 random effects meta-analysis of the 6 studies with results presented for exposure groups was

1 1.25 (95% CI 0.87, 1.79). As with the overall liver cancer meta-analyses, the meta-analyses of
2 the highest-exposure groups were dominated by one study (Raaschou-Nielsen), which provided
3 almost 60% of the weight. The RRp estimate from the primary random effects meta-analysis
4 with null RR estimates (i.e., 1.0) included for Hansen and Zhao to address (potential) reporting
5 bias (see above) was 1.22 (0.87, 1.71) (see Figure C.11). The inclusion of these 2 additional
6 studies contributed about 10% of the total weight. No single study was overly influential
7 (removal of individual studies resulted in RRp estimates that ranged from 1.16 to 1.24) and the
8 RRp estimate was not highly sensitive to alternate RR estimate selections (RRp estimates with
9 alternate selections ranged from 1.18 to 1.20; Table C.12). In addition, there was no observable
10 heterogeneity across the studies for any of the meta-analyses conducted with the highest-
11 exposure groups. However, none of the RRp estimates was statistically significant.

12 Furthermore, the RRp estimates for the highest-exposure groups were all less than the
13 RRp estimate for an overall effect on liver cancer (Section C.4.2.2 and Table C.12). This
14 anomalous result is driven by the fact that the RR estimates for the highest-exposure groups were
15 less than the overall RR estimates for Blair and, especially, Raaschou-Nielsen, which contributes
16 the majority of the weight to the meta-analyses. The liver cancer results are relatively under-
17 powered with respect to numbers of studies and number of cases, and the Raaschou-Nielsen
18 study, which dominates the analysis, uses duration of employment as an exposure-level surrogate
19 for liver cancer, and duration of employment is a notoriously weak exposure metric. Thus, the
20 anomalous finding that the RRp estimates for the highest-exposure groups were all less than the
21 RRp estimate for an overall effect does not rule out an effect of TCE on liver cancer; however, it
22 certainly does not provide additional support for such an effect.

23 **C.4.3. Discussion of liver cancer meta-analysis results**

24 For the most part, the meta-analyses of the overall effect of TCE exposure on liver (and
25 gall bladder/biliary passages) cancer suggest a small, statistically significant increase in risk.
26 The pooled estimate from the primary random effects meta-analysis of the 9 (all cohort) studies
27 was 1.36 (95% CI 1.10, 1.67). The analysis was dominated by one large study that contributed
28 almost 60% of the weight. When this study was removed, the RRp estimate did not change, but
29 it was no longer statistically significant (RRp = 1.36; 95% CI 0.98, 1.89). The pooled estimate
30 was not overly influenced by any other single study, nor was it overly sensitive to individual RR
31 estimate selections. The largest downward impacts were from the removal of the Anttila study,
32 resulting in a RRp estimate of 1.31 (95% CI 1.05, 1.63), and from the substitution of the Boice
33 (1999) RR estimate for potential routine exposure with that for any potential exposure, resulting
34 in a RRp estimate of 1.30 (1.07, 1.58). Substituting RR estimates for liver cancer alone for the 3

1 studies that provided these results yielded an RRp estimate of 1.32 (1.02, 1.70). There was no
2 evidence of publication bias in this dataset, and there was no observable heterogeneity across the
3 study results.

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

19 Thus, while there is the suggestion of an increased risk for liver cancer associated with
20 TCE exposure, the statistical significance of the pooled estimates is dependent on one study,
21 which provides the majority of the weight in the meta-analyses. Removal of this study does not
22 change the RRp estimate; however, it becomes minimally nonsignificant ($p = 0.064$).
23 Furthermore, meta-analysis results for the highest-exposure groups yielded *lower* RRp estimates
24 than for an overall effect. These results do not rule out an effect of TCE on liver cancer, because
25 the liver cancer results are relatively under-powered with respect to numbers of studies and
26 number of cases and the overwhelming study in terms of weight uses the weak exposure
27 surrogate of duration of employment for categorizing exposure level; however, at present, there
28 is only modest support for such an effect.

29 **C.5. Discussion of strengths, limitations, and uncertainties in the meta-analyses**

30 Meta-analysis provides a systematic way of objectively and quantitatively combining the
31 results of multiple studies to obtain a summary effect estimate. Use of meta-analysis can help
32 risk assessors avoid some of the potential pitfalls in overly relying on a single study or in making
33 more subjective qualitative judgments about the apparent weight of evidence across studies.
34 Combining the results of smaller studies also increases the statistical power to observe an effect,

1 if one exists. In addition, meta-analysis techniques assist in systematically investigating issues
2 such as potential publication bias and heterogeneity in a database.

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

11 Identifying all relevant studies may be hampered if publication bias has occurred.
12 Publication bias is a systematic error that can arise if statistically significant studies are more
13 likely to be published than nonsignificant studies. This can result in an upward bias on the effect
14 size measure, i.e., the relative risk estimate. To address this concern, potential publication bias
15 was investigated for the databases for which meta-analyses were undertaken. For the studies of
16 kidney cancer and liver cancer, there was no evidence of publication bias. For the studies of
17 lymphoma, there was some evidence of potential publication bias. It is uncertain whether this
18 reflects actual publication bias or rather an association between SE and effect size (as discussed
19 in Section C.1, a feature of publication bias is that smaller studies tend to have larger effect
20 sizes) resulting for some other reason, e.g., a difference in study populations or protocols in the
21 smaller studies. Furthermore, if there is publication bias in this dataset, it may be creating an
22 upward bias on the relative risk estimate, but this bias does not appear to completely account for
23 the finding of an increased lymphoma risk (see Section C.2.1.2).

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

33 For the studies of kidney cancer and liver cancer, there was no apparent heterogeneity
34 across the study results, i.e., random- and fixed-effects models gave identical summary
35 estimates. For the lymphoma studies, there was statistically significant heterogeneity ($p =$
36 0.048). When subgroup analyses were done for the cohort and case-control studies separately,

1 there was some residual heterogeneity in both groups, but in neither case was it statistically
2 significant. Further attempts to quantitatively investigate the heterogeneity were not pursued
3 because of limitations in the database. The sources of heterogeneity are an uncertainty in the
4 database of studies of TCE and lymphoma. Some potential sources of heterogeneity, which are
5 discussed qualitatively in Section C.2.3, include differences in exposure assessment or in the
6 intensity or prevalence of TCE exposures in the study population and differences in lymphoma
7 classification.

8 The joint occurrence of heterogeneity and potential publication bias in the database of
9 studies of TCE and lymphoma raises special concerns. Because of the heterogeneity, a random-
10 effects model should be used if these studies are to be combined; yet, the random-effects model
11 gives relatively large weight to small studies, which could exacerbate the potential impacts of
12 publication bias. For the lymphoma studies, the summary relative risk estimates from the
13 random- and fixed-effects models are not very different ($RR_p = 1.27$ [95% CI 1.04, 1.53] and
14 1.20 [1.06, 1.36], respectively); however, the confidence interval for the fixed-effect estimate
15 does not reflect the between-study variance and is thus overly narrow.

16 **C.6. Conclusions**

17 The strongest finding from the meta-analyses was for TCE and kidney cancer. The
18 summary estimate from the primary random-effects meta-analysis of the 14 studies was $RR_p =$
19 1.26 (95% CI 1.11, 1.42). There was no apparent heterogeneity across the study results (i.e.,
20 fixed-effect model gave same summary estimate), and there was no evidence of potential
21 publication bias. The summary estimate was robust across influence and sensitivity analyses; the
22 estimate was not markedly influenced by any single study, nor was it overly sensitive to
23 individual RR estimate selections. The findings from the meta-analyses of the highest exposure
24 groups for the studies that provided results categorized by exposure level were similarly robust.
25 The summary estimate was $RR_p = 1.55$ (95% CI 1.24, 1.94) for the 12 studies included in the
26 analysis. There was no apparent heterogeneity in the highest-exposure group results, and the
27 estimate was not markedly influenced by any single study, nor was it overly sensitive to
28 individual RR estimate selections. In sum, these robust results support a conclusion that TCE
29 exposure increases the risk of kidney cancer.

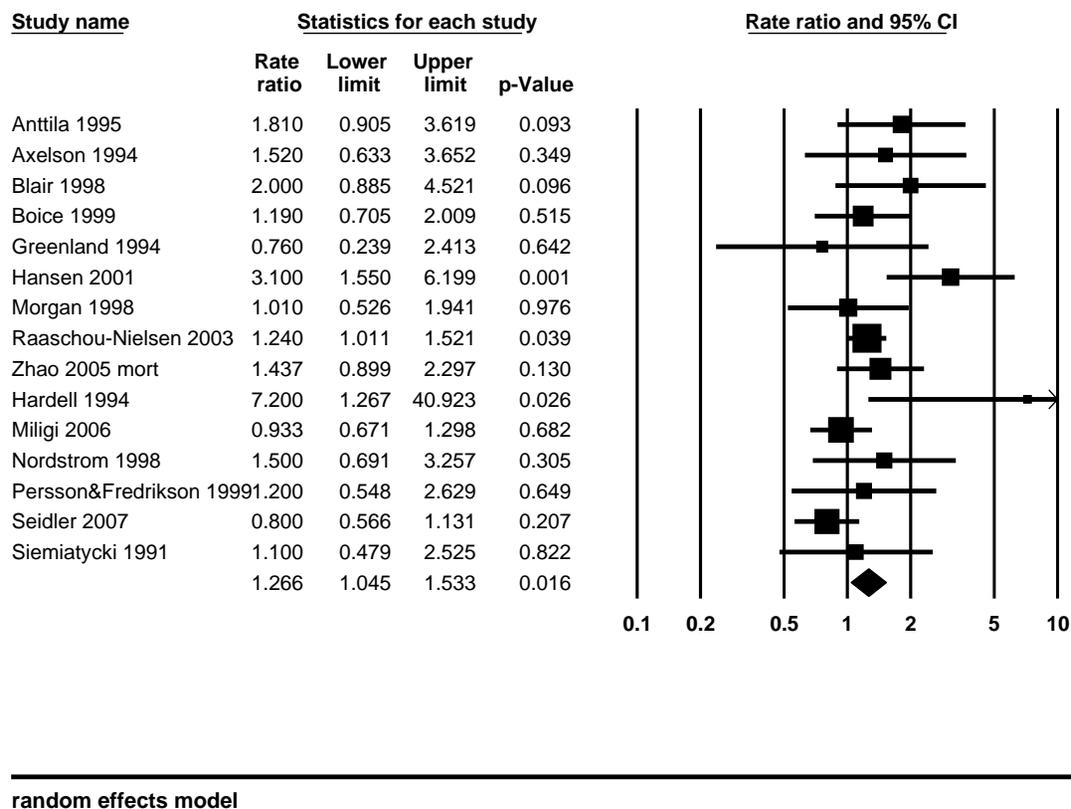
30 For the most part, the meta-analyses of the overall effect of TCE exposure on lymphoma
31 also suggest a small, statistically significant increase in risk. The summary estimate from the
32 primary random-effects meta-analysis of the 15 studies was 1.27 (95% CI 1.04, 1.53). This
33 result was not overly influenced by any single study, although the removal of one particular
34 study narrowly eliminated the statistical significance of the increased risk ($p = 0.051$). Nor was

1 the result overly sensitive to individual RR estimate selections, although use of a couple alternate
2 RR estimates considered clearly inferior also narrowly eliminated statistical significance of the
3 summary estimate. There is some evidence of potential publication bias in the lymphoma study
4 dataset; however, it is uncertain that this is actually publication bias rather than an association
5 between SE and effect size resulting for some other reason, e.g., a difference in study
6 populations or protocols in the smaller studies. Furthermore, if there is publication bias, it does
7 not appear to completely account for the findings of an increased lymphoma risk. There was
8 also statistically significant ($p = 0.048$) heterogeneity across the results of the 15 studies, and the
9 source(s) of this heterogeneity remains an uncertainty. The summary estimate from the meta-
10 analysis of the highest exposure groups for the 11 studies which provided results categorized by
11 exposure level was $RR_p = 1.50$ (95% CI 1.20, 1.88). This result for the highest exposure groups
12 was not overly influenced by any single study, nor was it overly sensitive to individual RR
13 estimate selections, and heterogeneity was not observed in any of the relevant analyses. The
14 robustness of the finding of an increased lymphoma risk for the highest exposure groups
15 strengthens the more moderate evidence from the meta-analyses for overall effect.

16 The meta-analyses of the overall effect of TCE exposure on liver (and gall bladder/biliary
17 passages) cancer also suggest a small, statistically significant increase in risk, but the study
18 database is more limited. The pooled estimate from the primary random-effects meta-analysis of
19 the 9 (all cohort) studies was 1.36 (95% CI 1.10, 1.67). The analysis was dominated by one
20 large study that contributed almost 60% of the weight. When this study was removed, the RR_p
21 estimate did not change, but it was less precise ($RR_p = 1.36$; 95% CI 0.98, 1.89). The pooled
22 estimate was not overly influenced by any other single study, nor was it overly sensitive to
23 individual RR estimate selections. There was no evidence of publication bias in this dataset, and
24 there was no observable heterogeneity across the study results. However, the findings from the
25 meta-analyses of the highest exposure groups for the studies that provided results categorized by
26 exposure level do not add support to the overall effect findings. The summary estimate was RR_p
27 $= 1.22$ (95% CI 0.87, 1.71) for the 8 studies included in the analysis, which is *lower* than the
28 summary estimate for the overall effect. This anomalous result is driven by the fact that the RR
29 estimate for the highest-exposure group in the individual study which contributes the majority of
30 the weight to the meta-analyses, although greater than 1.0, was less than the overall RR estimate
31 for the same study. In sum, these results do not rule out an effect of TCE on liver cancer,
32 because the liver cancer results are relatively under-powered with respect to numbers of studies
33 and number of cases and the overwhelming study in terms of weight uses the weak exposure
34 surrogate of duration of employment for categorizing exposure level; however, at present, there
35 is only modest support for an increased risk of liver cancer.

1

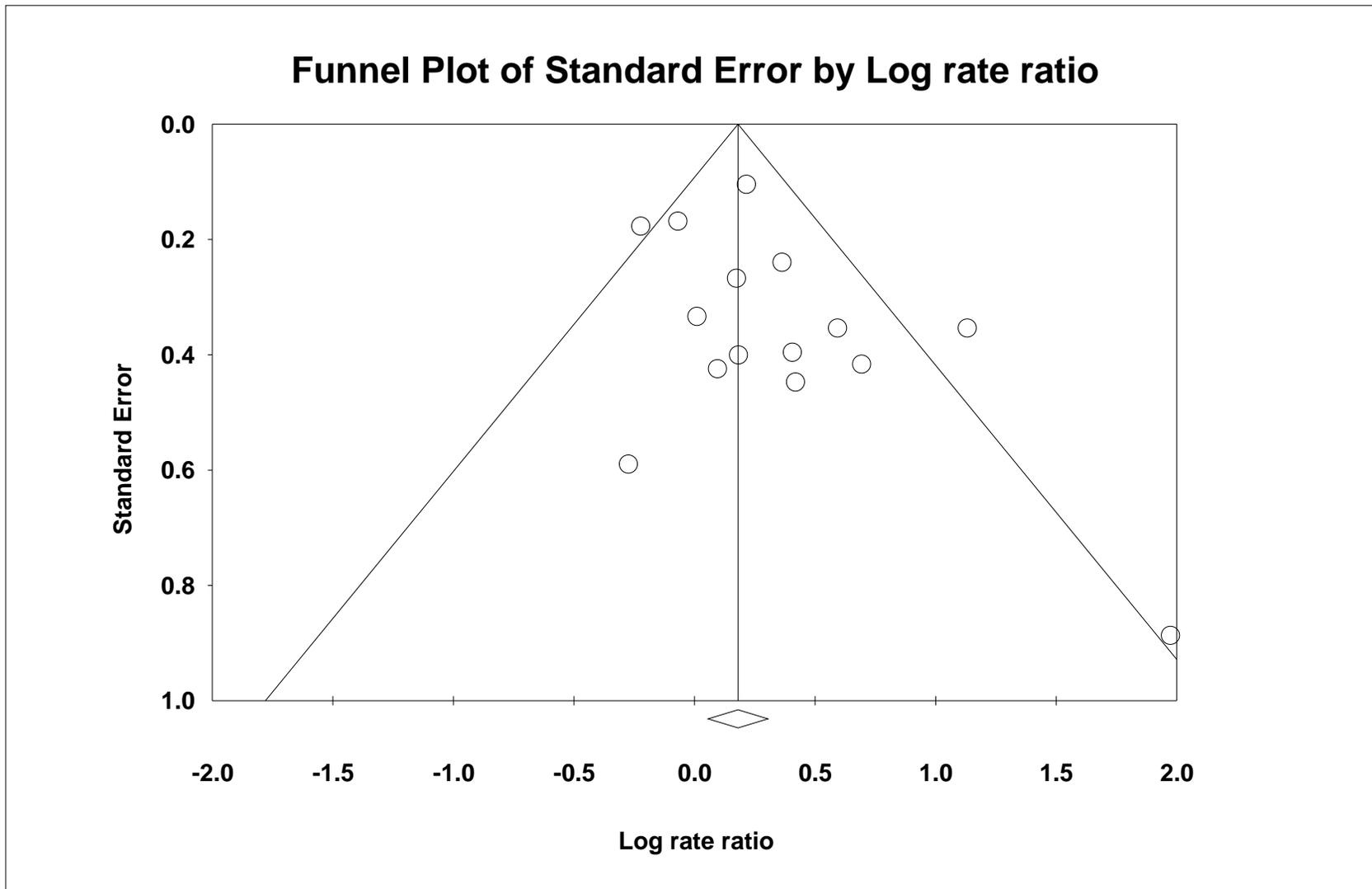
TCE and lymphoma



2

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

1

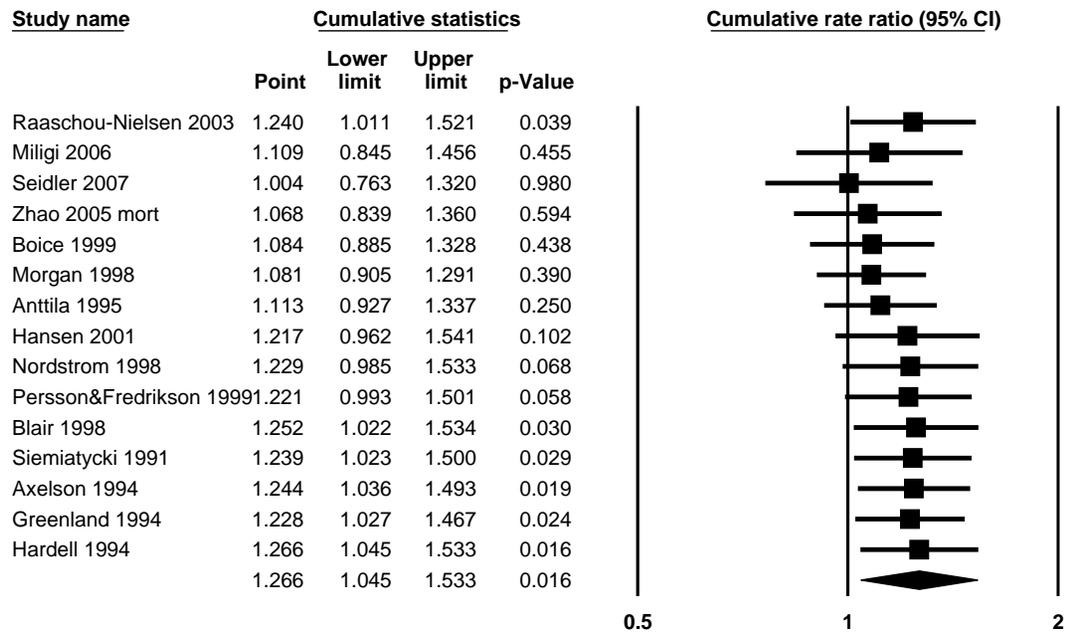


2

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

4

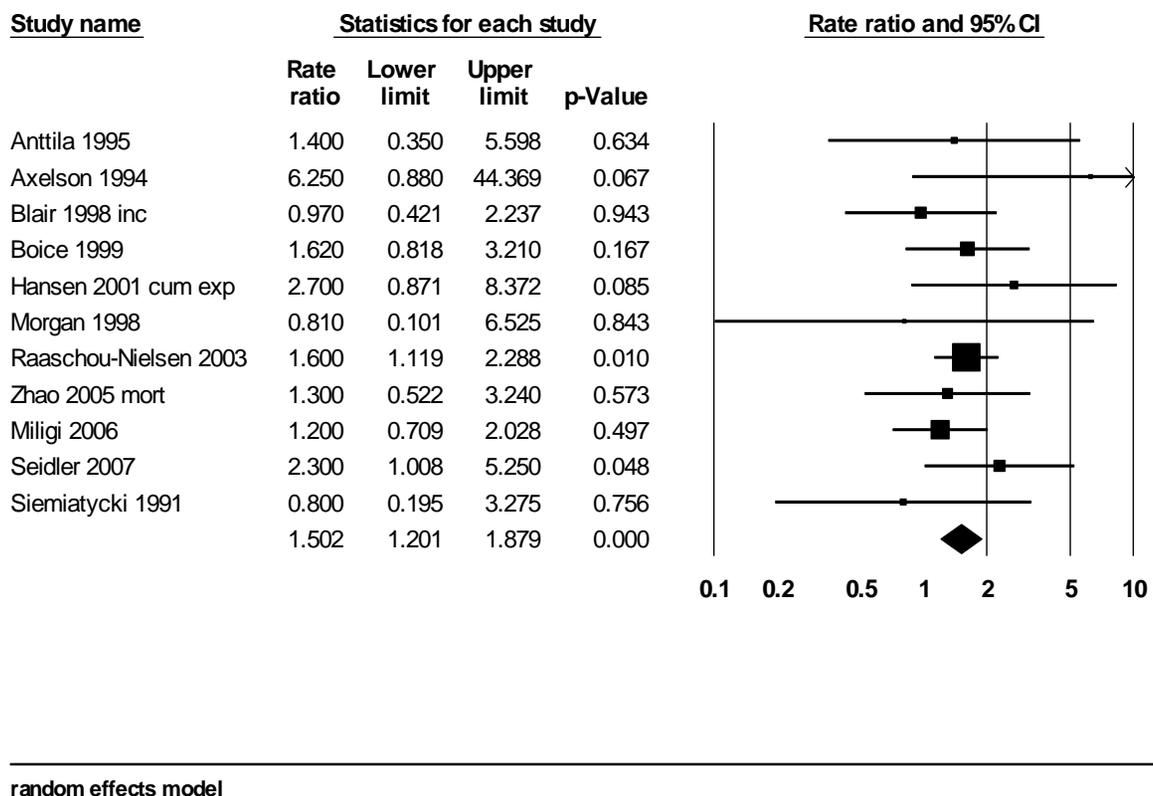
TCE and lymphoma



random effects model; cumulative analysis, sorted by SE

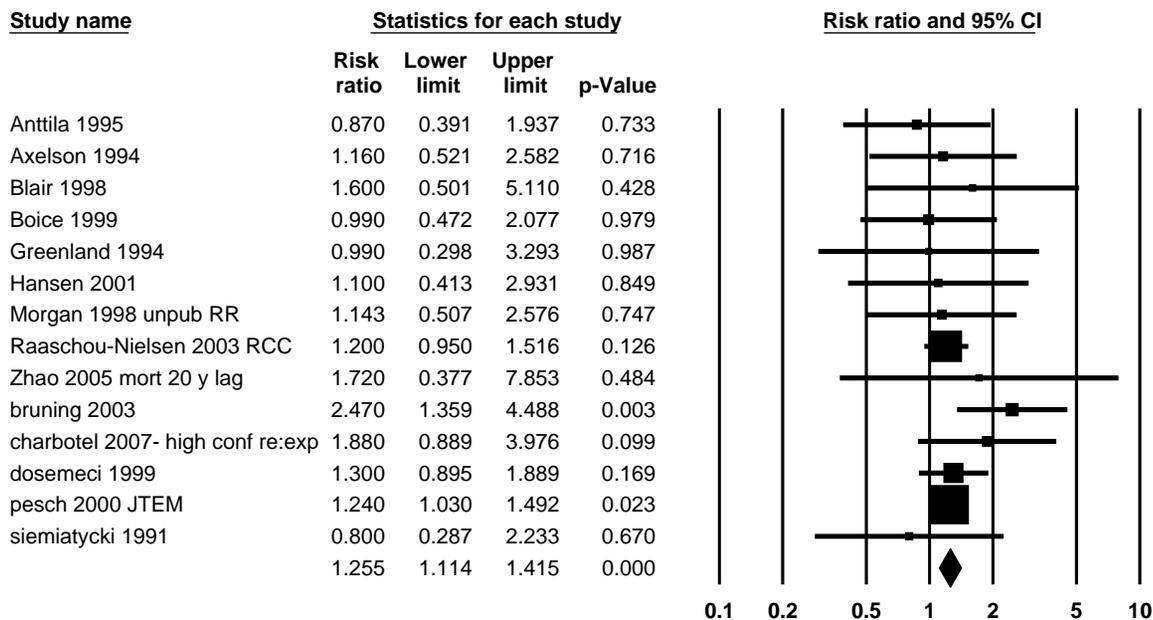
- 1
- 2 Figure C.3. Cumulative meta-analysis of TCE and lymphoma studies, progressively including studies with increasing SEs.

TCE and lymphoma - highest exposure groups



1
 2 Figure C.4. Meta-analysis of lymphoma and TCE exposure—highest exposure groups. (The pooled estimate is in the bottom row.
 3 Symbol sizes reflect relative weights of the studies. The horizontal midpoint of the bottom diamond represents the pooled RR
 4 estimate, and the horizontal extremes depict the 95% CI limits.)
 5

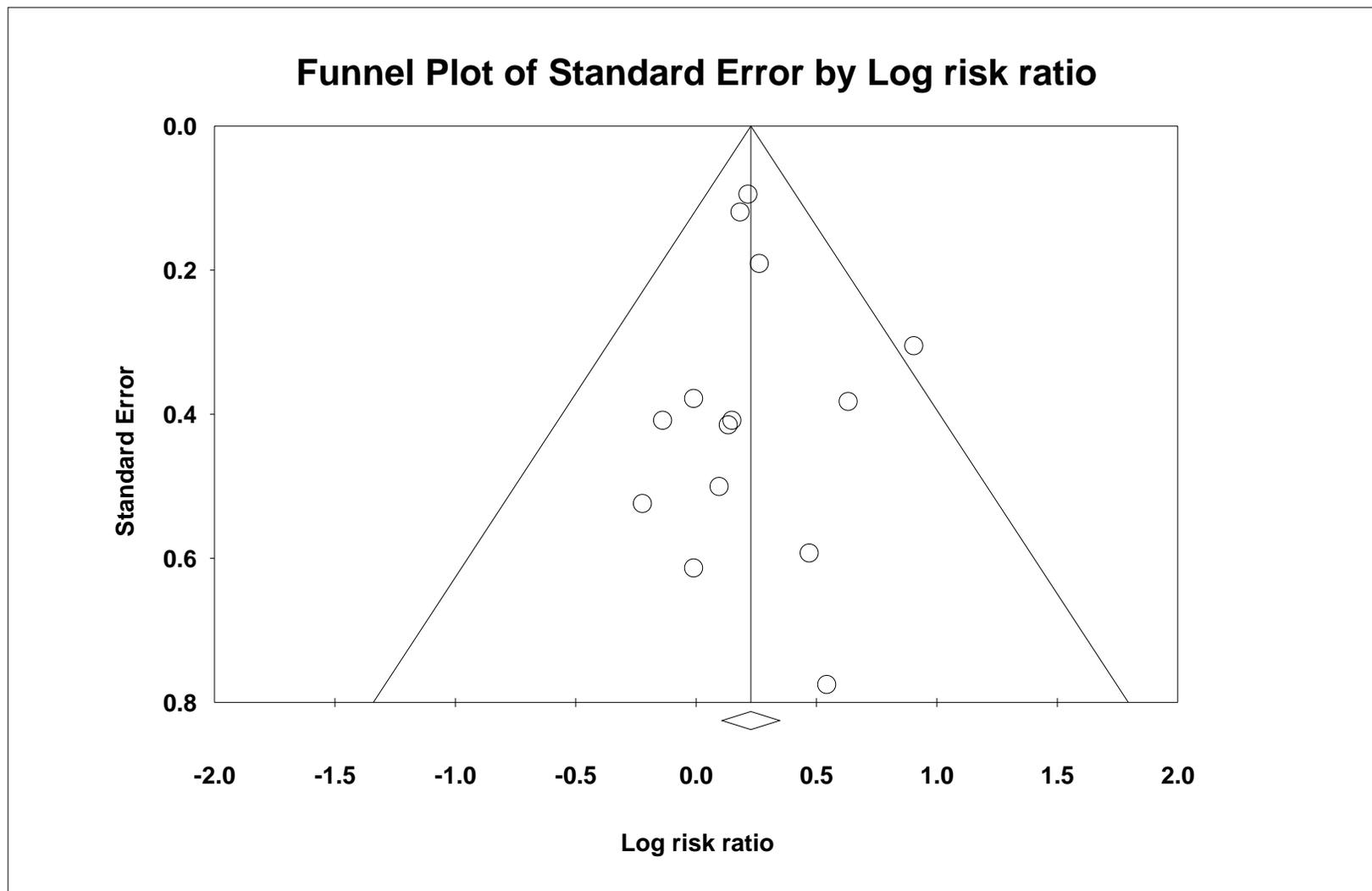
TCE and kidney cancer



random effects model; same for fixed

1
 2 Figure C.5. Meta-analysis of kidney cancer and overall TCE exposure. The pooled estimate is in the bottom row. Symbol sizes
 3 reflect relative weights of the studies. The horizontal midpoint of the bottom diamond represents the pooled RR estimate and the
 4 horizontal extremes depict the 95% CI limits.
 5

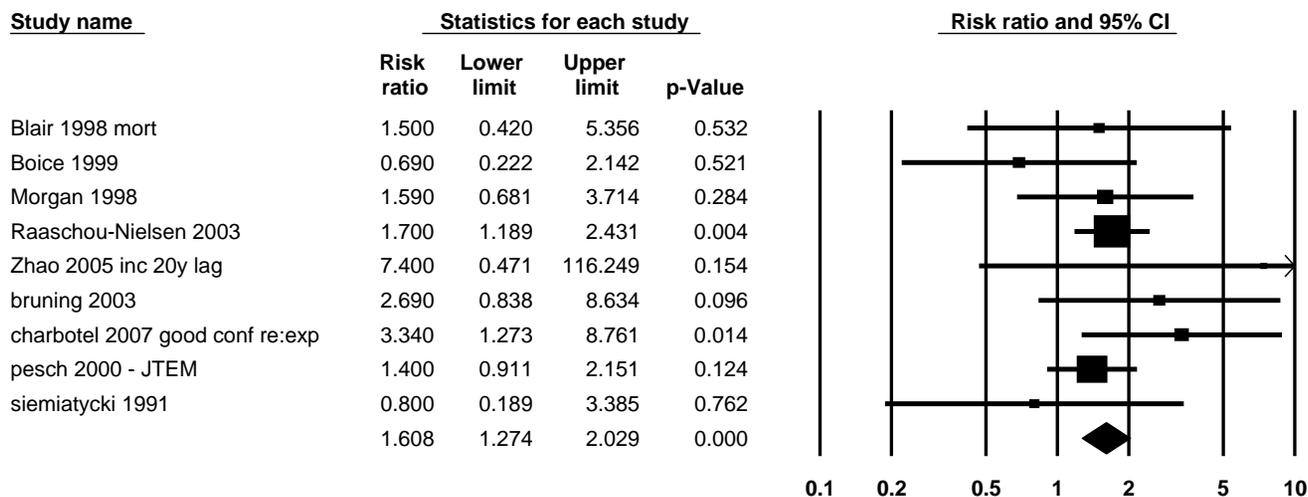
1



2

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

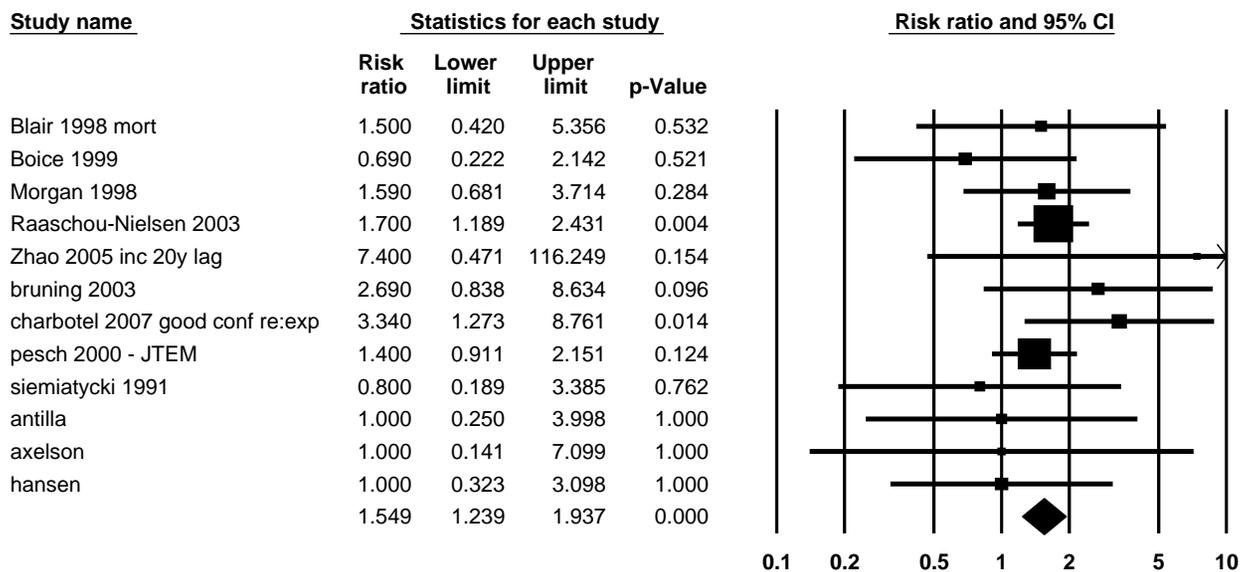
TCE and kidney cancer - highest exposure groups



random effects model; fixed effect same

- 1
- 2 Figure C.7. Meta-analysis of kidney cancer and TCE exposure – highest exposure groups. The pooled estimate is in the bottom row.
- 3 Symbol sizes reflect relative weights of the studies. The horizontal midpoint of the bottom diamond represents the pooled RR
- 4 estimate and the horizontal extremes depict the 95% CI limits.

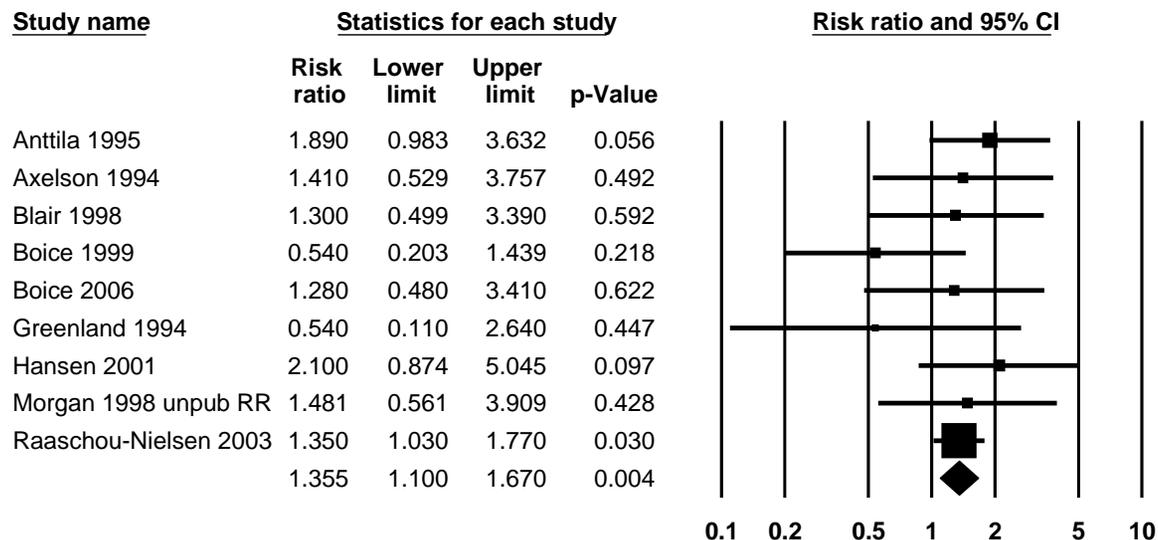
TCE and kidney cancer - highest exposure groups



random effects model; fixed effect same

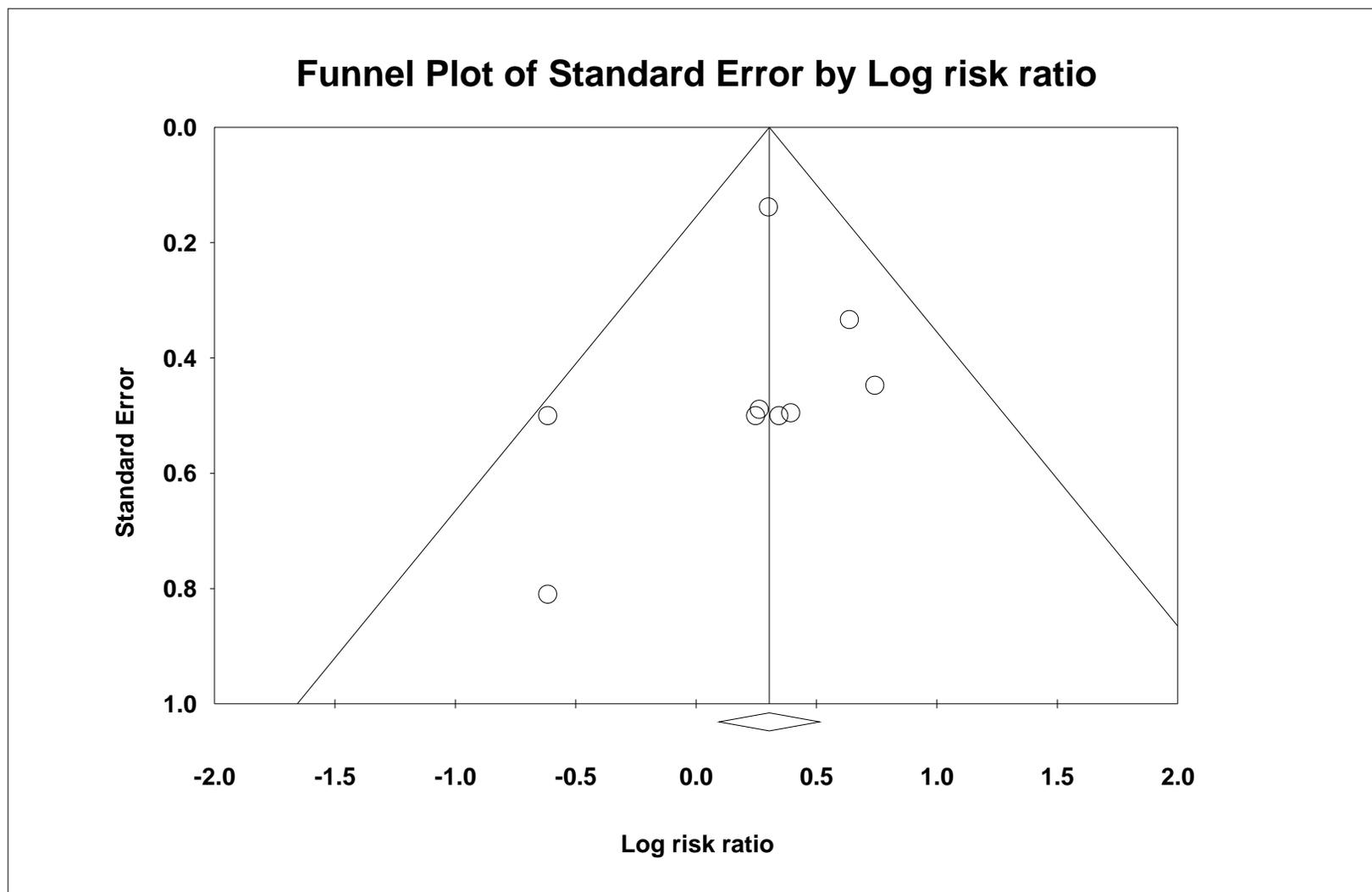
- 1
- 2 Figure C.8. Meta-analysis of kidney cancer and TCE exposure – highest exposure groups, with assumed null RR estimates for Anttila,
- 3 Axelson, and Hansen (see text).
- 4

TCE and liver cancer



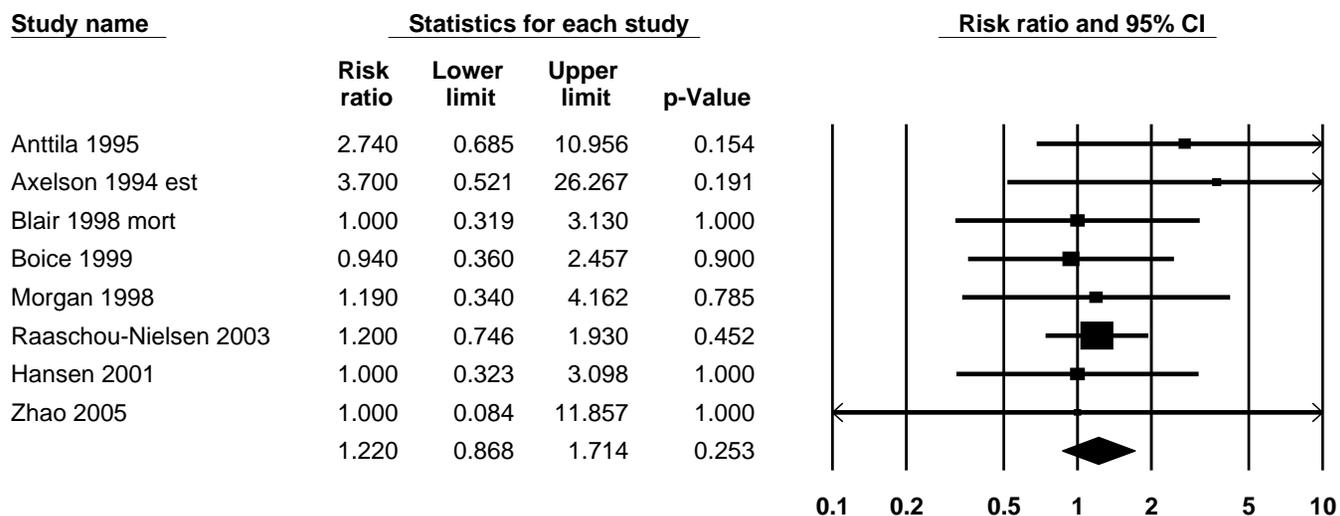
random effects model; same for fixed

1
 2 Figure C.9. Meta-analysis of liver cancer and TCE exposure. The pooled estimate is in the bottom row. Symbol sizes reflect relative
 3 weights of the studies. The horizontal midpoint of the bottom diamond represents the pooled RR estimate and the horizontal extremes
 4 depict the 95% CI limits.
 5



1
2 Figure C.10. Funnel plot of SE by log RR estimate for TCE and liver cancer studies

TCE and liver cancer - highest exposure groups



random effects model; same results with fixed effect model

1
 2 Figure C.11. Meta-analysis of liver cancer and TCE exposure – highest exposure groups, with assumed null RR estimates for Hansen
 3 and Zhao (see text).

4
 5

1 **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.
Blair et al. , 1998	2.0	0.9	4.6	mortality RR	0.693	0.416	none	ICD-8 200 + 202; adjusted for age, calendar time, and sex. referent group is workers with no chemical exposures.
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	nested case-control study. ICD-8 200-202.
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	1.24	1.01	1.52	SIR	0.215	0.104	none	ICD-7 200 + 202

al., 2003								
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)	Zhao 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 (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 (versus 33 in Zhao).

1 **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	Ann Arbor staging 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, 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

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

3 ^a the RR estimates are all ORs for incident cases

4 ^b not calculated

1 **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	15	random	1.27	1.04	1.53	significant (p = 0.048)	statistical significance not dependent on individual studies with the exception of one study whose exclusion yielded RRp estimate with p = 0.051
		fixed	1.20	1.06	1.36		
cohort	9	random	1.40	1.14	1.71	not significant	not significant difference between CC and cohort studies
		fixed	1.34	1.15	1.57		significant difference between CC and cohort studies
case-control	6	random	1.05	0.77	1.43	not significant	
		fixed	0.97	0.79	1.20		
alternate RR selections ^a	15	random	1.22	0.99	1.49	significant	with estimated Zhao overall RR for incidence rather than mortality
	15	random	1.24	1.00	1.54	significant	with Boice (2006) study rather than Zhao
	15	random	1.26	1.04	1.52	not significant (p = 0.051)	with estimated female contribution to Axelson
	15	random	1.26	1.05	1.51	significant (p = 0.048)	with Boice (1999) any potential exposure SMR
	15	random	1.29	1.06	1.57	not significant (p = 0.052)	with Morgan et al. (1998) unpublished RR

INTER-AGENCY REVIEW DRAFT – DO NOT QUOTE OR CITE

highest exposure groups	11	random	1.50	1.20	1.88	none observable (fixed = random)	with Blair incidence estimate and Zhao mortality estimate; statistical significance not dependent on single study. with mortality estimates from both: 1.57 (1.25, 1.96). with incidence estimates: 1.46 (1.12, 1.90). see Table C.5 for other alternate RR selections.
		fixed	1.50	1.20	1.88		

1 ^a changing the primary analysis by one alternate RR each time; more details on alternate RR estimates in text

1 **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.
Axelsson et al., 1994	6.25	0.16	34.83	≥ 2 y exp 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.
Blair et al., 1998	0.97	0.42	2.2	> 25 unit-y	0.589	0.495	1.8 (0.68, 4.8) mortality RR	incidence RR. ICD 200 + 202. Male and female results presented separately and combined (see text). referent group is workers with no chemical exposures.
Boice et al., 1999	1.62	0.82	3.22	≥ 5 y exp	0.482	0.349	none	mortality RR. ICD 200 + 202. For potential routine or intermittent exposure. adjusted for DoB, dates 1 st 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 months*mg/m ³	0.993	0.577	3.7 (1.0, 9.5) for ≥ 75 months exp duration 2.9 (0.79, 7.5) for ≥ 19 mg/m ³ mean exp	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	0.81	0.1	6.49	high	-0.211	1.06	1.31 (0.28, 6.08) for	mortality RR. ICD 200 only. adjusted for age and sex.

al., 1998				cumulative exp score			med/high peak vs. low/no	
Raaschou-Nielsen et al., 2003	1.6	1.1	2.2	≥ 5 y in subcohort with expected higher exp levels	0.470	0.183	none	SIR. ICD 200 + 202.
Zhao et al., 2005	1.30	0.52	3.23	high exp 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-exp gp) than do incidence results (1 in high-exp gp).
Miligi et al., 2006	1.2	0.7	2.0	med/high exp intensity	0.182	0.268	1.0 (0.5, 2.6) for med/high intensity and > 15 y exp	incidence OR. NHL + CLL (see section C.2.1.1).
Seidler et al., 2007	2.3	1.0	5.2	> 35 ppm-y	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, 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

1 ^a mean personal trichloroacetic acid in urine. 1 μmol/L = 0.1634 mg/L.
2

1 **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	1.50	1.20	1.88	none observable (fixed = random)	statistical significance not dependent on single study.
alternate RR selections ^a	random	1.57	1.25	1.96	none observable	with Blair mortality
	random	1.54	1.23	1.92	none observable	with Hansen duration
	random	1.53	1.20	1.93	none observable	with Miligi with >15 y
	random	1.52	1.22	1.90	none observable	with Hansen mean exposure
	random	1.51	1.21	1.89	none observable	with Morgan peak
	random	1.50	1.20	1.88	none observable	with estimated female contribution for Axelson
	random	1.46	1.12	1.90	some, but not statistically significant (p = 0.34)	with Zhao incidence. 1.47 (95% CI 1.18, 1.85) with fixed effect model.

2 ^a changing the primary analysis by one alternate RR estimate each time

1
2

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.
Blair et al., 1998	1.6	0.5	5.1	mortality RR	0.470	0.592	none	ICD-8 189. adjusted for age, calendar time, and sex. referent group is workers with no chemical exposures.
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.

INTER-AGENCY REVIEW DRAFT – DO NOT QUOTE OR CITE

<p>Zhao et al., 2005</p>	<p>1.7</p>	<p>0.38</p>	<p>7.9</p>	<p>mortality RR</p>	<p>0.542</p>	<p>0.775</p>	<p>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)</p>	<p>ICD-9 189. males only. adjusted for age, SES, time since first employment, exposure to other carcinogens. 20-year 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.</p>
---------------------------------	------------	-------------	------------	---------------------	--------------	--------------	---	--

1

1 **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

2 ^a the RR estimates are all ORs for incident cases

3 ^b not calculated

1 **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	1.26	1.11	1.42	none observable	statistical significance not dependent on single study; no apparent publication bias
		fixed	1.26	1.11	1.42		
cohort	9	random	1.16	0.96	1.41	none observable	not significant difference between CC and cohort studies
		fixed	1.16	0.96	1.41		not significant difference between CC and cohort studies
case-control	5	random	1.41	1.08	1.83	not significant	
		fixed	1.32	1.13	1.54		
alternate RR selections ^a	14	random	1.25–1.26	1.11–1.12	1.41–1.42	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	1.12	1.42	none obs	with Morgan published SMR
	14	random	1.25	1.12	1.40	none obs	with Raaschou-Nielsen all kidney cancer
	14	random	1.24	1.10	1.40	none obs	with Brüning longest job held in

	14	random	1.26	1.12	1.42	none obs	industry with TCE with Charbotel full study
	14	random	1.19	1.07	1.32	none obs	with Pesch JEM
highest exposure groups	9	random	1.61	1.27	2.03	none obs	
	12	random	1.55	1.24	1.94	none obs	using RR = 1 for Anttila, Axelson, and Hansen (see text). see Table C.10 for alternate RR selection results (RRp 1.43–1.57)

1 ^a changing the primary analysis by one alternate RR each time

1 **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 y exp and 100+ mg/L U-TCA			1.0 assumed	reported high exposure group results for some cancer sites but not kidney.
Blair et al. , 1998	1.5	0.4	5.1	> 25 unit-y	0.405	0.649	0.9 (0.3, 3.2) incidence RR	mortality RR. ICD-8 189. male and female results presented separately and combined (see text). referent group is workers with no chemical exposures.
Boice et al. , 1999	0.69	0.22	2.12	≥ 5 y exp	-0.371	0.578	none	mortality RR. ICD-9 189.0-189.2. for potential routine or intermittent exposure. adjusted for DoB, dates 1 st and last employed, race, and sex. referent group is workers not exposed to any solvent.
Hansen et al. , 2001				≥ 1080 months*mg/m3			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 exp 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 y in subcohort with expected	0.531	0.183	1.4 (0.99, 1.9) ICD-7 180 ≥5 y in total cohort	SIR. RCC.

INTER-AGENCY REVIEW DRAFT – DO NOT QUOTE OR CITE

Zhao et al., 2005	7.40	0.47	116	higher exp levels high exp score	2.00	1.41	mortality RR: 1.82 (0.09, 38.6) inc RR no lag: 7.71 (0.65, 91.4) mort RR no lag : 0.96 (0.09, 9.91) Boice 2006 mort RR: 2.12 (0.63, 7.11) for ≥ 5 y as test stand mechanic ; 3.13 (0.74,13.2) for ≥ 4 test-y engine flush	incidence RR. ICD-9 189. males only. adjusted for age, SES, time since first employment, exposure to other carcinogens. 20-year 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 y 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 2.16 (1.02, 4.60) for high cum + peaks 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.

							2.73 (1.06, 7.07) for high cum in full study	
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

1 ^a mean personal trichloroacetic acid in urine. 1 μmol/L = 0.1634 mg/L.

1 **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	1.61	1.27	2.03	none observable	
primary analysis	random	1.55	1.24	1.94	none observable (fixed = random)	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 observable	with Blair incidence
	random	1.57	1.26	1.96	none observable	with Morgan peak metric
	random	1.43	1.16	1.77	none observable	with Raaschou-Nielsen for all kidney cancer ≥ 5 y in total cohort
	random	1.53–1.55	1.22–1.24	1.91–1.94	none observable	with Zhao incidence unlagged and mortality with and without lag
	random	1.55–1.56	1.24–1.25	1.93–1.95	none observable	with Boice alternates for Zhao (see text)
	random	1.53–1.54	1.23	1.91–1.93	none observable	with Charbotel high cumulative dose + peaks in subgroup and with and without

	random	1.43	1.17	1.76	none observable	peaks in full study with Pesch JEM
--	--------	------	------	------	--------------------	---------------------------------------

1 ^a changing the primary analysis by one alternate RR each time

2

1
2

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.
Blair et al., 1998	1.3	0.5	3.4	mortality RR	0.262	0.489	1.7 (0.2, 16.2) for ICD-8 155	ICD-8 155 + 156. adjusted for age, calendar time, and sex. referent group is workers with no chemical exposures.
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	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

INTER-AGENCY REVIEW DRAFT – DO NOT QUOTE OR CITE

al., 2003									distribution.
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 (2006) used in lieu of Zhao et al. (2005) because Zhao et al. do not report liver cancer results. Boice et al. (2006) cohort overlaps Zhao cohort.	

1

1 **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	1.36	1.10	1.67	none observable	statistical significance not dependent on single study, except for Raaschou-Nielsen, without which p = 0.06; no apparent publication bias
		fixed	1.36	1.10	1.67		
all studies; liver cancer only, when available	9	random	1.32	1.02	1.70	none observable	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	1.35	1.09	1.67	none observable	with 1.0 assumed for Zhao in lieu of Boice (2006) (see text)
	9	random	1.30	1.07	1.58	none observable	with Boice (1999) any potential exposure rather than potential routine exposure
	9	random	1.35	1.10	1.67	none observable	with estimated female contribution to Axelson
	9	random	1.32	1.08	1.62	none observable	with Morgan published SMR
highest	6	random	1.25	0.87	1.79	none	

INTER-AGENCY REVIEW DRAFT – DO NOT QUOTE OR CITE

exposure groups						observable	
	8	random	1.22	0.87	1.71	none observable	using RR = 1 for Hansen and Zhao (see text).
	7-8	random	1.18-1.20	0.81-0.86	1.67-1.74	none observable	using alternate selections for Morgan and Raaschou-Nielsen and excluding Axelson ^a

1 ^a changing the primary analysis by one alternate RR each time

1 **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 exp gp (i.e., ≥ 2y and 100+ U-TCA), so combined with < 2y and 100+ subgp and females, estimating the expected #s (see text).
Blair et al., 1998	1.0	0.33	3.2	> 25 unit-y	0	0.582	none (see text.)	mortality RR. ICD-8 155 + 156. male and female results presented separately and combined (see text). referent group is workers with no chemical exposures.
Boice et al., 1999	0.94	0.36	2.46	≥ 5 y exp	-0.062	0.490	none	mortality RR. ICD-9 155 + 156. for potential routine or intermittent exposure. adjusted for DoB, dates 1 st and last employed, race, and sex. referent group is workers not exposed to any solvent.
Hansen et al., 2001				≥ 1080 months*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 exp 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 y	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.

Zhao et al., 2005	high exp score	1.0 assumed	no liver results reported.
------------------------------	----------------	-------------	----------------------------

1 ^a mean personal trichloroacetic acid in urine. 1 μmol/L = 0.1634 mg/L.

2

3

1 **References**

- 2
- 3 Anttila A, Pukkala E, Sallmén M, Hernberg S, Hemminki K. (1995) Cancer incidence among
4 Finnish workers exposed to halogenated hydrocarbons. *J Occup Environ Med*
5 37:797–806.
- 6 Axelson O, Selden A, Andersson K, Hogstedt C. (1994). Updated and expanded 1 Swedish
7 cohort study on trichloroethylene and cancer risk. *J Occup Environ* 36:556–562.
- 8 Banks PM (1992) Changes in diagnosis of non-Hodgkin’s lymphomas over time. *Cancer*
9 *Research* 52:5453s–5455s.
- 10 Begg CB, Mazumdar M. 1994. Operating characteristics of a rank correlation test for
11 publication bias. *Biometrics* 50:1088-1101.
- 12 Blair A, Hartge P, Stewart PA, McAdams M, Lubin J. (1998) Mortality and cancer incidence of
13 aircraft maintenance workers exposed to trichloroethylene and other organic solvents and
14 chemicals: extended follow up. *Occup Environ Med* 55:161–171.
- 15 Boice JD, Marano DE, Fryzek JP, Sadler CJ, McLaughlin JK. (1999) Mortality among aircraft
16 manufacturing workers. *Occup Environ Med* 56:581–597.
- 17 Boice JD, Marano DE, Cohen SS, Mumma MR, Blott WJ, Brill AB, Fryzek JP, Henderson BE,
18 McLaughlin JK. (2006) Mortality among Rocketdyne worker who tested rocket engines,
19 1948-1999. *J Occup Environ Med* 48:1070–1092.
- 20 Breslow NE, Day NE (1980) *Statistical Methods in Cancer Research, Volume I – The Analysis*
21 *of Case-Control Studies*. IARC Scientific Publications No. 32, International Agency for
22 *Research on Cancer, Lyon.*
- 23 Breslow NE, Day NE (1987) *Statistical Methods in Cancer Research, Volume II – The Design*
24 *and Analysis of Cohort Studies*. IARC Scientific Publications No. 82, International
25 *Agency for Research on Cancer, Lyon.*
- 26 Brüning T, Pesch B, Wiesenhütter B, Rabstein S, Lammert M, Baumüller A, Bolt HM. 2003.
27 Renal cell cancer risk and occupational exposure to trichloroethylene: results of a
28 consecutive case-control study in Arnsberg, Germany. *Am J Ind Med.* 43:274–85.
- 29 Charbotel B, Fevotte J, Hours M, Martin JL, Bergeret A. (2006) Case-control study on renal cell
30 cancer and occupational exposure to trichloroethylene. Part II: Epidemiological aspects.
31 *Ann Occup Hyg* 50(8):777–87.

- 1 Cherrie, J.W., H. Kromhout, and S. Semple. 2001. The importance of reliable exposure estimates
2 in deciding whether trichloroethylene can cause kidney cancer [letter]. *J Cancer Res Clin*
3 *Oncol* 127:400–402.
- 4 Copeland, K.T., H. Checkoway, A.J. McMichael, and R.H. Holbrook. 1977. Bias due to
5 misclassification in the estimation of relative risk. *Am. J. Epidemiol.* 105:488-495.
- 6 DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. *Controlled Clinical Trials*
7 7:177–188.
- 8 Dosemeci, M; Cocco, P; Chow, WH. (1999) Gender differences in risk of renal cell carcinoma
9 and occupational exposures to chlorinated aliphatic hydrocarbons. *Am J Ind Med*
10 36(1):54–9.
- 11 Duval S, Tweedie R (2000) A nonparametric “trim and fill” method of accounting for
12 publication bias in meta-analysis. *J Am Stat Assoc* 95:89–98.
- 13 Egger M, Smith GD, Schneider M, Minder C (1997) Bias in meta-analysis detected by a simple,
14 graphical test. *Brit Med J* 315:629–634.
- 15 Environmental Health Strategies. 1997. Final Report. Written correspondence from Paul A.
16 Cammer, Ph.D., Trichloroethylene Issues Group, to Cheryl Siegel Scott, US
17 Environmental Protection Agency. December 22, 1997.
- 18 Fevotte J, Charbotel B, Muller-Beaute P, Martin J-L, Hours, Bergeret A. 2006. Case-control
19 study on renal cell cancer and occupational exposure to trichloroethylene. Part I:
20 Exposure assessment. *Ann Occup Hyg* 50:765–775.
- 21 Greenland S, Salvan A, Wegman DH, Hallock MF, Smith TH. 1994. A case-control study of
22 cancer mortality at the transformer-assembly facility. *Int Arch Occup Environ Health*
23 66:49–54.
- 24 Hansen J, Raaschou-Nielsen O, Christensen JM, Johansen I, McLaughlin JK, Lipworth L, Blot
25 WJ, Olsen JH. 2001. Cancer incidence among Danish workers exposed to
26 trichloroethylene. *J Occup Environ Med* 43:133–139.
- 27 Hardell L, Eriksson M, Degerman A. 1994. Exposure to phenoxyacetic acids, chlorophenols, or
28 organic solvents in relation to histopathology, stage, and anatomical localization of non-
29 Hodgkin’s lymphoma. *Cancer Res* 54:2386–2389.
- 30 Harris NL, Jaffe ES, Diebold J, Flandrin G, Muller-Hermelink HK, Vardiman J (2000)
31 Lymphoma classification – from controversy to consensus: the R.E.A.L. and WHO
32 classification of lymphoid neoplasms. *Ann Oncol* 11 Suppl 1:3–10.

- 1 Mandel JH, Kelsh MA, Mink PJ, Alexander DD, Kalmes RM, Weingart M, Yost L, Goodman M
2 (2006) Occupational trichloroethylene exposure and non-Hodgkin's lymphoma: a meta-
3 analysis and review. *Occup Environ Med* 63:597–607.
- 4 Mannetje, A., K. Steenland, M. Attfield, P. Boffetta, H. Checkoway, N. DeKlerk, and R.S.
5 Koskela. 2002. Exposure-response analysis and risk assessment for silica and silicosis
6 mortality in a pooled analysis of six cohorts. *Occup. Environ. Med.* 59:723-728.
- 7 McGuire, V., L.M. Nelson, T.D. Koepsell, H. Checkoway, and W.T. Longstreth, Jr. 1998.
8 Assessment of occupational exposures in community-based case-control studies. *Annu.*
9 *Rev. Public Health* 19:35-53.
- 10 Miligi L, Costantini AS, Benvenuti A, Kreibel D, Bolejack V, et al. 2006. Occupational
11 exposure to solvents and the risk of lymphomas. *Epidemiol* 17:552–561.
- 12 Morgan RW, Kelsh MA, Zhao K, Heringer S. 1998. Mortality of aerospace workers exposure
13 to trichloroethylene. *Epidemiol* 9:424–431
- 14 Nelson, L.M., W.T. Longstreth, Jr., T.D. Koepsell, H. Checkoway, and G. van Belle. 1994.
15 Completeness and accuracy of interview data from proxy respondents: Demographic,
16 medical, and life-style factors. *Epidemiology* 5:204-217.
- 17 Nordstrom M, Hardell L, Hagberg H, Rask-Andersen A. 1998. Occupational exposures, animal
18 exposure and smoking as risk factors for hairy cell leukaemia evaluated in a case-control
19 study. *Br J Cancer* 77:2048–2052.
- 20 NRC (National Research Council). 2006. Assessing the Human Health Risks of
21 Trichloroethylene. Key Scientific Issues. National Academy Press, Washington, DC.
- 22 Persson B, Fredriksson M, Olsen K, Boeryd B, Axelson O. 1993. Some occupational exposure
23 as risk factors for malignant lymphomas. *Cancer* 72:1773–1778.
- 24 Pesch B, Haerting J, Ranft U, Klimpet A, Oelschägel, Schill W, and the MURC Study Group.
25 2000b. Occupational risk factors for renal cell carcinoma: agent-specific results from a
26 case-control study in Germany. *Int J Epidemiol* 29:1014–1024.
- 27 Raaschou-Nielsen O, Hansen J, McLaughlin JK, Kolstad H, Christensen JM, Tarone RE, Olsen
28 JH. 2003. Cancer risk among workers at Danish companies using trichloroethylene: a
29 cohort study. *Am J Epidemiol* 158:1182–1192.
- 30 Rothman KJ, Greenland S (1998) *Modern Epidemiology*, Second Edition. Lippincott Williams
31 & Wilkins, Philadelphia.

- 1 Seidler A, Mohner M, Berger J, Mester B, Deeg E, Elsner G, Neiters A, Becker N. 2007.
2 Solvent exposure and malignant lymphoma: a population-based case-control study in
3 Germany. *J Occup Med Toxicol* 2:2.
- 4 Siemiatycki J. 1991. Risk Factors for Cancer in the Workplace. J Siemiatycki, Ed. Boca
5 Raton: CRC Press.
- 6 Woolf B (1955) On estimating the relationship between blood group and disease. *Ann Human*
7 *Genet* 19:251–253.
- 8 Zhao Y, Krishnadasan A, Kennedy N, Morgenstern H, Ritz B. 2005. Estimated effects of
9 solvents and mineral oils on cancer incidence and Mortality in a cohort of aerospace
10 workers. *Am J Ind Med* 48:249–258.
- 11