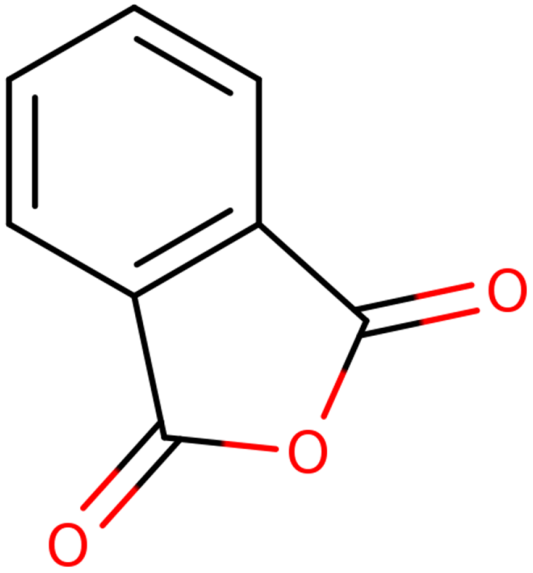
United States
Environmental Protection Agency

March 2026
Office of Chemical Safety and
Pollution Prevention

Data Quality Evaluation Information for Human Health Hazard Epidemiology for Phthalic Anhydride

Systematic Review Support Document for the Draft Risk Evaluation

CASRN: 85-44-9



March 2026

This supplemental file contains the data quality evaluation results for epidemiology data sources that met the PECO (Population, Exposure, Comparator or Scenario, and Outcomes) screening criteria and further filtering criteria for the *Draft Human Health Hazard Assessment for Phthalic Anhydride*. EPA conducted data quality evaluation based on author-reported descriptions and results; additional analyses (*e.g.*, statistical analyses performed during data integration into the risk evaluation) potentially conducted by EPA are not contained in this supplemental file. EPA used the TSCA systematic review process described in the *Draft Systematic Review Protocol Supporting TSCA Risk Evaluations for Chemical Substances* (also referred to as '2021 Draft Systematic Review Protocol'). Any updated steps in the systematic review process since the publication of the 2021 Draft Systematic Review Protocol are described in the *Draft Systematic Review Protocol for Phthalic Anhydride*.

Table of Contents

Mixture: Acid anhydrides (phthalic anhydride, maleic anhydride, trimellitic anhyride)

Study Citation:	Nielsen, J., Bensryd, I., Almquist, H., Dahlqvist, M., Welinder, H., Alexandersson, R., Skerfving, S. (1991). Serum IgE and lung function in workers exposed to phthalic anhydride. International Archives of Occupational and Environmental Health 63(3):199-204.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Lung function (spirometry including forced expiratory volume, vital capacity. maximal expiratory flow at 50% and 75% of forced expiratory volume, closing volume, volume of trapped gas, carbon monoxide transfer factor, and pulmonary clearance of diethylenetriamine pentaacetate (DTPA)); respiratory symptoms (chronic bronchitis, dry cough).
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5178100

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	Geographic setting was reported but temporal was not; inclusion/exclusion criteria were not articulated, and participation rate at different steps was not reported.
	Metric 2:	Attrition	Low	Numbers of individuals were not reported at important stages of study were not reported.
	Metric 3:	Comparison Group	Medium	There is indirect evidence that the groups were similar (authors reported that "The exposed workers and the control subjects were matched with regard to age and smoking habits" and provided limited comparative information).
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	Sampling was conducted for airborne contaminants and reported as total acid anhydrides (no data for PA alone) across 2 plants. Authors reported that controls were "not exposed to harmful chemical dust or smoke to any significant extent" but did not conduct any sampling for the control group.
	Metric 5:	Exposure Levels	Low	Reports 2 levels of exposure (exposed/unexposed)
	Metric 6:	Temporality	Low	PA's historical exposure levels were not estimated. PA's exposure sampling was for 7.9 hours of total sampling time and collected 24 samples. It is unclear whether exposures fall within relevant exposure windows for the outcome of interest.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	The outcome was assessed using self-reported symptoms, spirometry, volume of trapped gas, closing volume, carbon monoxide transfer factor, and pulmonary clearance of inhaled radioactive diethylenetriamenepenta-acetate. The authors did not report details of how the outcome assessments were performed on subjects (timing, sequence, etc.)
	Metric 8:	Reporting Bias	Medium	All outcomes were reported, but continuous data were given as median and range without SD/SE.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	Some covariates were considered by matching, e.g., "The exposed workers and the control subjects were matched with regard to age and smoking habits." All subjects were men and "Before the investigation no subject had any known lung disease." Authors evaluated atopy by history of symptoms and skin prick tests, and "There was no significant difference between the exposed workers and the control subjects with regard to smoking and history of atopy."

Continued on next page . . .

. . . continued from previous page

Study Citation:	Nielsen, J., Bensryd, I., Almquist, H., Dahlqvist, M., Welinder, H., Alexandersson, R., Skerfving, S. (1991). Serum IgE and lung function in workers exposed to phthalic anhydride. International Archives of Occupational and Environmental Health 63(3):199-204.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Lung function (spirometry including forced expiratory volume, vital capacity. maximal expiratory flow at 50% and 75% of forced expiratory volume, closing volume, volume of trapped gas, carbon monoxide transfer factor, and pulmonary clearance of diethylenetriamine pentaacetate (DTPA)); respiratory symptoms (chronic bronchitis, dry cough).
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5178100

Domain	Metric		Rating	Comments
	Metric 10:	Covariate Characterization	Medium	Smoking and history of atopy were assessed by interview and no information on the validity was provided. Atopy was also assessed by skin prick test.
	Metric 11:	Co-exposure Counfounding	Low	No adjustment for co-exposures was made and there is direct evidence for unbalanced provision of co-exposures. Exposed subjects had co-exposures to solvents and other acid anhydrides. Controls were ”employed in the municipal engineering department (repairmen and gardeners)” and ”were not exposed to harmful chemical dust or smoke to any significant extent.” Potential exposures to other chemicals and/or pesticides were not considered.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	Low	Study design was appropriate to evaluate the association between exposure and health outcomes in a cross-sectional study. For example, the manuscript showed that ”For comparison of distributions between different groups (exposed workers and controls; groups with different symptoms), the Mann-Whitney U-test was used. For comparison of symptoms in the exposed group and control group, Fisher’s test was applied. All stated P-values involved two-tailed analysis; differences were considered to be statistically significant at P $<$ 0 05.”However, the statistical analyses did not put co-exposure pollutants into consideration to adjust the association between exposure and health outcomes.
	Metric 13:	Statistical Power	Medium	Numbers of subjects were low but adequate to detect differences in some metrics.
	Metric 14:	Reproducibility of Analyses	Medium	Description sufficient to understand what was done and be conceptually reproducible
	Metric 15:	Statistical Analysis	N/A	statistical models were not applied

Additional Comments: Lung function and atopy/allergic responses were measured in 23 male workers exposed to phthalic anhydride and other acid anhydrides and compared with measurements in 18 male repairmen and gardeners employed in a municipal engineering department. Respiratory outcomes included self-reported symptoms, spirometry, volume of trapped gas, closing volume, carbon monoxide transfer factor, and DTPA clearance. Total acid anhydride levels in individual air samples averaged 6.6 mg/m3 (TWA) during phthalic anhydride loading. Exposed subjects had significantly higher MEF50 and MEF25, and significantly higher prevalence of non-specific bronchial hyperreactivity. No other metrics were significantly different. Exposed subjects had co-exposures to other acid anhydrides and solvents.

Overall Quality Determination — Low

Study Citation:	Nielsen, J., Welinder, H., Schutz, A., Skerfving, S. (1988). Specific serum antibodies against phthalic anhydride in occupationally exposed subjects. Journal of Allergy and Clinical Immunology 82(1):126-133.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Respiratory symptoms (chronic bronchitis)
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176341

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	Geographic setting was reported but temporal was not; inclusion/exclusion criteria were not articulated, and participation rate at different steps was not reported.
	Metric 2:	Attrition	Low	Numbers of individuals were not reported at important stages of study were not reported.
	Metric 3:	Comparison Group	Medium	There is indirect evidence that the groups were similar; authors reported that control subjects with similar age distribution and smoking habits were randomly selected; Exposed groups consisted of men and women (2), while the reference group only consisted of males. limited comparative information was reported.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	Sampling was conducted for airborne contaminants across 2 plants. Authors reported that controls (workers at a food-processing factory) "had not been exposed to any harmful chemical, smoke, or dust in their work environment " but did not conduct any sampling for the control group. In addition, the detailed employment records for exposed and control groups were not provided.
	Metric 5:	Exposure Levels	Low	Reports 3 qualitative levels of exposure (heavily exposed/lightly exposed/unexposed); Exposures of the 2 plants were similar so workers were "pooled". In addition, the lightly exposed areas in both plants only had 5 samples and the detected levels were <0.1 mg/m3. The exposure level of the control group is assumed to be zero, so the range of exposure in the population is limited.
	Metric 6:	Temporality	Medium	Temporality is established, but it is unclear whether exposures fall within relevant exposure windows for the outcome of interest. The average exposure duration for heavily exposed workers was 13 years (range 0-43 years) and 12 years for slightly exposed workers (range: 0.3-40 years).
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	The outcome was assessed using self-reported symptoms. Physical examinations were performed; spirometry results were reported for those having asthma and chronic bronchitis; however, the authors did not report details of how the outcome assessments were performed on subjects (timing, sequence, etc.)
	Metric 8:	Reporting Bias	High	All outcomes were reported as incidence and % of exposure group.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Nielsen, J., Welinder, H., Schutz, A., Skerfving, S. (1988). Specific serum antibodies against phthalic anhydride in occupationally exposed subjects. Journal of Allergy and Clinical Immunology 82(1):126-133.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Respiratory symptoms (chronic bronchitis)
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176341

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	High	Some covariates were considered by matching; authors reported that control subjects with similar age distribution and smoking habits. It was noted that there was no significant difference between the exposure groups with regard to smoking and history of atopy.
	Metric 10:	Covariate Characterization	Low	Smoking and history of atopy were assessed by interview and no information on the validity was provided. Atopy for common allergens was also assessed by skin prick test.
	Metric 11:	Co-exposure Counfounding	Low	No adjustment for co-exposures was made and there is direct evidence for unbalanced provision of co-exposures. Exposed subjects had co-exposures to solvents and other anhydrides. Controls were employed in a food-processing factory and "had not been exposed to any harmful chemical, smoke, or dust in their work environment." Potential exposures to other chemicals and/or pesticides were not considered.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	Low	"For comparison of proportions, the chi-square test was used, and when expected numbers were < 5, Fisher's exact test was used." Association between pairs of variables were tested using Spearman's rank-correlation test. This cross-sectional study only can show that there could be an association between PAD and lung functions, but cannot answer if only PAD caused lunch functions. There are two reasons for this unclear causality: (1) co-exposure from other chemicals in the plants were not assessed, (2) historical PAD exposure in the plants were not provided.
	Metric 13:	Statistical Power	Medium	Numbers of subjects were low but adequate to detect differences in some metrics; there were no significant differences in respiratory outcomes (chronic bronchitis).
	Metric 14:	Reproducibility of Analyses	Medium	Description sufficient to understand what was done and be conceptually reproducible.
	Metric 15:	Statistical Analysis	N/A	Statistical models were not applied.

Additional Comments: Lung function and atopy/allergic responses were measured in 60 workers (58 men, 2 women) exposed to phthalic anhydride and other acid anhydrides and compared with measurements in 22 male workers employed in a food-processing factory. Exposure groups included heavy exposure (n=35), low exposure (n=25), and nonexposed (n=22). Respiratory outcomes, including chronic bronchitis, included self-reported symptoms. Phthalic anhydride levels in individual air samples averaged 6.6 mg/m3 (TWA) during phthalic anhydride loading. 6 (17%) of heavily exposed subjects had chronic bronchitis and one (4%) of low exposed workers had chronic bronchitis; the difference was not statistically significant.

Overall Quality Determination Medium

Study Citation:	Riboli, E., Bai, E., Berrino, F., Merisi, A. (1983). Mortality from lung cancer in an acetylene and phthalic anhydride plant: a case-referent study. Scandinavian Journal of Work, Environment and Health 9(6):455-462.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Lung cancer mortality
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63774

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	In this retrospective case-control study in Italy, subjects were deceased residents of Pioltello, notably near a Societa ltaliana Serie Acetica Sintetica (SISAS) chemical factory where workers were exposed to phthalic anhydride. Subjects were recruited without knowledge of exposure status. Male subjects were identified, through the municipal death register of Pioltello, who had a cause of death attributed to lung cancer between January 1, 1976 – December 31, 1979. Referents were made up of the first 2 persons listed after each death record in the register, within 5 years of age as the case, who did not die of respiratory cancer. Age and date of death were comparable between cases and referents; sex of the referents were not specified. 53 cases and 106 referents were selected; setting, inclusion and exclusion criteria, and methods of case ascertainment were reported. The study reports that the death register is likely complete since reporting is compulsory by law.
	Metric 2:	Attrition	Medium	All subjects, once identified through the death register, were traced via contact with next of kin (typically widows, sometimes with other relatives). 6 of 53 cases and 7 of 106 referents were not able to be traced; this is reportedly due to subject's families moving or having no relatives living in the town, which are reasonable justifications not expected to be related to the subjects' exposure or outcome status. Further exclusions were made if they could not confirm death due to primary lung cancer (as opposed to other cancers that metastasized into the lungs) via identification in the Regional Discharge Diagnosis Information System, or if the subject's family did not confirm primary lung cancer. 4 subjects were determined to have cancer with lung metastasis and were thus removed from the analysis. The final study size consisted of 43 cases and 99 referents. There is no evidence that attrition was inappropriately handled.
	Metric 3:	Comparison Group	Medium	There is indirect evidence that cases and controls are similar. Cases and controls were recruited from the same eligible population within the same time frame using the register of deaths of the Pioltello municipality. Referents were made up of the first 2 persons listed after each death record in the register, within 5 years of age as the case, who did not die of respiratory cancer. Age and date of death were noted to be comparable between cases and referents; sex of the referents is not specified (cases only included males). None of the referents had a diagnosis of lung cancer based on the Regional Discharge Diagnosis Information System or family interviews. There are some potential concerns for healthy worker bias, as not all participants were reported to be employed. However, this is somewhat mitigated by some analyses that compare workers occupationally exposed at the SISAS plant to those exposed to potential lung carcinogens at other workplaces.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Riboli, E., Bai, E., Berrino, F., Merisi, A. (1983). Mortality from lung cancer in an acetylene and phthalic anhydride plant: a case-referent study. Scandinavian Journal of Work, Environment and Health 9(6):455-462.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Lung cancer mortality
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63774

Domain	Metric		Rating	Comments
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Low	An attempt was made to obtain complete occupational histories through interviews with next of kin; information on job duties and period of employment were collected. The study was designed to determine if exposures of phthalic anhydride and other suspected carcinogens (soot and phthalates) from employment at one chemical manufacturing plant (SISAS, 2-20 years) resulted in lung cancers. The study did not measure specific exposures to phthalic anhydride. In addition, it was noted the possibility of exposure to lung carcinogens in a few subjects may have occurred while working for other companies. Certain job functions and duties were not associated specifically with only phthalic anhydride exposure. The use of interviews with next of kin to determine job function is not the most reliable, and has potential for exposure misclassification, but this is not expected to the large enough to significantly impact the results. Exposure assessment was reported to have been completed without knowledge of case status.
	Metric 5:	Exposure Levels	Low	Exposure was categorized into three groups: exposed at the SISAS plant, exposed to lung carcinogens at other places of employment, or not exposed to occupational lung carcinogens. There is no quantitative information on exposure levels or range of exposures for subjects.
	Metric 6:	Temporality	High	Temporality of exposure and outcome is established; it was noted that the period after production (exposure) was compatible with the latency period between exposure and the outcome (death from lung cancer)
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	The lung cancer mortality outcome was assessed using information from the Register of Deaths of the Pioltello municipality death certificates and validated using the Regional Discharge Diagnosis Information system and interviews with next of kin by medical doctors, in order to distinguish between primary lung cancer and other cancer that metastasized to the lungs. There is overall high certainty in the outcome definition.
	Metric 8:	Reporting Bias	Medium	The outcome of death by lung cancer in cases and referents were reported. Authors reported the number of cases in each exposure group in the table. Odds ratios for the risk of dying from lung cancer were reported with respective confidence intervals.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Riboli, E., Bai, E., Berrino, F., Merisi, A. (1983). Mortality from lung cancer in an acetylene and phthalic anhydride plant: a case-referent study. Scandinavian Journal of Work, Environment and Health 9(6):455-462.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Lung cancer mortality
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63774

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	Medium	Cases and controls were matched for age (within 5 years). Considerations of age and date of death were accounted for in selection of subjects. Occupational history and smoking habits were also considered as covariates; light smokers (<10 cigarettes/day) were combined with those who never smoked. Sex and socioeconomic status were not adjusted for; it is reported that cases included males only; however it is not specified if referents only included males. Results were stratified by cigarette use and age. While the study does not explain in-depth why covariates were or were not chosen, there is no evidence of substantial residual confounding.
	Metric 10:	Covariate Characterization	Medium	Potential covariates were assessed through the death records (age of death) and interviews with next of kin that established work histories and smoking habits using recall questions from next of kin. There is potential error in asking next of kin to report on smoking habits, but other covariates were assessed using adequate methods.
	Metric 11:	Co-exposure Counfounding	Low	The study reported that SISAS employees were potentially exposed to many chemicals including acetylene, phthalic anhydride and their derivatives, soot and phthalates. The study was designed to assess all occupational exposures at SISAS and risk of death by lung cancer and not for any individual chemical. Co-exposures were not adjusted for. Exposures to lung carcinogens at other employers was possible, and no information on relative exposure levels was provided.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The study design was adequate to assess the association between exposure to all occupational exposures in the SISAS plant and risk of death by lung cancer, but not for any specific chemical exposure. The study authors noted they used statistical methods described by Mantel And Haenszel (1959, 1963) to derive chi square associations, point estimates for odds ratios, and trend tests. Standardized risk ratios estimations were conducted using programs developed by Rothman and Boice. The study's approach is appropriate to assess the relationship between potential phthalic anhydride exposure and lung cancer mortality.
	Metric 13:	Statistical Power	Medium	Statistical power was not calculated. The OR estimates for SISAS exposure and dying by lung cancer was based on a small number of cases in each level of smokers resulting in large variability (n=43). The number of cases and controls is adequate to detect an effect in the total exposed population; but not for subgroups of the population age and smoking category, as evidenced by wide confidence intervals reported for analyses stratified by cigarette exposure.
	Metric 14:	Reproducibility of Analyses	Medium	The methods are described in sufficient detail to reproduce the results with access to the analytic data.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Riboli, E., Bai, E., Berrino, F., Merisi, A. (1983). Mortality from lung cancer in an acetylene and phthalic anhydride plant: a case-referent study. Scandinavian Journal of Work, Environment and Health 9(6):455-462.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Lung cancer mortality
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63774

Domain	Metric	Rating	Comments
	Metric 15: Statistical Analysis	High	The study calculates the chi-square for association as well as odds ratios using the Mantel and Haenszel method. The computation is sufficiently reported and appropriate, with test-based confidence limits calculated as well.

Additional Comments: A retrospective case-referent study was conducted in a population in Milano Province, Italy where there is an acetylene and phthalic anhydride manufacturing plant (SISAS). Cases of deaths by lung cancer (n=43) were determined from death records (and validated through clinical and interviews of next of kin) and compared to referents (n=99). Standardized risk ratios (adjusted for age and smoking habits) for dying with lung cancer associated with occupational exposure to a number of chemicals (phthalic anhydride, acetylene, and their derivatives, soot, phthalates) were calculated. The study was designed to determine if occupational exposures from employment at the chemical manufacturing plant (SISAS) or exposures to suspected carcinogens from other occupations resulted in lung cancers. Exposures were categorized as exposure from subjects who worked at the SISAS plant for 2-20 years (S+), subjects who have ever worked at a job likely to have exposure to lung carcinogens (E+), or no occupational exposure (E-). There is no quantitative information on exposure levels or range of exposures and risk ratios were not calculated for exposure specifically to phthalic anhydride. The risk of dying due to lung cancer of subjects employed at SISAS was 5.6 (95% CI 1.9-16.2) compared to referents; the risk for those occupationally exposed to lung carcinogens, but not employed at SISAS was 1.7% (95% CI 0.9-3.5). It was noted by the authors that exposure to soot and phthalates at SISAS contributed to the increased risk of lung cancer.

Overall Quality Determination Medium

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Forced vital capacity (FVC), forced expiratory volume in 1 second (FEV1), abnormal chest x-ray findings
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	In this cross-sectional occupational surveillance study, workers from 9 Koppers Coal Tar plants were examined from January-October 1979. Only one of these plants (the Chicago plant) produced phthalic anhydride. Participation at the relevant phthalic anhydride plant (i.e., the Chicago plant) was approximately 46% (105/230), while participation from the total workforce across all plants was 51% (453/888). The study authors note that participation at the Chicago plant was non-representative for hourly workers relative to salaried workers. The study does not discuss recruitment or selection processes in detail but does indicate that workers could participate in the study during normal work hours without impact on their pay. Overall, details are sparse and it is unclear if those who did not participate may have been more or less exposed compared to those who were included.
	Metric 2:	Attrition	High	One participant out of 105 in the Chicago did not have a blood test taken and was excluded from analyses that included serum measures. No reason is provided for why the blood test was not taken, but there was no other indication of attrition or exclusion in the Chicago plant group or in the total included group.
	Metric 3:	Comparison Group	Low	Demographic details on sex and race are provided for all plants. The distribution of race and sex is roughly similar in the Chicago plant to the overall distribution of race and sex across the included workforce. However, these demographic differences are not controlled for in statistical analyses. There is also no discussion of age and whether groups had differences based on age.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Low	In comparisons between different plants, all Chicago plant workers were grouped together. Descriptions of the Chicago plant were very limited beyond stating that phthalic anhydride was used. A sub-group analysis of the Chicago plant separates out workers by ”type” (Phthalic, Maleic, or Coal Tar). It is unclear if phthalic anhydride workers were all actually formally exposed to phthalic anhydride, and it is not clear whether the other types of workers were not exposed. There is a large potential for exposure misclassification if not all tasks or employees were in contact with phthalic anhydride, and since exposure was estimated solely using professional judgment, exposure misclassification cannot be ruled out.
	Metric 5:	Exposure Levels	Low	The study only reports two levels of exposure, exposed and unexposed. No quantitative exposure information is available.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Forced vital capacity (FVC), forced expiratory volume in 1 second (FEV1), abnormal chest x-ray findings
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
	Metric 6:	Temporality	Low	Work history or duration of employment was not discussed in the analysis. It is clear that the outcomes were measured after some degree of exposure, since outcomes were assessed in an occupational context. However, it is unclear how long these workers may have been exposed, and it is unclear whether the employees were free of any of the reported health outcomes prior to enrollment into the study. The temporality of exposure and outcome is uncertain.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Specific methods were not described, but description of the study suggests standard clinical methods were used (in line with NIOSH recommendations). Examination forms were provided in the appendices.
	Metric 8:	Reporting Bias	Medium	All outlined outcomes are reported in the results, however, data is provided as the proportion of individuals with ”abnormal” measurements compared against clinical standards.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Low	There was no evidence of adjustment for potential confounders, however, some were discussed in-text for medical interpretations. The distribution of sex and race is provided across exposure group, but there is no discussion of age. Covariates were not adjusted for or stratified in relation to outcomes.
	Metric 10:	Covariate Characterization	Medium	Questionnaires were provided in the appendix. It was not reported whether this was a validated questionnaire, however, there was no evidence to suggest it was an invalid instrument.
	Metric 11:	Co-exposure Counfounding	Low	Exposure to other coal tar components was described, including potential health effects. Workers in plants other than the Chicago plant were exposed to other occupational agents, and it is unclear whether those exposures contributed to the incidence of health outcomes. Within the Chicago plant, the study splits the sample into workers exposed to phthalic anhydride, maleic anhydride, or coal tar. It is not confirmed that phthalic anhydride workers were only exposed to phthalic anhydride.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The design limited the ability to determine which exposures were relevant to each health effect, however the design is sufficient to answer the study’s question as to whether there are ”abnormal” medical findings among workers at the studied coal tar plants.
	Metric 13:	Statistical Power	Low	Statistical power was not reported. A total of 105 employees completed exams from the Chicago plant. However, the analysis within the Chicago plant only identifies 14 workers who were exposed to phthalic anhydride, which may be too small of a sample size to detect an effect.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Forced vital capacity (FVC), forced expiratory volume in 1 second (FEV1), abnormal chest x-ray findings
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what has been done and to be conceptually reproducible with access to the analytic data.
	Metric 15:	Statistical Analysis	N/A	Not applicable, no formal statistical analysis was conducted.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	N/A	Not applicable, no biomarkers of exposure were measured.
	Metric 17:	Effect Biomarker	High	All studied effect biomarkers were demonstrated to be related to adverse health outcomes and were collected from serum or urine samples.
	Metric 18:	Method Sensitivity	N/A	Limits of detection not relevant for the medical testing employed.
	Metric 19:	Biomarker Stability	Low	There is no description of the storage history and/or stability data for urine or serum samples.
	Metric 20:	Sample Contamination	Medium	There is no discussion of contamination.
	Metric 21:	Method Requirements	Medium	While no formal descriptions are provided, given the medical context of the examinations there is some confidence that accurate detection methodologies were used.
	Metric 22:	Matrix Adjustment	Medium	No discussion of matrices is described. This would be relevant for effect biomarkers measured in urine, but there is no evidence that there was no adjustment for creatinine.

Additional Comments: This occupational health surveillance study focused on several coal tar facilities in the United States. At the Chicago plant phthalic anhydride was present. The aim of the study was to compare medical findings across a wide range of health outcomes to established clinical ranges, thus no formal statistical analysis was available that compared exposed vs. unexposed. An examination of the reported data does not indicate that there were significant differences between workers exposed to phthalic anhydride in the Chicago plant and the overall population of workers, or workers at the Chicago plant not reported to be exposed to phthalic anhydride; however, this cannot be completely determined without a formal statistical analysis. There were several large concerns raised with the study, including the lack of an exposure measurement. All employees in the phthalic anhydride plant were considered exposed, however, without surveillance data or employment records, there is a large potential for exposure misclassification. There were also sparse details regarding recruitment, and a lack of consideration of potentially relevant covariates such as age. The workers in this study were all exposed to other occupational agents, thus confounding by other exposures cannot be ruled out.

Overall Quality Determination Low

Study Citation:	Wernfors, M., Nielsen, J., Schütz, A., Skerfving, S. (1986). Phthalic anhydride-induced occupational asthma. International Archives of Allergy and Applied Immunology 79(1):77-82.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Irritation of the upper airways, rhinitis, asthma, PA-induced asthma, chronic productive bronchitis, spirometry
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176303

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	This occupational cohort study examined workers from two plants that utilized flaked phthalic anhydride. Two separate groups were recruited - current employees, and former employees. There is no discussion of how the currently employed were recruited or selected, and it is unclear whether the 48 current employees represent a large proportion of the total eligible population and whether or not the selected employees were similar to the eligible population. More details were provided for former employees, as the study specifies that inquiries were mailed to 85 former employees, and that 70 (82%) replied. It is not specified whether they contacted all former employees, and there is no other description of the total eligible population.Four plants (A, B, C, and D) are mentioned as part of the study, but exposure is only described for plants A and B, and no further description of plants is provided. It is unclear whether participants were recruited from all plants, or just A and B. The study later states that asthma was reported in Plant A by 11 of 42 former and 5 of 28 present employees, and in Plant B (4/23 and 1/12). These numbers imply that 8 present employees and 5 former employees may have come from Plant C and D to reach the total numbers of 48 current employees and 70 former employees, but this information is not explicitly spelled out, and the differences in numbers may be due to other factors. Due to a lack of detailed information, selection bias cannot be ruled out.
	Metric 2:	Attrition	High	No attrition was reported. Results tables do not indicate any loss or exclusion.
	Metric 3:	Comparison Group	Low	The study authors do not address whether groups were similar, and covariates were not addressed in the statistical analysis. No demographic information (beyond smoking) was provided.
Domain 2: Exposure Characterization				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Wernfors, M., Nielsen, J., Schütz, A., Skerfving, S. (1986). Phthalic anhydride-induced occupational asthma. International Archives of Allergy and Applied Immunology 79(1):77-82.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Irritation of the upper airways, rhinitis, asthma, PA-induced asthma, chronic productive bronchitis, spirometry
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176303

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	Low	PA levels were determined in plants A and B by use of two different methods. Some plant A personal samples were collected in glass fiber filters and PA was determined using UV spectrophotometry. The LOD for this method was 0.1 mg/m3 per 200L air sample. For other samples in plant A, and those in plant B, personal samples were collected using battery-operated pumps and respirable particle samplers in some cases. In the second method was determined using HPLC, and the LOD was 0.03 mg/m3 per 20L air sample. The methods for exposure assessment in plants C and D are not described, and it is unclear whether participants were even recruited from those two plants. However, estimates previously calculated above in Metric 1 imply that the total number of participants from Plant C and D may have been only 13, a relatively small proportion of the total sample. It is unclear whether these exposure levels were detected among the actual participants of the present study, and it is unclear whether the exposure levels are applicable to former employees. There is potential for exposure misclassification due to the inconsistency of exposure measurements and the uncertainty of whether these exposure measurements apply to the total sample population.
	Metric 5:	Exposure Levels	Low	PA measurements were reported by plant and work operation. Participants were all treated as ”exposed”. Some comparisons were made with former employees, however, it’s not clear if they were assigned a separate exposure level.
	Metric 6:	Temporality	Medium	The study authors describe a ”latency period between start of employment and onset of respiratory symptoms” which ”ranged from 1 month to 16 years.” Temporality was established, however, it was not entirely clear if these were all incident cases.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Questionnaires were used to assess the presence of relevant conditions (e.g., chronic bronchitis, asthma, etc.). Spirometry was performed using a Vitalograph, with specific pathological spirometric values defined. A subset of participants were subjected to skin-prick testing with 15 common allergens - these specific allergens are not provided, but the source of the allergens is specified to be Allergologisk Laboratorium (Copenhagen, Denmark). Diagnostic definitions for positive skin-prick tests are provided.
	Metric 8:	Reporting Bias	Medium	Outcomes mentioned in the abstract, introduction, and methods were mostly addressed. Study authors provided spirometry results only for asthmatic employees. Most results were simple counts or proportions. Ranges were provided in some cases.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Wernfors, M., Nielsen, J., Schütz, A., Skerfving, S. (1986). Phthalic anhydride-induced occupational asthma. International Archives of Allergy and Applied Immunology 79(1):77-82.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Irritation of the upper airways, rhinitis, asthma, PA-induced asthma, chronic productive bronchitis, spirometry
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176303

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	Low	Covariates were largely not addressed. There was some qualitative consideration of smoking in the analysis, but no formal stratification. The range of employment duration is provided, but it is not clear how the range relates to exposure and outcome groups. Other potential covariates, such as age, are not described. Thus, considerations were not made for confounders and the distribution of primary covariates and potential co-founders is not reported.
	Metric 10:	Covariate Characterization	Medium	Smoking was assessed through survey and questionnaire. Duration of employment was likely assessed via company records.
	Metric 11:	Co-exposure Counfounding	Medium	There is some potential for co-exposures in this setting, particularly with other anhydrides (i.e., maelic and trimellitic), however, the study authors describe the use of these materials as infrequent.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	Low	Prevalence and severity of symptoms were analyzed in an occupational cohort. The design of the study limited understanding dose-response relationships and clear comparisons with unexposed or less exposed individuals.
	Metric 13:	Statistical Power	Medium	Statistical power was not reported, however, the number of current and former employees was likely sufficient to detect an effect.
	Metric 14:	Reproducibility of Analyses	Low	The analysis was not described in detail, as it is unclear which exact Plants participants originated from and how exactly exposure was determined.
	Metric 15:	Statistical Analysis	High	Chi-square tests were used to compare different groups (i.e., current and former employees, asthmatic vs non-asthmatic, etc.). No concerns were identified.

Additional Comments: This study utilized current and former employees of a phthalic anhydride resin plant to investigate the effects of PAD on respiratory and immune endpoints. There were a few concerns regarding this study, including the participation rate and lack of detail regarding recruitment of current workers. Additionally, there was some uncertainty on exposure classification. Exposure values were only measured in two of four plants, and exposure was assumed for all employees. The analytical design also limits the ability to assess differences in outcomes by exposure level.

Overall Quality Determination Low

Study Citation:	Riboli, E., Bai, E., Berrino, F., Merisi, A. (1983). Mortality from lung cancer in an acetylene and phthalic anhydride plant: a case-referent study. Scandinavian Journal of Work, Environment and Health 9(6):455-462.
Health Outcome(s) Assessed:	Mortality
Reported Health Effect(s):	Lung cancer mortality
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63774

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	In this retrospective case-control study in Italy, subjects were deceased residents of Pioltello, notably near a Societa ltaliana Serie Acetica Sintetica (SISAS) chemical factory where workers were exposed to phthalic anhydride. Subjects were recruited without knowledge of exposure status. Male subjects were identified, through the municipal death register of Pioltello, who had a cause of death attributed to lung cancer between January 1, 1976 – December 31, 1979. Referents were made up of the first 2 persons listed after each death record in the register, within 5 years of age as the case, who did not die of respiratory cancer. Age and date of death were comparable between cases and referents; sex of the referents were not specified. 53 cases and 106 referents were selected; setting, inclusion and exclusion criteria, and methods of case ascertainment were reported. The study reports that the death register is likely complete since reporting is compulsory by law.
	Metric 2:	Attrition	Medium	All subjects, once identified through the death register, were traced via contact with next of kin (typically widows, sometimes with other relatives). 6 of 53 cases and 7 of 106 referents were not able to be traced; this is reportedly due to subject's families moving or having no relatives living in the town, which are reasonable justifications not expected to be related to the subjects' exposure or outcome status. Further exclusions were made if they could not confirm death due to primary lung cancer (as opposed to other cancers that metastasized into the lungs) via identification in the Regional Discharge Diagnosis Information System, or if the subject's family did not confirm primary lung cancer. 4 subjects were determined to have cancer with lung metastasis and were thus removed from the analysis. The final study size consisted of 43 cases and 99 referents. There is no evidence that attrition was inappropriately handled.
	Metric 3:	Comparison Group	Medium	There is indirect evidence that cases and controls are similar. Cases and controls were recruited from the same eligible population within the same time frame using the register of deaths of the Pioltello municipality. Referents were made up of the first 2 persons listed after each death record in the register, within 5 years of age as the case, who did not die of respiratory cancer. Age and date of death were noted to be comparable between cases and referents; sex of the referents is not specified (cases only included males). None of the referents had a diagnosis of lung cancer based on the Regional Discharge Diagnosis Information System or family interviews. There are some potential concerns for healthy worker bias, as not all participants were reported to be employed. However, this is somewhat mitigated by some analyses that compare workers occupationally exposed at the SISAS plant to those exposed to potential lung carcinogens at other workplaces.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Riboli, E., Bai, E., Berrino, F., Merisi, A. (1983). Mortality from lung cancer in an acetylene and phthalic anhydride plant: a case-referent study. Scandinavian Journal of Work, Environment and Health 9(6):455-462.
Health Outcome(s) Assessed:	Mortality
Reported Health Effect(s):	Lung cancer mortality
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63774

Domain	Metric		Rating	Comments
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Low	An attempt was made to obtain complete occupational histories through interviews with next of kin; information on job duties and period of employment were collected. The study was designed to determine if exposures of phthalic anhydride and other suspected carcinogens (soot and phthalates) from employment at one chemical manufacturing plant (SISAS, 2-20 years) resulted in lung cancers. The study did not measure specific exposures to phthalic anhydride. In addition, it was noted the possibility of exposure to lung carcinogens in a few subjects may have occurred while working for other companies. Certain job functions and duties were not associated specifically with only phthalic anhydride exposure. The use of interviews with next of kin to determine job function is not the most reliable, and has potential for exposure misclassification, but this is not expected to the large enough to significantly impact the results. Exposure assessment was reported to have been completed without knowledge of case status.
	Metric 5:	Exposure Levels	Low	Exposure was categorized into three groups: exposed at the SISAS plant, exposed to lung carcinogens at other places of employment, or not exposed to occupational lung carcinogens. There is no quantitative information on exposure levels or range of exposures for subjects.
	Metric 6:	Temporality	High	Temporality of exposure and outcome is established; it was noted that the period after production (exposure) was compatible with the latency period between exposure and the outcome (death from lung cancer)
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	The lung cancer mortality outcome was assessed using information from the Register of Deaths of the Pioltello municipality death certificates and validated using the Regional Discharge Diagnosis Information system and interviews with next of kin by medical doctors, in order to distinguish between primary lung cancer and other cancer that metastasized to the lungs. There is overall high certainty in the outcome definition.
	Metric 8:	Reporting Bias	Medium	The outcome of death by lung cancer in cases and referents were reported. Authors reported the number of cases in each exposure group in the table. Odds ratios for the risk of dying from lung cancer were reported with respective confidence intervals.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Riboli, E., Bai, E., Berrino, F., Merisi, A. (1983). Mortality from lung cancer in an acetylene and phthalic anhydride plant: a case-referent study. Scandinavian Journal of Work, Environment and Health 9(6):455-462.
Health Outcome(s) Assessed:	Mortality
Reported Health Effect(s):	Lung cancer mortality
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63774

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	Medium	Cases and controls were matched for age (within 5 years). Considerations of age and date of death were accounted for in selection of subjects. Occupational history and smoking habits were also considered as covariates; light smokers (<10 cigarettes/day) were combined with those who never smoked. Sex and socioeconomic status were not adjusted for; it is reported that cases included males only; however it is not specified if referents only included males. Results were stratified by cigarette use and age. While the study does not explain in-depth why covariates were or were not chosen, there is no evidence of substantial residual confounding.
	Metric 10:	Covariate Characterization	Medium	Potential covariates were assessed through the death records (age of death) and interviews with next of kin that established work histories and smoking habits using recall questions from next of kin. There is potential error in asking next of kin to report on smoking habits, but other covariates were assessed using adequate methods.
	Metric 11:	Co-exposure Counfounding	Low	The study reported that SISAS employees were potentially exposed to many chemicals including acetylene, phthalic anhydride and their derivatives, soot and phthalates. The study was designed to assess all occupational exposures at SISAS and risk of death by lung cancer and not for any individual chemical. Co-exposures were not adjusted for. Exposures to lung carcinogens at other employers was possible, and no information on relative exposure levels was provided.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The study design was adequate to assess the association between exposure to all occupational exposures in the SISAS plant and risk of death by lung cancer, but not for any specific chemical exposure. The study authors noted they used statistical methods described by Mantel And Haenszel (1959, 1963) to derive chi square associations, point estimates for odds ratios, and trend tests. Standardized risk ratios estimations were conducted using programs developed by Rothman and Boice. The study's approach is appropriate to assess the relationship between potential phthalic anhydride exposure and lung cancer mortality.
	Metric 13:	Statistical Power	Medium	Statistical power was not calculated. The OR estimates for SISAS exposure and dying by lung cancer was based on a small number of cases in each level of smokers resulting in large variability (n=43). The number of cases and controls is adequate to detect an effect in the total exposed population; but not for subgroups of the population age and smoking category, as evidenced by wide confidence intervals reported for analyses stratified by cigarette exposure.
	Metric 14:	Reproducibility of Analyses	Medium	The methods are described in sufficient detail to reproduce the results with access to the analytic data.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Riboli, E., Bai, E., Berrino, F., Merisi, A. (1983). Mortality from lung cancer in an acetylene and phthalic anhydride plant: a case-referent study. Scandinavian Journal of Work, Environment and Health 9(6):455-462.
Health Outcome(s) Assessed:	Mortality
Reported Health Effect(s):	Lung cancer mortality
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63774

Domain	Metric	Rating	Comments
	Metric 15: Statistical Analysis	High	The study calculates the chi-square for association as well as odds ratios using the Mantel and Haenszel method. The computation is sufficiently reported and appropriate, with test-based confidence limits calculated as well.

Additional Comments: A retrospective case-referent study was conducted in a population in Milano Province, Italy where there is an acetylene and phthalic anhydride manufacturing plant (SISAS). Cases of deaths by lung cancer (n=43) were determined from death records (and validated through clinical and interviews of next of kin) and compared to referents (n=99). Standardized risk ratios (adjusted for age and smoking habits) for dying with lung cancer associated with occupational exposure to a number of chemicals (phthalic anhydride, acetylene, and their derivatives, soot, phthalates) were calculated. The study was designed to determine if occupational exposures from employment at the chemical manufacturing plant (SISAS) or exposures to suspected carcinogens from other occupations resulted in lung cancers. Exposures were categorized as exposure from subjects who worked at the SISAS plant for 2-20 years (S+), subjects who have ever worked at a job likely to have exposure to lung carcinogens (E+), or no occupational exposure (E-). There is no quantitative information on exposure levels or range of exposures and risk ratios were not calculated for exposure specifically to phthalic anhydride. The risk of dying due to lung cancer of subjects employed at SISAS was 5.6 (95% CI 1.9-16.2) compared to referents; the risk for those occupationally exposed to lung carcinogens, but not employed at SISAS was 1.7% (95% CI 0.9-3.5). It was noted by the authors that exposure to soot and phthalates at SISAS contributed to the increased risk of lung cancer.

Overall Quality Determination Medium

Study Citation:	Riboli, E., Bai, E., Berrino, F., Merisi, A. (1983). Mortality from lung cancer in an acetylene and phthalic anhydride plant: a case-referent study. Scandinavian Journal of Work, Environment and Health 9(6):455-462.
Health Outcome(s) Assessed:	Cancer/Carcinogenesis
Reported Health Effect(s):	Lung cancer mortality
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63774

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	In this retrospective case-control study in Italy, subjects were deceased residents of Pioltello, notably near a Societa ltaliana Serie Acetica Sintetica (SISAS) chemical factory where workers were exposed to phthalic anhydride. Subjects were recruited without knowledge of exposure status. Male subjects were identified, through the municipal death register of Pioltello, who had a cause of death attributed to lung cancer between January 1, 1976 – December 31, 1979. Referents were made up of the first 2 persons listed after each death record in the register, within 5 years of age as the case, who did not die of respiratory cancer. Age and date of death were comparable between cases and referents; sex of the referents were not specified. 53 cases and 106 referents were selected; setting, inclusion and exclusion criteria, and methods of case ascertainment were reported. The study reports that the death register is likely complete since reporting is compulsory by law.
	Metric 2:	Attrition	Medium	All subjects, once identified through the death register, were traced via contact with next of kin (typically widows, sometimes with other relatives). 6 of 53 cases and 7 of 106 referents were not able to be traced; this is reportedly due to subject's families moving or having no relatives living in the town, which are reasonable justifications not expected to be related to the subjects' exposure or outcome status. Further exclusions were made if they could not confirm death due to primary lung cancer (as opposed to other cancers that metastasized into the lungs) via identification in the Regional Discharge Diagnosis Information System, or if the subject's family did not confirm primary lung cancer. 4 subjects were determined to have cancer with lung metastasis and were thus removed from the analysis. The final study size consisted of 43 cases and 99 referents. There is no evidence that attrition was inappropriately handled.
	Metric 3:	Comparison Group	Medium	There is indirect evidence that cases and controls are similar. Cases and controls were recruited from the same eligible population within the same time frame using the register of deaths of the Pioltello municipality. Referents were made up of the first 2 persons listed after each death record in the register, within 5 years of age as the case, who did not die of respiratory cancer. Age and date of death were noted to be comparable between cases and referents; sex of the referents is not specified (cases only included males). None of the referents had a diagnosis of lung cancer based on the Regional Discharge Diagnosis Information System or family interviews. There are some potential concerns for healthy worker bias, as not all participants were reported to be employed. However, this is somewhat mitigated by some analyses that compare workers occupationally exposed at the SISAS plant to those exposed to potential lung carcinogens at other workplaces.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Riboli, E., Bai, E., Berrino, F., Merisi, A. (1983). Mortality from lung cancer in an acetylene and phthalic anhydride plant: a case-referent study. Scandinavian Journal of Work, Environment and Health 9(6):455-462.
Health Outcome(s) Assessed:	Cancer/Carcinogenesis
Reported Health Effect(s):	Lung cancer mortality
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63774

Domain	Metric		Rating	Comments
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Low	An attempt was made to obtain complete occupational histories through interviews with next of kin; information on job duties and period of employment were collected. The study was designed to determine if exposures of phthalic anhydride and other suspected carcinogens (soot and phthalates) from employment at one chemical manufacturing plant (SISAS, 2-20 years) resulted in lung cancers. The study did not measure specific exposures to phthalic anhydride. In addition, it was noted the possibility of exposure to lung carcinogens in a few subjects may have occurred while working for other companies. Certain job functions and duties were not associated specifically with only phthalic anhydride exposure. The use of interviews with next of kin to determine job function is not the most reliable, and has potential for exposure misclassification, but this is not expected to the large enough to significantly impact the results. Exposure assessment was reported to have been completed without knowledge of case status.
	Metric 5:	Exposure Levels	Low	Exposure was categorized into three groups: exposed at the SISAS plant, exposed to lung carcinogens at other places of employment, or not exposed to occupational lung carcinogens. There is no quantitative information on exposure levels or range of exposures for subjects.
	Metric 6:	Temporality	High	Temporality of exposure and outcome is established; it was noted that the period after production (exposure) was compatible with the latency period between exposure and the outcome (death from lung cancer)
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	The lung cancer mortality outcome was assessed using information from the Register of Deaths of the Pioltello municipality death certificates and validated using the Regional Discharge Diagnosis Information system and interviews with next of kin by medical doctors, in order to distinguish between primary lung cancer and other cancer that metastasized to the lungs. There is overall high certainty in the outcome definition.
	Metric 8:	Reporting Bias	Medium	The outcome of death by lung cancer in cases and referents were reported. Authors reported the number of cases in each exposure group in the table. Odds ratios for the risk of dying from lung cancer were reported with respective confidence intervals.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Riboli, E., Bai, E., Berrino, F., Merisi, A. (1983). Mortality from lung cancer in an acetylene and phthalic anhydride plant: a case-referent study. Scandinavian Journal of Work, Environment and Health 9(6):455-462.
Health Outcome(s) Assessed:	Cancer/Carcinogenesis
Reported Health Effect(s):	Lung cancer mortality
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63774

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	Medium	Cases and controls were matched for age (within 5 years). Considerations of age and date of death were accounted for in selection of subjects. Occupational history and smoking habits were also considered as covariates; light smokers (<10 cigarettes/day) were combined with those who never smoked. Sex and socioeconomic status were not adjusted for; it is reported that cases included males only; however it is not specified if referents only included males. Results were stratified by cigarette use and age. While the study does not explain in-depth why covariates were or were not chosen, there is no evidence of substantial residual confounding.
	Metric 10:	Covariate Characterization	Medium	Potential covariates were assessed through the death records (age of death) and interviews with next of kin that established work histories and smoking habits using recall questions from next of kin. There is potential error in asking next of kin to report on smoking habits, but other covariates were assessed using adequate methods.
	Metric 11:	Co-exposure Counfounding	Low	The study reported that SISAS employees were potentially exposed to many chemicals including acetylene, phthalic anhydride and their derivatives, soot and phthalates. The study was designed to assess all occupational exposures at SISAS and risk of death by lung cancer and not for any individual chemical. Co-exposures were not adjusted for. Exposures to lung carcinogens at other employers was possible, and no information on relative exposure levels was provided.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The study design was adequate to assess the association between exposure to all occupational exposures in the SISAS plant and risk of death by lung cancer, but not for any specific chemical exposure. The study authors noted they used statistical methods described by Mantel And Haenszel (1959, 1963) to derive chi square associations, point estimates for odds ratios, and trend tests. Standardized risk ratios estimations were conducted using programs developed by Rothman and Boice. The study's approach is appropriate to assess the relationship between potential phthalic anhydride exposure and lung cancer mortality.
	Metric 13:	Statistical Power	Medium	Statistical power was not calculated. The OR estimates for SISAS exposure and dying by lung cancer was based on a small number of cases in each level of smokers resulting in large variability (n=43). The number of cases and controls is adequate to detect an effect in the total exposed population; but not for subgroups of the population age and smoking category, as evidenced by wide confidence intervals reported for analyses stratified by cigarette exposure.
	Metric 14:	Reproducibility of Analyses	Medium	The methods are described in sufficient detail to reproduce the results with access to the analytic data.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Riboli, E., Bai, E., Berrino, F., Merisi, A. (1983). Mortality from lung cancer in an acetylene and phthalic anhydride plant: a case-referent study. Scandinavian Journal of Work, Environment and Health 9(6):455-462.
Health Outcome(s) Assessed:	Cancer/Carcinogenesis
Reported Health Effect(s):	Lung cancer mortality
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63774

Domain	Metric	Rating	Comments
	Metric 15: Statistical Analysis	High	The study calculates the chi-square for association as well as odds ratios using the Mantel and Haenszel method. The computation is sufficiently reported and appropriate, with test-based confidence limits calculated as well.

Additional Comments: A retrospective case-referent study was conducted in a population in Milano Province, Italy where there is an acetylene and phthalic anhydride manufacturing plant (SISAS). Cases of deaths by lung cancer (n=43) were determined from death records (and validated through clinical and interviews of next of kin) and compared to referents (n=99). Standardized risk ratios (adjusted for age and smoking habits) for dying with lung cancer associated with occupational exposure to a number of chemicals (phthalic anhydride, acetylene, and their derivatives, soot, phthalates) were calculated. The study was designed to determine if occupational exposures from employment at the chemical manufacturing plant (SISAS) or exposures to suspected carcinogens from other occupations resulted in lung cancers. Exposures were categorized as exposure from subjects who worked at the SISAS plant for 2-20 years (S+), subjects who have ever worked at a job likely to have exposure to lung carcinogens (E+), or no occupational exposure (E-). There is no quantitative information on exposure levels or range of exposures and risk ratios were not calculated for exposure specifically to phthalic anhydride. The risk of dying due to lung cancer of subjects employed at SISAS was 5.6 (95% CI 1.9-16.2) compared to referents; the risk for those occupationally exposed to lung carcinogens, but not employed at SISAS was 1.7% (95% CI 0.9-3.5). It was noted by the authors that exposure to soot and phthalates at SISAS contributed to the increased risk of lung cancer.

Overall Quality Determination Medium

Study Citation:	TOMA, (1979). 1978 Cross-sectional health study of workers at the Bridgeville plant of Koppers Company, Inc.
Health Outcome(s) Assessed:	Cancer/Carcinogenesis
Reported Health Effect(s):	Malignant skin tumors
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	1480908 Linked HERO ID(s): 1480908, 63805

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	In this cross-sectional study, a plant producing phthalic acid in Bridgeville, Pennsylvania was investigated in 1978. Workers were recruited via a letter prepared by plant management to each employee. Of 285 total workers at the plant, 174 workers participated in the study for at least one outcome (61%). The study reports that the distribution by age between participants and the total workforce was similar, but the study specifies that they were unable to compare years of service between participants and non-participants. Years of service may correlate to higher levels of exposure and a higher likelihood of outcome incidence; however, there is no evidence that those who refused to participate had higher levels of exposure or were more likely to experience adverse health outcomes.
	Metric 2:	Attrition	Medium	1 individual out of 174 participants did not participate in clinical testing, and five only partially completed clinical testing. A total of 25 did not participate in the physician's examination. While no specific reasons are provided for non-participation, the overall amount of attrition is relatively minimal. Participants were removed from their specific analysis if they were missing relevant outcome data, which is an acceptable way to handle these missing data.
	Metric 3:	Comparison Group	Low	In this cross-sectional study, the prevalence of skin conditions was compared to the prevalence of skin conditions in USA white males age 34-44. The source for this data is not described, and the sex and race of the workforce sample are not described.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Low	Phthalic acid was reported to be used in the plant, and there was potential exposure among all participants. Plant officials identified workers who routinely worked in areas where phthalic anhydride was found, and a special testing protocol specific to those workers was identified (Process Group A). The extent to which other participants were exposed to phthalic anhydride is unclear. No quantitative measures of phthalic anhydride were reported.
	Metric 5:	Exposure Levels	Low	While no analysis contains multiple levels of exposure, there are separate analyses performed on the total sample of workers and those in Process Group A (n=6), who are expected to be exposed to higher levels of phthalic anhydride. Thus, it is possible to compare the number of outcomes reported in the total group compared to the subgroup, serving as two separate levels of exposure. However, only these two exposure levels are available for analysis.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1979). 1978 Cross-sectional health study of workers at the Bridgeville plant of Koppers Company, Inc.
Health Outcome(s) Assessed:	Cancer/Carcinogenesis
Reported Health Effect(s):	Malignant skin tumors
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	1480908 Linked HERO ID(s): 1480908, 63805

Domain	Metric		Rating	Comments
	Metric 6:	Temporality	Low	The temporality of exposure and outcome is uncertain. While exposure definitely occurred before the outcome was measured, the cross-sectional nature of the outcome examination precludes certainty that participants did not experience adverse outcomes prior to exposure. 50% of the sample were employed for 5 years or less, thus it cannot be determined whether the outcome status of workers is recent.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Low	Specific methods were not described for any outcome measurement, including skin conditions. All outcomes were reported to be either laboratory tests or physical examinations by medical professionals. While there may be some confidence in the assessment given the medical environment, the lack of specificity of tests for any outcome is concerning. It is also not specified whether examiners knew whether participants were expected to be highly exposed to phthalic anhydride via Process Group A.
	Metric 8:	Reporting Bias	High	All outlined outcomes are reported in the results. For skin conditions, the number of participants examined and the number of abnormal outcomes observed are reported. Observed and expected percentages are presented for each skin condition.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Low	Data are presented on the distribution of participants by years of service and age, but these are not formally adjusted for in a statistical model. Participants are compared to ”expected” averages from white US males adults aged 34-44. No stratification is performed by age, race, or sex.
	Metric 10:	Covariate Characterization	Medium	Questionnaires were provided in the appendix. It was not reported whether this was a validated questionnaire, however, there was no evidence to suggest it was an invalid instrument.
	Metric 11:	Co-exposure Counfounding	Low	Exposure to other occupational chemicals is described, such as benzene, naphthalene, and acrylonitrile. Exposure to other chemicals is contextualized in some but not all interpretations, and no quantitative data is reported on levels of exposure. Results are provided for specific ”process groups” where exposure is likely to be higher for specific chemicals. However, this does not rule out co-exposure, and the process group for phthalic anhydride includes another chemical (naphthalene).
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cross-sectional study design was sufficient to detect whether there were clinical abnormalities among a sample of workers in a chemical plant.
	Metric 13:	Statistical Power	Low	Statistical power was not reported. A total of 143 individuals completed all exams; but the incidence of outcomes (n=4 at maximum) is likely too small to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	It would be possible to reproduce this study given access to the analytic data.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1979). 1978 Cross-sectional health study of workers at the Bridgeville plant of Koppers Company, Inc.
Health Outcome(s) Assessed:	Cancer/Carcinogenesis
Reported Health Effect(s):	Malignant skin tumors
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	1480908 Linked HERO ID(s): 1480908, 63805

Domain	Metric	Rating	Comments
	Metric 15: Statistical Analysis	N/A	Description of analyses was not present but appears to be simple comparisons to reference ranges. A chi-square goodness of fit test is described, but only in the context of differences in age categories between participants.

Additional Comments: This occupational surveillance study examined numerous employees from a facility utilizing phthalic anhydride and assessed common clinical indicators of adverse effects for the respiratory, hepatic, renal, integumentary, and hematology/immune systems. Simple comparisons were made between recorded values in participants with historical reference ranges; there is an exception for integumentary outcomes, where they compared the prevalence of outcomes in the sample to the expected prevalence among US white males aged 34-44. There were several concerns with this study, including lack of robust exposure assessment for phthalic anhydride which could lead to severe exposure misclassification, and a lack of detailed reporting on how exactly outcomes were measured. There is also a lack of clarity in the statistical analyses used, primarily in their determination of ”significant.” The incidence of all skin conditions was lower in the workforce sample than would be expected compared to the reference population, but the incidence of outcomes was likely too low to detect an effect.

Overall Quality Determination Low

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Cancer/Carcinogenesis
Reported Health Effect(s):	Sputum cytology, urine cytology, C-reactive protein
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	In this cross-sectional occupational surveillance study, workers from 9 Koppers Coal Tar plants were examined from January-October 1979. Only one of these plants (the Chicago plant) produced phthalic anhydride. Participation at the relevant phthalic anhydride plant (i.e., the Chicago plant) was approximately 46% (105/230), while participation from the total workforce across all plants was 51% (453/888). The study authors note that participation at the Chicago plant was non-representative for hourly workers relative to salaried workers. The study does not discuss recruitment or selection processes in detail but does indicate that workers could participate in the study during normal work hours without impact on their pay. Overall, details are sparse and it is unclear if those who did not participate may have been more or less exposed compared to those who were included.
	Metric 2:	Attrition	High	One participant out of 105 in the Chicago did not have a blood test taken and was excluded from analyses that included serum measures. No reason is provided for why the blood test was not taken, but there was no other indication of attrition or exclusion in the Chicago plant group or in the total included group.
	Metric 3:	Comparison Group	Low	Demographic details on sex and race are provided for all plants. The distribution of race and sex is roughly similar in the Chicago plant to the overall distribution of race and sex across the included workforce. However, these demographic differences are not controlled for in statistical analyses. There is also no discussion of age and whether groups had differences based on age.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Low	In comparisons between different plants, all Chicago plant workers were grouped together. Descriptions of the Chicago plant were very limited beyond stating that phthalic anhydride was used. A sub-group analysis of the Chicago plant separates out workers by ”type” (Phthalic, Maleic, or Coal Tar). It is unclear if phthalic anhydride workers were all actually formally exposed to phthalic anhydride, and it is not clear whether the other types of workers were not exposed. There is a large potential for exposure misclassification if not all tasks or employees were in contact with phthalic anhydride, and since exposure was estimated solely using professional judgment, exposure misclassification cannot be ruled out.
	Metric 5:	Exposure Levels	Low	The study only reports two levels of exposure, exposed and unexposed. No quantitative exposure information is available.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Cancer/Carcinogenesis
Reported Health Effect(s):	Sputum cytology, urine cytology, C-reactive protein
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
	Metric 6:	Temporality	Low	Work history or duration of employment was not discussed in the analysis. It is clear that the outcomes were measured after some degree of exposure, since outcomes were assessed in an occupational context. However, it is unclear how long these workers may have been exposed, and it is unclear whether the employees were free of any of the reported health outcomes prior to enrollment into the study. The temporality of exposure and outcome is uncertain.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Specific methods were not described, but description of the study suggests standard clinical methods were used (in line with NIOSH recommendations). Examination forms were provided in the appendices.
	Metric 8:	Reporting Bias	Medium	All outlined outcomes are reported in the results, however, data is provided as the proportion of individuals with ”abnormal” measurements compared against clinical standards.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Low	There was no evidence of adjustment for potential confounders, however, some were discussed in-text for medical interpretations. The distribution of sex and race is provided across exposure group, but there is no discussion of age. Covariates were not adjusted for or stratified in relation to outcomes.
	Metric 10:	Covariate Characterization	Medium	Questionnaires were provided in the appendix. It was not reported whether this was a validated questionnaire, however, there was no evidence to suggest it was an invalid instrument.
	Metric 11:	Co-exposure Counfounding	Low	Exposure to other coal tar components was described, including potential health effects. Workers in plants other than the Chicago plant were exposed to other occupational agents, and it is unclear whether those exposures contributed to the incidence of health outcomes. Within the Chicago plant, the study splits the sample into workers exposed to phthalic anhydride, maleic anhydride, or coal tar. It is not confirmed that phthalic anhydride workers were only exposed to phthalic anhydride.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The design limited the ability to determine which exposures were relevant to each health effect, however the design is sufficient to answer the study’s question as to whether there are ”abnormal” medical findings among workers at the studied coal tar plants.
	Metric 13:	Statistical Power	Low	Statistical power was not reported. A total of 105 employees completed exams from the Chicago plant. However, the analysis within the Chicago plant only identifies 14 workers who were exposed to phthalic anhydride, which may be too small of a sample size to detect an effect.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Cancer/Carcinogenesis
Reported Health Effect(s):	Sputum cytology, urine cytology, C-reactive protein
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what has been done and to be conceptually reproducible with access to the analytic data.
	Metric 15:	Statistical Analysis	N/A	Not applicable, no formal statistical analysis was conducted.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	N/A	Not applicable, no biomarkers of exposure were measured.
	Metric 17:	Effect Biomarker	High	All studied effect biomarkers were demonstrated to be related to adverse health outcomes and were collected from serum or urine samples.
	Metric 18:	Method Sensitivity	N/A	Limits of detection not relevant for the medical testing employed.
	Metric 19:	Biomarker Stability	Low	There is no description of the storage history and/or stability data for urine or serum samples.
	Metric 20:	Sample Contamination	Medium	There is no discussion of contamination.
	Metric 21:	Method Requirements	Medium	While no formal descriptions are provided, given the medical context of the examinations there is some confidence that accurate detection methodologies were used.
	Metric 22:	Matrix Adjustment	Medium	No discussion of matrices is described. This would be relevant for effect biomarkers measured in urine, but there is no evidence that there was no adjustment for creatinine.

Additional Comments: This occupational health surveillance study focused on several coal tar facilities in the United States. At the Chicago plant phthalic anhydride was present. The aim of the study was to compare medical findings across a wide range of health outcomes to established clinical ranges, thus no formal statistical analysis was available that compared exposed vs. unexposed. An examination of the reported data does not indicate that there were significant differences between workers exposed to phthalic anhydride in the Chicago plant and the overall population of workers, or workers at the Chicago plant not reported to be exposed to phthalic anhydride; however, this cannot be completely determined without a formal statistical analysis. There were several large concerns raised with the study, including the lack of an exposure measurement. All employees in the phthalic anhydride plant were considered exposed, however, without surveillance data or employment records, there is a large potential for exposure misclassification. There were also sparse details regarding recruitment, and a lack of consideration of potentially relevant covariates such as age. The workers in this study were all exposed to other occupational agents, thus confounding by other exposures cannot be ruled out.

Overall Quality Determination Low

Study Citation:	TOMA, (1982). Occupational health evaluation of the Bridgeville, Pennsylvania plant of Koppers Company, Inc. Organic Material Group. Final report [86870001543].
Health Outcome(s) Assessed:	Cancer/Carcinogenesis
Reported Health Effect(s):	Malignant skin tumors
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63805 Linked HERO ID(s): 1480908, 63805

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	In this cross-sectional study, a plant producing phthalic acid in Bridgeville, Pennsylvania was investigated in 1981 after an initial investigation in 1978 as described in HEROID: 1480908. This updated investigation uses a modified version of the original protocol. This iteration of the study does not specify how exactly employees were recruited, but in the 1978 study employees were recruited via a letter prepared by the management company. Of 246 total workers at the plant, 139 workers participated in the study for at least one outcome (56.5%). The study reports that "there do not appear to be major differences in participation among different age groups." The study was unable to compare years of service between participants and non-participants. Years of service may correlate to higher levels of exposure and a higher likelihood of outcome incidence; however, there is no evidence that those who refused to participate had higher levels of exposure or were more likely to experience adverse health outcomes. 82/139 participants were also evaluated in the 1978 study.
	Metric 2:	Attrition	High	Only three individuals did not complete the full medical evaluation. One participant resigned prior to completion, and two refused the physician's exam. There is no evidence to suggest this would appreciably impact the results, or that attrition was related to exposure and outcome.
	Metric 3:	Comparison Group	Low	Subjects were compared with historical reference ranges for significant skin pathology obtained from the National Center for Health Statistics on 1/26/1977. The age/sex/race distribution of the source population used to compute the prevalence of skin condition is not specified, which may be inappropriate given that the study sample is 92.8% male.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Low	Phthalic acid was reported to be used in the plant, and there was potential exposure among all participants. However, there is no specific evidence that any workers were exposed to phthalic acid, and no quantitative measures are available. Thus, exposure was only estimated using professional judgment.
	Metric 5:	Exposure Levels	Low	Workers were compared to a general population, thus there was only a comparison between exposed/unexposed.
	Metric 6:	Temporality	Low	The temporality of exposure and outcome is uncertain.
Domain 3: Outcome Assessment				

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1982). Occupational health evaluation of the Bridgeville, Pennsylvania plant of Koppers Company, Inc. Organic Material Group. Final report [86870001543].
Health Outcome(s) Assessed:	Cancer/Carcinogenesis
Reported Health Effect(s):	Malignant skin tumors
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63805 Linked HERO ID(s): 1480908, 63805

Domain	Metric		Rating	Comments
	Metric 7:	Outcome Measurement or Characterization	Low	Specific methods were not described for any outcome measurement, including skin conditions. All outcomes were reported to be either laboratory tests or physical examinations by medical professionals. While there may be some confidence in the assessment given the medical environment, the lack of specificity of tests for any outcome is concerning.
	Metric 8:	Reporting Bias	Medium	All outlined outcomes are reported in the results. For skin conditions, the number of participants examined and the number of abnormal outcomes observed are reported.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Low	Data are presented on the distribution of participants by years of service and age, but these are not formally adjusted for in a statistical model. No stratification is performed by age, race, or sex, which may be relevant for the comparison the reference population from the National Center for Health Statistics.
	Metric 10:	Covariate Characterization	Medium	Questionnaires were provided in the appendix. It was not reported whether this was a validated questionnaire, however, there was no evidence to suggest it was an invalid instrument.
	Metric 11:	Co-exposure Counfounding	Low	Exposure to other occupational chemicals is described, such as benzene, naphthalene, and acrylonitrile. Exposure to other chemicals is contextualized in some but not all interpretations, and no quantitative data is reported on levels of exposure. Per the first investigation in HEROID 1480908, workers were assigned to different process areas where exposures were expected to be different. Thus, there is direct evidence that there was an unbalanced provision of co-exposures across the primary study groups that were not appropriately adjusted for
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cross-sectional study design was sufficient to detect whether there were clinical abnormalities among a sample of workers in a chemical plant.
	Metric 13:	Statistical Power	Medium	Statistical power was not reported. A total of 143 individuals completed all exams; but the incidence of outcomes (n=6 at maximum) is likely too small to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	It would be possible to reproduce this study given access to the analytic data.
	Metric 15:	Statistical Analysis	N/A	Description of analyses was not present but appears to be simple comparisons to reference ranges. A chi-square goodness of fit test is described, but only in the context of differences in age categories between participants.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1982). Occupational health evaluation of the Bridgeville, Pennsylvania plant of Koppers Company, Inc. Organic Material Group. Final report [86870001543].
Health Outcome(s) Assessed:	Cancer/Carcinogenesis
Reported Health Effect(s):	Malignant skin tumors
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63805 Linked HERO ID(s): 1480908, 63805

Domain	Metric	Rating	Comments
Additional Comments:	This occupational surveillance study examined numerous employees from a facility utilizing phthalic anhydride. Simple comparisons were made with historical reference ranges. There were several issues with this study, including lack of robust exposure assessment for phthalic anhydride which could lead to severe exposure misclassification. Numerous other occupational exposures were described (i.e., benzene).This 1981 occupational surveillance study was a follow-up investigation to one conducted in 1978 (HEROID: 1480908) with a different subset of the workforce; however, 82 subjects participated in both investigations. The study examined numerous employees from a facility utilizing phthalic anhydride and assessed common clinical indicators of adverse effects for the respiratory, hepatic, renal, integumentary, and hematology/immune systems. Simple comparisons were made between recorded values in participants with historical reference ranges; there is an exception for integumentary outcomes, where they compared the prevalence of outcomes in the sample to the expected prevalence per the National Center for Health Statistics data. There were several concerns with this study, including lack of robust exposure assessment for phthalic anhydride which could lead to severe exposure misclassification, and a lack of detailed reporting on how exactly outcomes were measured. There is also a lack of clarity in the statistical analyses used, primarily in their determination of ”significant.” The incidence of all skin conditions was lower in the workforce sample than would be expected compared to the reference population, but the incidence of outcomes was likely too low to detect an effect.		

Overall Quality Determination Low

Study Citation:	TOMA, (1979). 1978 Cross-sectional health study of workers at the Bridgeville plant of Koppers Company, Inc.
Health Outcome(s) Assessed:	Skin and Connective Tissue
Reported Health Effect(s):	Fungal skin infection, acne, benign growths, malignant skin tumors, keratosis, eczema, folliculitis
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	1480908 Linked HERO ID(s): 1480908, 63805

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	In this cross-sectional study, a plant producing phthalic acid in Bridgeville, Pennsylvania was investigated in 1978. Workers were recruited via a letter prepared by plant management to each employee. Of 285 total workers at the plant, 174 workers participated in the study for at least one outcome (61%). The study reports that the distribution by age between participants and the total workforce was similar, but the study specifies that they were unable to compare years of service between participants and non-participants. Years of service may correlate to higher levels of exposure and a higher likelihood of outcome incidence; however, there is no evidence that those who refused to participate had higher levels of exposure or were more likely to experience adverse health outcomes.
	Metric 2:	Attrition	Medium	1 individual out of 174 participants did not participate in clinical testing, and five only partially completed clinical testing. A total of 25 did not participate in the physician's examination. While no specific reasons are provided for non-participation, the overall amount of attrition is relatively minimal. Participants were removed from their specific analysis if they were missing relevant outcome data, which is an acceptable way to handle these missing data.
	Metric 3:	Comparison Group	Low	In this cross-sectional study, the prevalence of skin conditions was compared to the prevalence of skin conditions in USA white males age 34-44. The source for this data is not described, and the sex and race of the workforce sample are not described.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Low	Phthalic acid was reported to be used in the plant, and there was potential exposure among all participants. Plant officials identified workers who routinely worked in areas where phthalic anhydride was found, and a special testing protocol specific to those workers was identified (Process Group A). The extent to which other participants were exposed to phthalic anhydride is unclear. No quantitative measures of phthalic anhydride were reported.
	Metric 5:	Exposure Levels	Low	While no analysis contains multiple levels of exposure, there are separate analyses performed on the total sample of workers and those in Process Group A (n=6), who are expected to be exposed to higher levels of phthalic anhydride. Thus, it is possible to compare the number of outcomes reported in the total group compared to the subgroup, serving as two separate levels of exposure. However, only these two exposure levels are available for analysis.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1979). 1978 Cross-sectional health study of workers at the Bridgeville plant of Koppers Company, Inc.
Health Outcome(s) Assessed:	Skin and Connective Tissue
Reported Health Effect(s):	Fungal skin infection, acne, benign growths, malignant skin tumors, keratosis, eczema, folliculitis
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	1480908 Linked HERO ID(s): 1480908, 63805

Domain	Metric		Rating	Comments
	Metric 6:	Temporality	Low	The temporality of exposure and outcome is uncertain. While exposure definitely occurred before the outcome was measured, the cross-sectional nature of the outcome examination precludes certainty that participants did not experience adverse outcomes prior to exposure. 50% of the sample were employed for 5 years or less, thus it cannot be determined whether the outcome status of workers is recent.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Low	Specific methods were not described for any outcome measurement, including skin conditions. All outcomes were reported to be either laboratory tests or physical examinations by medical professionals. While there may be some confidence in the assessment given the medical environment, the lack of specificity of tests for any outcome is concerning. It is also not specified whether examiners knew whether participants were expected to be highly exposed to phthalic anhydride via Process Group A.
	Metric 8:	Reporting Bias	High	All outlined outcomes are reported in the results. For skin conditions, the number of participants examined and the number of abnormal outcomes observed are reported. Observed and expected percentages are presented for each skin condition.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Low	Data are presented on the distribution of participants by years of service and age, but these are not formally adjusted for in a statistical model. Participants are compared to ”expected” averages from white US males adults aged 34-44. No stratification is performed by age, race, or sex.
	Metric 10:	Covariate Characterization	Medium	Questionnaires were provided in the appendix. It was not reported whether this was a validated questionnaire, however, there was no evidence to suggest it was an invalid instrument.
	Metric 11:	Co-exposure Counfounding	Low	Exposure to other occupational chemicals is described, such as benzene, naphthalene, and acrylonitrile. Exposure to other chemicals is contextualized in some but not all interpretations, and no quantitative data is reported on levels of exposure. Results are provided for specific ”process groups” where exposure is likely to be higher for specific chemicals. However, this does not rule out co-exposure, and the process group for phthalic anhydride includes another chemical (naphthalene).
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cross-sectional study design was sufficient to detect whether there were clinical abnormalities among a sample of workers in a chemical plant.
	Metric 13:	Statistical Power	Low	Statistical power was not reported. A total of 143 individuals completed all exams; but the incidence of outcomes (n=4 at maximum) is likely too small to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	It would be possible to reproduce this study given access to the analytic data.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1979). 1978 Cross-sectional health study of workers at the Bridgeville plant of Koppers Company, Inc.
Health Outcome(s) Assessed:	Skin and Connective Tissue
Reported Health Effect(s):	Fungal skin infection, acne, benign growths, malignant skin tumors, keratosis, eczema, folliculitis
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	1480908 Linked HERO ID(s): 1480908, 63805

Domain	Metric	Rating	Comments
	Metric 15: Statistical Analysis	N/A	Description of analyses was not present but appears to be simple comparisons to reference ranges. A chi-square goodness of fit test is described, but only in the context of differences in age categories between participants.

Additional Comments: This occupational surveillance study examined numerous employees from a facility utilizing phthalic anhydride and assessed common clinical indicators of adverse effects for the respiratory, hepatic, renal, integumentary, and hematology/immune systems. Simple comparisons were made between recorded values in participants with historical reference ranges; there is an exception for integumentary outcomes, where they compared the prevalence of outcomes in the sample to the expected prevalence among US white males aged 34-44. There were several concerns with this study, including lack of robust exposure assessment for phthalic anhydride which could lead to severe exposure misclassification, and a lack of detailed reporting on how exactly outcomes were measured. There is also a lack of clarity in the statistical analyses used, primarily in their determination of ”significant.” The incidence of all skin conditions was lower in the workforce sample than would be expected compared to the reference population, but the incidence of outcomes was likely too low to detect an effect.

Overall Quality Determination Low

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Skin and Connective Tissue
Reported Health Effect(s):	Benign skin growths, other skin conditions (keratosis, eczema, folliculitis)
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	In this cross-sectional occupational surveillance study, workers from 9 Koppers Coal Tar plants were examined from January-October 1979. Only one of these plants (the Chicago plant) produced phthalic anhydride. Participation at the relevant phthalic anhydride plant (i.e., the Chicago plant) was approximately 46% (105/230), while participation from the total workforce across all plants was 51% (453/888). The study authors note that participation at the Chicago plant was non-representative for hourly workers relative to salaried workers. The study does not discuss recruitment or selection processes in detail but does indicate that workers could participate in the study during normal work hours without impact on their pay. Overall, details are sparse and it is unclear if those who did not participate may have been more or less exposed compared to those who were included.
	Metric 2:	Attrition	High	One participant out of 105 in the Chicago did not have a blood test taken and was excluded from analyses that included serum measures. No reason is provided for why the blood test was not taken, but there was no other indication of attrition or exclusion in the Chicago plant group or in the total included group.
	Metric 3:	Comparison Group	Low	Demographic details on sex and race are provided for all plants. The distribution of race and sex is roughly similar in the Chicago plant to the overall distribution of race and sex across the included workforce. However, these demographic differences are not controlled for in statistical analyses. There is also no discussion of age and whether groups had differences based on age.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Low	In comparisons between different plants, all Chicago plant workers were grouped together. Descriptions of the Chicago plant were very limited beyond stating that phthalic anhydride was used. A sub-group analysis of the Chicago plant separates out workers by ”type” (Phthalic, Maleic, or Coal Tar). It is unclear if phthalic anhydride workers were all actually formally exposed to phthalic anhydride, and it is not clear whether the other types of workers were not exposed. There is a large potential for exposure misclassification if not all tasks or employees were in contact with phthalic anhydride, and since exposure was estimated solely using professional judgment, exposure misclassification cannot be ruled out.
	Metric 5:	Exposure Levels	Low	The study only reports two levels of exposure, exposed and unexposed. No quantitative exposure information is available.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Skin and Connective Tissue
Reported Health Effect(s):	Benign skin growths, other skin conditions (keratosis, eczema, folliculitis)
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
	Metric 6:	Temporality	Low	Work history or duration of employment was not discussed in the analysis. It is clear that the outcomes were measured after some degree of exposure, since outcomes were assessed in an occupational context. However, it is unclear how long these workers may have been exposed, and it is unclear whether the employees were free of any of the reported health outcomes prior to enrollment into the study. The temporality of exposure and outcome is uncertain.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Specific methods were not described, but description of the study suggests standard clinical methods were used (in line with NIOSH recommendations). Examination forms were provided in the appendices.
	Metric 8:	Reporting Bias	Medium	All outlined outcomes are reported in the results, however, data is provided as the proportion of individuals with ”abnormal” measurements compared against clinical standards.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Low	There was no evidence of adjustment for potential confounders, however, some were discussed in-text for medical interpretations. The distribution of sex and race is provided across exposure group, but there is no discussion of age. Covariates were not adjusted for or stratified in relation to outcomes.
	Metric 10:	Covariate Characterization	Medium	Questionnaires were provided in the appendix. It was not reported whether this was a validated questionnaire, however, there was no evidence to suggest it was an invalid instrument.
	Metric 11:	Co-exposure Counfounding	Low	Exposure to other coal tar components was described, including potential health effects. Workers in plants other than the Chicago plant were exposed to other occupational agents, and it is unclear whether those exposures contributed to the incidence of health outcomes. Within the Chicago plant, the study splits the sample into workers exposed to phthalic anhydride, maleic anhydride, or coal tar. It is not confirmed that phthalic anhydride workers were only exposed to phthalic anhydride.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The design limited the ability to determine which exposures were relevant to each health effect, however the design is sufficient to answer the study’s question as to whether there are ”abnormal” medical findings among workers at the studied coal tar plants.
	Metric 13:	Statistical Power	Low	Statistical power was not reported. A total of 105 employees completed exams from the Chicago plant. However, the analysis within the Chicago plant only identifies 14 workers who were exposed to phthalic anhydride, which may be too small of a sample size to detect an effect.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Skin and Connective Tissue
Reported Health Effect(s):	Benign skin growths, other skin conditions (keratosis, eczema, folliculitis)
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what has been done and to be conceptually reproducible with access to the analytic data.
	Metric 15:	Statistical Analysis	N/A	Not applicable, no formal statistical analysis was conducted.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	N/A	Not applicable, no biomarkers of exposure were measured.
	Metric 17:	Effect Biomarker	High	All studied effect biomarkers were demonstrated to be related to adverse health outcomes and were collected from serum or urine samples.
	Metric 18:	Method Sensitivity	N/A	Limits of detection not relevant for the medical testing employed.
	Metric 19:	Biomarker Stability	Low	There is no description of the storage history and/or stability data for urine or serum samples.
	Metric 20:	Sample Contamination	Medium	There is no discussion of contamination.
	Metric 21:	Method Requirements	Medium	While no formal descriptions are provided, given the medical context of the examinations there is some confidence that accurate detection methodologies were used.
	Metric 22:	Matrix Adjustment	Medium	No discussion of matrices is described. This would be relevant for effect biomarkers measured in urine, but there is no evidence that there was no adjustment for creatinine.

Additional Comments: This occupational health surveillance study focused on several coal tar facilities in the United States. At the Chicago plant phthalic anhydride was present. The aim of the study was to compare medical findings across a wide range of health outcomes to established clinical ranges, thus no formal statistical analysis was available that compared exposed vs. unexposed. An examination of the reported data does not indicate that there were significant differences between workers exposed to phthalic anhydride in the Chicago plant and the overall population of workers, or workers at the Chicago plant not reported to be exposed to phthalic anhydride; however, this cannot be completely determined without a formal statistical analysis. There were several large concerns raised with the study, including the lack of an exposure measurement. All employees in the phthalic anhydride plant were considered exposed, however, without surveillance data or employment records, there is a large potential for exposure misclassification. There were also sparse details regarding recruitment, and a lack of consideration of potentially relevant covariates such as age. The workers in this study were all exposed to other occupational agents, thus confounding by other exposures cannot be ruled out.

Overall Quality Determination — Low

Study Citation:	TOMA, (1982). Occupational health evaluation of the Bridgeville, Pennsylvania plant of Koppers Company, Inc. Organic Material Group. Final report [86870001543].
Health Outcome(s) Assessed:	Skin and Connective Tissue
Reported Health Effect(s):	Fungal skin infection, acne, benign growths, malignant skin tumors, keratosis, eczema, folliculitis
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63805 Linked HERO ID(s): 1480908, 63805

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	In this cross-sectional study, a plant producing phthalic acid in Bridgeville, Pennsylvania was investigated in 1981 after an initial investigation in 1978 as described in HEROID: 1480908. This updated investigation uses a modified version of the original protocol. This iteration of the study does not specify how exactly employees were recruited, but in the 1978 study employees were recruited via a letter prepared by the management company. Of 246 total workers at the plant, 139 workers participated in the study for at least one outcome (56.5%). The study reports that "there do not appear to be major differences in participation among different age groups." The study was unable to compare years of service between participants and non-participants. Years of service may correlate to higher levels of exposure and a higher likelihood of outcome incidence; however, there is no evidence that those who refused to participate had higher levels of exposure or were more likely to experience adverse health outcomes. 82/139 participants were also evaluated in the 1978 study.
	Metric 2:	Attrition	High	Only three individuals did not complete the full medical evaluation. One participant resigned prior to completion, and two refused the physician's exam. There is no evidence to suggest this would appreciably impact the results, or that attrition was related to exposure and outcome.
	Metric 3:	Comparison Group	Low	Subjects were compared with historical reference ranges for significant skin pathology obtained from the National Center for Health Statistics on 1/26/1977. The age/sex/race distribution of the source population used to compute the prevalence of skin condition is not specified, which may be inappropriate given that the study sample is 92.8% male.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Low	Phthalic acid was reported to be used in the plant, and there was potential exposure among all participants. However, there is no specific evidence that any workers were exposed to phthalic acid, and no quantitative measures are available. Thus, exposure was only estimated using professional judgment.
	Metric 5:	Exposure Levels	Low	Workers were compared to a general population, thus there was only a comparison between exposed/unexposed.
	Metric 6:	Temporality	Low	The temporality of exposure and outcome is uncertain.
Domain 3: Outcome Assessment				

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1982). Occupational health evaluation of the Bridgeville, Pennsylvania plant of Koppers Company, Inc. Organic Material Group. Final report [86870001543].
Health Outcome(s) Assessed:	Skin and Connective Tissue
Reported Health Effect(s):	Fungal skin infection, acne, benign growths, malignant skin tumors, keratosis, eczema, folliculitis
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63805 Linked HERO ID(s): 1480908, 63805

Domain	Metric		Rating	Comments
	Metric 7:	Outcome Measurement or Characterization	Low	Specific methods were not described for any outcome measurement, including skin conditions. All outcomes were reported to be either laboratory tests or physical examinations by medical professionals. While there may be some confidence in the assessment given the medical environment, the lack of specificity of tests for any outcome is concerning.
	Metric 8:	Reporting Bias	Medium	All outlined outcomes are reported in the results. For skin conditions, the number of participants examined and the number of abnormal outcomes observed are reported.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Low	Data are presented on the distribution of participants by years of service and age, but these are not formally adjusted for in a statistical model. No stratification is performed by age, race, or sex, which may be relevant for the comparison the reference population from the National Center for Health Statistics.
	Metric 10:	Covariate Characterization	Medium	Questionnaires were provided in the appendix. It was not reported whether this was a validated questionnaire, however, there was no evidence to suggest it was an invalid instrument.
	Metric 11:	Co-exposure Counfounding	Low	Exposure to other occupational chemicals is described, such as benzene, naphthalene, and acrylonitrile. Exposure to other chemicals is contextualized in some but not all interpretations, and no quantitative data is reported on levels of exposure. Per the first investigation in HEROID 1480908, workers were assigned to different process areas where exposures were expected to be different. Thus, there is direct evidence that there was an unbalanced provision of co-exposures across the primary study groups that were not appropriately adjusted for
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cross-sectional study design was sufficient to detect whether there were clinical abnormalities among a sample of workers in a chemical plant.
	Metric 13:	Statistical Power	Medium	Statistical power was not reported. A total of 143 individuals completed all exams; but the incidence of outcomes (n=6 at maximum) is likely too small to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	It would be possible to reproduce this study given access to the analytic data.
	Metric 15:	Statistical Analysis	N/A	Description of analyses was not present but appears to be simple comparisons to reference ranges. A chi-square goodness of fit test is described, but only in the context of differences in age categories between participants.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1982). Occupational health evaluation of the Bridgeville, Pennsylvania plant of Koppers Company, Inc. Organic Material Group. Final report [86870001543].
Health Outcome(s) Assessed:	Skin and Connective Tissue
Reported Health Effect(s):	Fungal skin infection, acne, benign growths, malignant skin tumors, keratosis, eczema, folliculitis
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	63805 Linked HERO ID(s): 1480908, 63805

Domain	Metric	Rating	Comments
Additional Comments:	This occupational surveillance study examined numerous employees from a facility utilizing phthalic anhydride. Simple comparisons were made with historical reference ranges. There were several issues with this study, including lack of robust exposure assessment for phthalic anhydride which could lead to severe exposure misclassification. Numerous other occupational exposures were described (i.e., benzene).This 1981 occupational surveillance study was a follow-up investigation to one conducted in 1978 (HEROID: 1480908) with a different subset of the workforce; however, 82 subjects participated in both investigations. The study examined numerous employees from a facility utilizing phthalic anhydride and assessed common clinical indicators of adverse effects for the respiratory, hepatic, renal, integumentary, and hematology/immune systems. Simple comparisons were made between recorded values in participants with historical reference ranges; there is an exception for integumentary outcomes, where they compared the prevalence of outcomes in the sample to the expected prevalence per the National Center for Health Statistics data. There were several concerns with this study, including lack of robust exposure assessment for phthalic anhydride which could lead to severe exposure misclassification, and a lack of detailed reporting on how exactly outcomes were measured. There is also a lack of clarity in the statistical analyses used, primarily in their determination of ”significant.” The incidence of all skin conditions was lower in the workforce sample than would be expected compared to the reference population, but the incidence of outcomes was likely too low to detect an effect.		

Overall Quality Determination Low

Study Citation:	Mettang, T., Thomas, S., Kiefer, T., Fischer, F. P., Kuhlmann, U., Wodarz, R., Rettenmeier, A. W. (1996). Uraemic pruritus and exposure to di(2-ethylhexyl) phthalate (DEHP) in haemodialysis patients. Nephrology, Dialysis, Transplantation 11(12):2439-2443.
Health Outcome(s) Assessed:	Renal/Kidney
Reported Health Effect(s):	uraemic pruritus score
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	673485

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	Out of 27 patients eligible for inclusion in the study, 6 patients were excluded. Three patients were excluded due to pre-existing conditions and one was excluded because they were taking anti-histaminic medication, which are both reasonable metrics for exclusion. However, two patients were excluded because they were ”unable to complete the questionnaire.” It is possible that these patients had more extreme cases of ureamic pruritius and thus should have been included to best represent the eligible population. However, since this only applies to two patients it is unlikely to have a significant effect on the exposure-outcome distribution.
	Metric 2:	Attrition	High	There is no reported loss to follow-up or attrition in the study. Exposure and outcome data is complete for all patients.
	Metric 3:	Comparison Group	High	Controls and cases were both pulled from the same population of adult patients on regular hemodialysis treatment during the same timeframe. The authors report that 7 out of 21 patients included did not report symptoms of ureamic pruritus, and thus were used as controls in the study.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Blood samples were taken before and after a 4-hour dialysis session for each patient and then analyzed for di(2-ethylhexyl) phthalate and its hydrolysis products, which includes phthalic acid.
	Metric 5:	Exposure Levels	Low	Serum concentrations of phthalic acid are reported separately for cases and controls, both pre and post hemodialysis. For cases, the mean serum concentration of phthalic acid was 0.207 ug/ml (SD: 0.106) pre-HD and 0.135 ug/ml (SD: 0.081) post-HD. For controls, the mean serum concentration was 0.160 ug/ml (SD: 0.101) pre-HD and 0.097 ug/ml (SD: 0.07) post-HD. A Kruskal-Walis test found no significant difference in concentration between cases and controls.
	Metric 6:	Temporality	Low	The study measures phthalic acid concentrations before and after hemodialysis - however, the outcome of uraemic pruritus is only measured once pre-hemodialysis. There is no evidence that exposure may have predated the outcome. The study was designed to investigate whether post-HD concentrations of phthalic acid were higher in cases of ureamic pruritus, but this does not establish temporality. The authors mention that they recorded temporality of pruritus in relation to hemodialysis when assessing cases of pruritus, but it is unclear how this information was used.
Domain 3: Outcome Assessment				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Mettang, T., Thomas, S., Kiefer, T., Fischer, F. P., Kuhlmann, U., Wodarz, R., Rettenmeier, A. W. (1996). Uraemic pruritus and exposure to di(2-ethylhexyl) phthalate (DEHP) in haemodialysis patients. Nephrology, Dialysis, Transplantation 11(12):2439-2443.
Health Outcome(s) Assessed:	Renal/Kidney
Reported Health Effect(s):	uraemic pruritus score
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	673485

Domain	Metric		Rating	Comments
	Metric 7:	Outcome Measurement or Characterization	Low	The study authors questioned patients as to whether they had symptoms of pruritus. It is unclear whether, or how, they distinguished uraemic pruritus and regular pruritus. They potentially used a question about its temporal relationship to dialysis, but this is not explicitly stated. The authors also used a modified "Liu Jing Duo [1987]" score to evaluate the intensity, distribution, and sleep disturbance. No information was accessible on the method's validity.
	Metric 8:	Reporting Bias	High	A description of all measured outcomes is provided, as are effect estimates in the form of mean values of serum phthalic acid for cases and controls with standard deviations. Spearman rank correlation tests were also performed and reported for pruritus score and phthalic acid concentration.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Low	No information is provided regarding covariate adjustment in the methods section. The results section mentions that "no correlation was found between serum triglyceride or cholesterol levels, age, or duration of dialysis and pre- or post-dialysis concentrations of DEHP, MEHP, PA or 2-EH." However, no data is shown demonstrating this. No information on differences in covariates between cases and controls is provided.
	Metric 10:	Covariate Characterization	Low	It is unclear how exactly covariates were characterized. It is assumed that serum triglyceride and cholesterol concentrations were measured using reasonably valid methods. Age and duration of dialysis are assumed to be based off of questionnaire data.
	Metric 11:	Co-exposure Counfounding	Low	Potential co-exposures di(2-ethylhexyl) phthalate, mono-(2-ethylhexyl phthalate, and 2-ethylhexanol were evaluated. However, the statistical methods used to examine PA exposures and pruritus (Spearman rank correlation and Kruskal-Wallis tests) do not accommodate adjustment for confounders.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	Low	The study design does not ensure temporality. Spearman rank correlations and Kruskal-Wallis tests were used to assess PA exposures and pruritus, instead of more appropriate methods that would allow for control of confounders (e.g., regression analyses).
	Metric 13:	Statistical Power	Low	Overall participation is low - only 21 patients were included in the final analysis and only 7 patients were specified as controls.
	Metric 14:	Reproducibility of Analyses	Medium	The methods section adequately describes the steps taken to achieve the results provided.
	Metric 15:	Statistical Analysis	Low	The assumption that the variables are independent and ordinal are met by the statistical modeling building process. However, the assumption that the relationship between phthalic acid concentrations and pruritus is monotonic is not clarified or assessed.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Mettang, T., Thomas, S., Kiefer, T., Fischer, F. P., Kuhlmann, U., Wodarz, R., Rettenmeier, A. W. (1996). Uraemic pruritus and exposure to di(2-ethylhexyl) phthalate (DEHP) in haemodialysis patients. Nephrology, Dialysis, Transplantation 11(12):2439-2443.
Health Outcome(s) Assessed:	Renal/Kidney
Reported Health Effect(s):	uraemic pruritus score
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	673485

Domain	Metric		Rating	Comments
	Metric 16:	Use of Biomarker of Exposure	Medium	Phthalic acid is directly measured from serum samples of participants. However, phthalic acid is the general metabolite of many phthalate esters.
	Metric 17:	Effect Biomarker	N/A	Not applicable, no biomarkers of effect were measured.
	Metric 18:	Method Sensitivity	Medium	The limit of detection is stated as 1 ng/ml serum for phthalic acid, or .001 ug/ml. No information is provided regarding the number of samples above or below the limit of detection, but the mean minus the standard deviation for phthalic acid is not lower than .001 ug/ml.
	Metric 19:	Biomarker Stability	Medium	The study design states that blood samples were immediately transferred to glass tubes, centrifuged, and stored at -20 deg C until analysis. No information is provided regarding potential losses due to thawing.
	Metric 20:	Sample Contamination	Medium	No information is provided regarding potential sample contamination.
	Metric 21:	Method Requirements	Medium	Quantitative analysis was performed by selected ion monitoring gas chromatography/mass spectrometry.
	Metric 22:	Matrix Adjustment	N/A	Not Applicable, no matrix adjustment necessary due to measurement in serum

Additional Comments: This case-control study uses Spearman rank correlation tests to assess the relationship between serum phthalic acid pre and post hemodialysis and a score for uraemic pruritus. No correlation was found between concentrations of phthalic acid and intensity of pruritus. Additionally, serum concentrations of phthalic acid did not differ significantly between cases of uraemic phthalic and controls. Concentrations of phthalic acid were found to significantly decrease post hemodialysis compared to pre hemodialysis. The study had significant issues with temporality and statistical modeling that limit the quality of the results.

Overall Quality Determination — Low

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Renal/Kidney
Reported Health Effect(s):	Blood urea nitrogen, cellular casts in urine, red blood cells in urine, white blood cells in urine, serum total protein, urine albumin, hyaline casts in urine
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	In this cross-sectional occupational surveillance study, workers from 9 Koppers Coal Tar plants were examined from January-October 1979. Only one of these plants (the Chicago plant) produced phthalic anhydride. Participation at the relevant phthalic anhydride plant (i.e., the Chicago plant) was approximately 46% (105/230), while participation from the total workforce across all plants was 51% (453/888). The study authors note that participation at the Chicago plant was non-representative for hourly workers relative to salaried workers. The study does not discuss recruitment or selection processes in detail but does indicate that workers could participate in the study during normal work hours without impact on their pay. Overall, details are sparse and it is unclear if those who did not participate may have been more or less exposed compared to those who were included.
	Metric 2:	Attrition	High	One participant out of 105 in the Chicago did not have a blood test taken and was excluded from analyses that included serum measures. No reason is provided for why the blood test was not taken, but there was no other indication of attrition or exclusion in the Chicago plant group or in the total included group.
	Metric 3:	Comparison Group	Low	Demographic details on sex and race are provided for all plants. The distribution of race and sex is roughly similar in the Chicago plant to the overall distribution of race and sex across the included workforce. However, these demographic differences are not controlled for in statistical analyses. There is also no discussion of age and whether groups had differences based on age.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Low	In comparisons between different plants, all Chicago plant workers were grouped together. Descriptions of the Chicago plant were very limited beyond stating that phthalic anhydride was used. A sub-group analysis of the Chicago plant separates out workers by ”type” (Phthalic, Maleic, or Coal Tar). It is unclear if phthalic anhydride workers were all actually formally exposed to phthalic anhydride, and it is not clear whether the other types of workers were not exposed. There is a large potential for exposure misclassification if not all tasks or employees were in contact with phthalic anhydride, and since exposure was estimated solely using professional judgment, exposure misclassification cannot be ruled out.
	Metric 5:	Exposure Levels	Low	The study only reports two levels of exposure, exposed and unexposed. No quantitative exposure information is available.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Renal/Kidney
Reported Health Effect(s):	Blood urea nitrogen, cellular casts in urine, red blood cells in urine, white blood cells in urine, serum total protein, urine albumin, hyaline casts in urine
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
	Metric 6:	Temporality	Low	Work history or duration of employment was not discussed in the analysis. It is clear that the outcomes were measured after some degree of exposure, since outcomes were assessed in an occupational context. However, it is unclear how long these workers may have been exposed, and it is unclear whether the employees were free of any of the reported health outcomes prior to enrollment into the study. The temporality of exposure and outcome is uncertain.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Specific methods were not described, but description of the study suggests standard clinical methods were used (in line with NIOSH recommendations). Examination forms were provided in the appendices.
	Metric 8:	Reporting Bias	Medium	All outlined outcomes are reported in the results, however, data is provided as the proportion of individuals with ”abnormal” measurements compared against clinical standards.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Low	There was no evidence of adjustment for potential confounders, however, some were discussed in-text for medical interpretations. The distribution of sex and race is provided across exposure group, but there is no discussion of age. Covariates were not adjusted for or stratified in relation to outcomes.
	Metric 10:	Covariate Characterization	Medium	Questionnaires were provided in the appendix. It was not reported whether this was a validated questionnaire, however, there was no evidence to suggest it was an invalid instrument.
	Metric 11:	Co-exposure Counfounding	Low	Exposure to other coal tar components was described, including potential health effects. Workers in plants other than the Chicago plant were exposed to other occupational agents, and it is unclear whether those exposures contributed to the incidence of health outcomes. Within the Chicago plant, the study splits the sample into workers exposed to phthalic anhydride, maleic anhydride, or coal tar. It is not confirmed that phthalic anhydride workers were only exposed to phthalic anhydride.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The design limited the ability to determine which exposures were relevant to each health effect, however the design is sufficient to answer the study’s question as to whether there are ”abnormal” medical findings among workers at the studied coal tar plants.
	Metric 13:	Statistical Power	Low	Statistical power was not reported. A total of 105 employees completed exams from the Chicago plant. However, the analysis within the Chicago plant only identifies 14 workers who were exposed to phthalic anhydride, which may be too small of a sample size to detect an effect.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Renal/Kidney
Reported Health Effect(s):	Blood urea nitrogen, cellular casts in urine, red blood cells in urine, white blood cells in urine, serum total protein, urine albumin, hyaline casts in urine
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what has been done and to be conceptually reproducible with access to the analytic data.
	Metric 15:	Statistical Analysis	N/A	Not applicable, no formal statistical analysis was conducted.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	N/A	Not applicable, no biomarkers of exposure were measured.
	Metric 17:	Effect Biomarker	High	All studied effect biomarkers were demonstrated to be related to adverse health outcomes and were collected from serum or urine samples.
	Metric 18:	Method Sensitivity	N/A	Limits of detection not relevant for the medical testing employed.
	Metric 19:	Biomarker Stability	Low	There is no description of the storage history and/or stability data for urine or serum samples.
	Metric 20:	Sample Contamination	Medium	There is no discussion of contamination.
	Metric 21:	Method Requirements	Medium	While no formal descriptions are provided, given the medical context of the examinations there is some confidence that accurate detection methodologies were used.
	Metric 22:	Matrix Adjustment	Medium	No discussion of matrices is described. This would be relevant for effect biomarkers measured in urine, but there is no evidence that there was no adjustment for creatinine.

Additional Comments: This occupational health surveillance study focused on several coal tar facilities in the United States. At the Chicago plant phthalic anhydride was present. The aim of the study was to compare medical findings across a wide range of health outcomes to established clinical ranges, thus no formal statistical analysis was available that compared exposed vs. unexposed. An examination of the reported data does not indicate that there were significant differences between workers exposed to phthalic anhydride in the Chicago plant and the overall population of workers, or workers at the Chicago plant not reported to be exposed to phthalic anhydride; however, this cannot be completely determined without a formal statistical analysis. There were several large concerns raised with the study, including the lack of an exposure measurement. All employees in the phthalic anhydride plant were considered exposed, however, without surveillance data or employment records, there is a large potential for exposure misclassification. There were also sparse details regarding recruitment, and a lack of consideration of potentially relevant covariates such as age. The workers in this study were all exposed to other occupational agents, thus confounding by other exposures cannot be ruled out.

Overall Quality Determination — Low

Study Citation:	Choi, H., Kim, J., Im, Y., Lee, S., Kim, Y. (2012). The association between some endocrine disruptors and hypospadias in biological samples. Journal of Environmental Science and Health, Part A: Toxic/Hazardous Substances & Environmental Engineering 47(13):2173-2179.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	Hypospadias
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	1332536

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	Participants in this study were volunteers from a medical college in Seoul. No other information on participant recruitment, selection, or inclusion/exclusion criteria was provided. Since key elements of the study design are not reported, selection bias cannot be ruled out. The final sample consisted of 80 controls, 80 hypospadias cases, and 40 mothers of hypospadias cases.
	Metric 2:	Attrition	Low	Participation rates/rates of attrition are not reported.
	Metric 3:	Comparison Group	Uninformative	Information was not reported to determine if participants with hypospadias were similar to the control group. The study provides no evidence that controls were free from the outcome, and key confounders (birth weight, maternal age, smoking) are not described, thus potentially critical differences in the case and control groups were not controlled for via matching or statistical adjustment.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Exposure to phthalic acid was measured via gas chromatography/mass spectrometry analysis of urine and serum samples. This is a well-established method that is reliable in determining exposure to phthalic acid. Quantification and analysis methods, including discussion of recovery/accuracy/precision, were extensively discussed.
	Metric 5:	Exposure Levels	Medium	Exposure is assessed as a linear variable, and phthalic acid is measured in all participants (both cases and controls).
	Metric 6:	Temporality	Low	There is no specific information provided on temporality. It is unclear how close urine and serum measurements were made to birth, at which point the outcome of hypospadias was likely to be measured. Despite this uncertainty, there is no evidence of an improper time order.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Uninformative	There is no information provided on how participants with hypospadias were selected, how/whether diagnosis was confirmed, or how controls were confirmed to be without hypospadias.
	Metric 8:	Reporting Bias	Medium	All of the study's analyses described in the methods are reported in the results.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Choi, H., Kim, J., Im, Y., Lee, S., Kim, Y. (2012). The association between some endocrine disruptors and hypospadias in biological samples. Journal of Environmental Science and Health, Part A: Toxic/Hazardous Substances & Environmental Engineering 47(13):2173-2179.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	Hypospadias
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	1332536

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	Low	Covariates/potential confounders are not discussed or mentioned in the methods or results of the paper. The introduction of the study indicates that some factors such as maternal age, birth weight, and smoking are associated with hypospadias, but these are not measured or used for adjustment. However, there is no evidence that these potential confounders differ by exposure.
	Metric 10:	Covariate Characterization	N/A	Not applicable, no confounders were assessed.
	Metric 11:	Co-exposure Counfounding	Medium	Other chemicals studied included DBP, MBP, MEHP, n-nonylphenol, t-octylphenol, and bisphenol A. While these are not controlled for in a multi-pollutant model, separate results are presented for each co-exposure.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	Low	Due to lack of information on participant selection, exposure assessment, and outcome ascertainment, it is difficult to confirm the study design and whether it was appropriate.
	Metric 13:	Statistical Power	Medium	The statistical power of the study was not calculated. However, the number of cases controls (n=80 each) is likely sufficient to detect an effect.
	Metric 14:	Reproducibility of Analyses	Low	Very little information on analytical methods provided. Statistical methods were straightforward, but information on distribution of exposure/outcome, transformations of variables, and handling of missing data or outliers was not provided. The specific statistical method used to determine statistical significance is not stated.
	Metric 15:	Statistical Analysis	Low	The statistical model used to determine p-values is not explicitly stated, which limits the ability to determine the appropriateness of the model creation.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	Phthalic acid is derived from multiple parent chemicals, but given the focus of this review on phthalic acid as a specific exposure, there is no concern related to phthalic acid deriving from multiple sources.
	Metric 17:	Effect Biomarker	N/A	Not applicable, no biomarkers of effect were measured.
	Metric 18:	Method Sensitivity	Medium	The LOD for phthalic acid is not specifically provided, but they indicate that the LOD for all chemicals in the urine samples was 0.09-1.45 ng/mL, and 0.20-2.45 ng/mL in plasma samples. Any of these levels would likely be sufficiently low.
	Metric 19:	Biomarker Stability	Medium	Samples were reported to be stored in glass containers and maintained at -20 Celsius. No stability data is mentioned.
	Metric 20:	Sample Contamination	Medium	Information on contamination not reported.
	Metric 21:	Method Requirements	Medium	Samples were analyzed using GC/MS.
	Metric 22:	Matrix Adjustment	Low	There was no discussion of adjusting urine measures for creatinine.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Choi, H., Kim, J., Im, Y., Lee, S., Kim, Y. (2012). The association between some endocrine disruptors and hypospadias in biological samples. Journal of Environmental Science and Health, Part A: Toxic/Hazardous Substances & Environmental Engineering 47(13):2173-2179.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	Hypospadias
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	1332536

Domain	Metric	Rating	Comments
Additional Comments:	In this study, urine and plasma levels of phthalic acid were compared between cases of hypospadias and controls. There is a critical lack of information overall, including no discussion of control selection and no discussion of how exactly the outcome was determined. Further, the study explicitly discusses the importance of several confounders such as maternal age and birth weight, but these are not measured in the study. Due to the overall lack of information, the quality of this study cannot be determined, and the information available would not be reliable in a hazard assessment.		

Overall Quality Determination Uninformative

Study Citation:	Philips, E. M., Kahn, L. G., Jaddoe, V., V.W., Shao, Y., Asimakopoulos, A. G., Kannan, K., Steegers, P., E.A., Trasande, L. (2018). First trimester urinary bisphenol and phthalate concentrations and time to pregnancy: A population-based cohort analysis. Journal of Clinical Endocrinology and Metabolism 103(9):3540–3547.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	Time to Pregnancy
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	4728822

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Eligibility criteria included enrollment at less than 18 weeks of gestation, information on TTP and infertility treatment use, singleton live born children, a 5-year postpartum visit, and a first trimester urine sample, resulting in a sample of 877 out of 2835. It is possible that those who used infertility treatment self-excluded due to stigma, but this is unlikely to have a significant effect on exposure-outcome distribution.
	Metric 2:	Attrition	High	The study only reported missing data for potential covariates, which were imputed using multiple imputation. The study reports that the percentage of missing values for any given covariate was less than or equal to 10%, except in the case of folic acid supplementation use (13.3%).
	Metric 3:	Comparison Group	High	The study design sufficiently describes the eligible population and inclusion/exclusion criteria. The study also states that demographic characteristics were similar between those who were included and those who were excluded from the analysis.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Concentrations were measured in a spot urine sample taken from each participation during a first trimester visit. Phthalic acid was analyzed as a proxy for total phthalate exposure in urine samples.
	Metric 5:	Exposure Levels	Medium	The range of exposure is sufficiently large, with a median of 55.59 ng/mL and an interquartile range of 29.71 to 118.08 ng/mL. The interquartile range is assumed to be the 25th-75th percentile, but this is not explicitly stated.
	Metric 6:	Temporality	Low	Urinary samples used to measure phthalic acid concentrations were taken during the first trimester of pregnancy, which is after the outcome of time to pregnancy would have been determined. It is unclear whether phthalic acid concentrations measured during early pregnancy accurately reflect exposure levels prior to conception.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Time to pregnancy was assessed using questionnaires during the first trimester enrollment visit. This creates an opportunity for recall bias since it is based on recollection. However, this is unlikely to differ based on exposure levels. For those who used infertility treatments, the time to pregnancy was set at 12 months to match the clinical definition of infertility.
	Metric 8:	Reporting Bias	High	All measured outcomes are reported as fecundability ratios with a 95% confidence interval.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Philips, E. M., Kahn, L. G., Jaddoe, V., V.W., Shao, Y., Asimakopoulos, A. G., Kannan, K., Steegers, P., E.A., Trasande, L. (2018). First trimester urinary bisphenol and phthalate concentrations and time to pregnancy: A population-based cohort analysis. Journal of Clinical Endocrinology and Metabolism 103(9):3540–3547.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	Time to Pregnancy
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	4728822

Domain	Metric		Rating	Comments
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	Maternal age, parity, maternal education, preconception folic acid supplementation, and urinary creatinine concentration were adjusted for in the model. Sensitivity analysis was also performed excluding participants who used infertility treatment.
	Metric 10:	Covariate Characterization	High	Questionnaires were used to evaluate potential confounding variables. Urinary creatinine concentrations were measured using urinary samples during the first trimester of pregnancy.
	Metric 11:	Co-exposure Counfounding	Medium	Other co-exposures included exposure to bisphenols, low-molecular weight phthalate metabolites, high-molecular weight phthalate metabolites, di-2-ethylhexylphthalate metabolites, and di-n-octylphthalate metabolites and were measured along with phthalic acid from urine samples at first trimester. All were adjusted for in phthalic acid analysis.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cohort study design is appropriate for the comparison of phthalic acid levels and time to pregnancy due to time to pregnancy being a common outcome. Cox proportional hazard models with a resulting fecundability ratio are an appropriate method of investigating the effect of phthalic acid exposure.
	Metric 13:	Statistical Power	Medium	The number of participants (n=877) is sufficiently large to guarantee a high enough statistical power. The subgroups that were analyzed were also sufficiently large.
	Metric 14:	Reproducibility of Analyses	Medium	The methods and statistical model building are thoroughly described and could be reproduced given the cohort data.
	Metric 15:	Statistical Analysis	High	The Cox proportional hazard models were transparently built. Multicollinearity tests were performed to assess whether or not covariates should be included in the model. Model assumptions were tested using plots of Schoenfeld residuals for continuous variables and log-minus-log plots for categorical variables. Phthalic acid was ln-transformed to reduce variability.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	Phthalic acid is directly measured in urine samples. However, phthalic acid is a general metabolite of many phthalate esters.
	Metric 17:	Effect Biomarker	N/A	Not applicable, no biomarkers of effect were measured.
	Metric 18:	Method Sensitivity	Medium	The limit of detection was reported as 1.11 ng/mL, and only 0.5% of values were below the limit of detection.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Philips, E. M., Kahn, L. G., Jaddoe, V., V.W., Shao, Y., Asimakopoulos, A. G., Kannan, K., Steegers, P., E.A., Trasande, L. (2018). First trimester urinary bisphenol and phthalate concentrations and time to pregnancy: A population-based cohort analysis. Journal of Clinical Endocrinology and Metabolism 103(9):3540–3547.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	Time to Pregnancy
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	4728822

Domain	Metric		Rating	Comments
	Metric 19:	Biomarker Stability	High	No information is reported regarding storage history or stability in the paper. However, the study is cited as having information regarding transportation (Philips et al. 2004, HERO ID: 4728366) contains detailed information on storage history.
	Metric 20:	Sample Contamination	High	No information is provided regarding potential contamination. The study cited for method details (Philips et al. 2004, HERO ID: 4728366) notes that samples were contamination-free based on analysis of blanks and describes the steps taken to ensure that data are reliable.
	Metric 21:	Method Requirements	High	No information is provided for phthalic acid analysis techniques in the paper, but per study cited for method details (Philips et al. 2004, HERO ID: 4728366) high performance liquid chromatography electrospray ionization-tandem mass spectrometry (HPLC-ESI-MS/MS) was used.
	Metric 22:	Matrix Adjustment	Medium	Phthalic acid was adjusted for urinary creatinine concentration - however, only adjusted results are provided, not unadjusted results.

Additional Comments: This prospective cohort study investigates associations between phthalic acid concentrations in first trimester urine samples with time to pregnancy (assessed via fecundability ratios). Significant associations between phthalic acid and fecundability ratios were found among women who did not use folic acid supplements before conception. The study is overall high quality, with strengths including the exposure assessment and control for confounders. The primary limitation of the study is the temporality of observed associations, as exposures were measured after the outcome of interest had occurred. It is unclear whether phthalic acid concentrations measured during early pregnancy accurately reflect exposure levels prior to conception.

Overall Quality Determination High

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Participants were selected from the Generation R prospective cohort of 8879 women. Selection criteria were stated. Exclusion criteria were lack of blood pressure measurements (n= 140) or women with pre-existing hypertension (n=23), resulting in a total sample of 1233 for analysis. Demographic characteristics of those included were generally similar to those excluded.
	Metric 2:	Attrition	High	The only reported missing data were several covariate measures, which were calculated using multiple imputation. No participants are reported as being lost to follow-up.
	Metric 3:	Comparison Group	High	Baseline demographic characteristics between participants such as age, education level, ethnicity, and parity were considered as potential confounders to adjust for any potential differences by exposure level.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Concentration of phthalic acid was measured directly in spot urine samples.
	Metric 5:	Exposure Levels	Medium	The exposure range is sufficiently large, with a median of 56.99 ng/ml and an inter-quartile range of 30.62-124.05. The inter-quartile range is assumed to be 25th-75th percentile, but this is not explicitly stated. (Supplemental Table 1)
	Metric 6:	Temporality	Medium	Outcome measures are reported during early and mid-pregnancy for blood samples, and umbilical and uterine measures were taken in mid and late pregnancy. Since the exposure is measured once during early pregnancy, a clear temporality is established. However, there is no consideration of whether or not the exposure window is appropriate to precede the effect.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Effect biomarkers were measured from blood samples. Placental vascular resistance was evaluated using flow velocity waveforms, and umbilical artery pulsatility index was measured in a free-floating loop of the umbilical cord. Placental weight and diagnosis of gestational hypertensive disorders were obtained from medical records and measured using standard protocols. Biomarker and placental weight were assessed using gold-standard methodology, and there is no evidence that the methods for hemodynamic function had low validity.
	Metric 8:	Reporting Bias	High	Measured outcomes are reported in tables/figures, and discussed in the text. Effect estimates are reported with 95% confidence intervals.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	High	Potential covariates were selected via causal diagram, literature review, and previous study results (Philips et al. 2018, HERO ID:4728366). Covariates that were adjusted for included maternal age, maternal pre-pregnancy BMI, parity, ethnicity, education, maternal smoking, maternal alcohol, folic acid supplementation, gestational age at time of measurement, and creatinine.
	Metric 10:	Covariate Characterization	High	Gestational age was established during the first ultrasound visit. All other covariates were evaluated via enrollment questionnaire.
	Metric 11:	Co-exposure Counfounding	Medium	Additional exposures that the study evaluated, such as bisphenols, low and high-molecular weight phthalate metabolites, di-2-ethylhexylphthalate metabolites, and di-n-octylphthalate metabolites were all measured and controlled for in analysis of phthalic acid.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cohort study design is appropriate for assessing the association between phthalic acids and common metrics of maternal hemodynamics and gestational hypertensive disorders. Multivariable linear regression was used, as was repeated measurement regression to evaluate relationships between phthalic acid concentrations and repeatedly measured outcomes, both of which are appropriate.
	Metric 13:	Statistical Power	Medium	The number of participants is high at 1233, and cases of gestational hypertension and pre-eclampsia are not exceedingly rare (n=40 and n=24, respectively). Some subgroups - especially for pre-eclampsia - are small, but the overall number of participants is sufficiently large.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is thorough and could be reproduced given the data used.
	Metric 15:	Statistical Analysis	Low	There is no mention of assumptions for multivariable linear regression or repeated measures analysis, and are thus not assessed.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	Phthalic acid is directly measured in urine. However, phthalic acid is a general metabolite of many phthalate esters.
	Metric 17:	Effect Biomarker	High	The introduction cites sources in support of the use of placental growth factor and soluble-tyrosine kinase as associates of impaired vascular proliferation, which may result in increased risk of gestational hypertensive disorders. (Saito and Nakashima, 2014, not available in HERO).
	Metric 18:	Method Sensitivity	Medium	Exposure: A limit of detection is not stated, but the study does clarify that only 0.3% of values were below the limit of detection. More information may be available in cited studies (Silva et al. 2004, HERO ID: 673528; Philips et al. 2018, HERO ID:4728366)Effect: No information is provided regarding limits of detection for effect biomarkers.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
	Metric 19:	Biomarker Stability	Medium	Exposure: Urine samples were collected in 100-ml polypropylene containers, refrigerated, aliquoted and frozen at -20 deg C within 24 hours. Information regarding potential losses due to thawing are not reported.Effect: No information is provided regarding stability of blood samples used for effect biomarkers.
	Metric 20:	Sample Contamination	Medium	No information regarding potential contamination is provided.
	Metric 21:	Method Requirements	High	Exposure: Phthalic acid was assessed using high-performance liquid chromatography-tandem mass spectroscopy analysis.Effect: Effect biomarkers were measured using an immune-electrochemoluminence assay.
	Metric 22:	Matrix Adjustment	Medium	All results, including basic models, are creatinine adjusted.

Additional Comments: This prospective cohort study compares concentrations of phthalic acid in early pregnancy with various pregnancy hypertensive outcomes, including placental growth factor and soluble fms-tyrosine kinase-1 concentrations, diagnosis of gestational hypertension, pre-eclampsia, placental weight, and pregnancy-related hemodynamic functions. The study is overall thorough and high in quality, with strengths including a wide range of exposure levels, key covariate adjustment, and appropriate statistical modeling. One notable weakness is the lack of validation of model assumptions.

Overall Quality Determination High

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Participants were selected from the Generation R prospective cohort of 8879 women. Selection criteria were stated. Exclusion criteria were lack of blood pressure measurements (n= 140) or women with pre-existing hypertension (n=23), resulting in a total sample of 1233 for analysis. Demographic characteristics of those included were generally similar to those excluded.
	Metric 2:	Attrition	High	The only reported missing data were several covariate measures, which were calculated using multiple imputation. No participants are reported as being lost to follow-up.
	Metric 3:	Comparison Group	High	Baseline demographic characteristics between participants such as age, education level, ethnicity, and parity were considered as potential confounders to adjust for any potential differences by exposure level.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Concentration of phthalic acid was measured directly in spot urine samples.
	Metric 5:	Exposure Levels	Medium	The exposure range is sufficiently large, with a median of 56.99 ng/ml and an inter-quartile range of 30.62-124.05. The inter-quartile range is assumed to be 25th-75th percentile, but this is not explicitly stated. (Supplemental Table 1)
	Metric 6:	Temporality	Medium	Outcome measures are reported during early and mid-pregnancy for blood samples, and umbilical and uterine measures were taken in mid and late pregnancy. Since the exposure is measured once during early pregnancy, a clear temporality is established. However, there is no consideration of whether or not the exposure window is appropriate to precede the effect.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Effect biomarkers were measured from blood samples. Placental vascular resistance was evaluated using flow velocity waveforms, and umbilical artery pulsatility index was measured in a free-floating loop of the umbilical cord. Placental weight and diagnosis of gestational hypertensive disorders were obtained from medical records and measured using standard protocols. Biomarker and placental weight were assessed using gold-standard methodology, and there is no evidence that the methods for hemodynamic function had low validity.
	Metric 8:	Reporting Bias	High	Measured outcomes are reported in tables/figures, and discussed in the text. Effect estimates are reported with 95% confidence intervals.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain		Metric	Rating	Comments
	Metric 9:	Covariate Adjustment	High	Potential covariates were selected via causal diagram, literature review, and previous study results (Philips et al. 2018, HERO ID:4728366). Covariates that were adjusted for included maternal age, maternal pre-pregnancy BMI, parity, ethnicity, education, maternal smoking, maternal alcohol, folic acid supplementation, gestational age at time of measurement, and creatinine.
	Metric 10:	Covariate Characterization	High	Gestational age was established during the first ultrasound visit. All other covariates were evaluated via enrollment questionnaire.
	Metric 11:	Co-exposure Counfounding	Medium	Additional exposures that the study evaluated, such as bisphenols, low and high-molecular weight phthalate metabolites, di-2-ethylhexylphthalate metabolites, and di-n-octylphthalate metabolites were all measured and controlled for in analysis of phthalic acid.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cohort study design is appropriate for assessing the association between phthalic acids and common metrics of maternal hemodynamics and gestational hypertensive disorders. Multivariable linear regression was used, as was repeated measurement regression to evaluate relationships between phthalic acid concentrations and repeatedly measured outcomes, both of which are appropriate.
	Metric 13:	Statistical Power	Medium	The number of participants is high at 1233, and cases of gestational hypertension and pre-eclampsia are not exceedingly rare (n=40 and n=24, respectively). Some subgroups - especially for pre-eclampsia - are small, but the overall number of participants is sufficiently large.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is thorough and could be reproduced given the data used.
	Metric 15:	Statistical Analysis	Low	There is no mention of assumptions for multivariable linear regression or repeated measures analysis, and are thus not assessed.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	Phthalic acid is directly measured in urine. However, phthalic acid is a general metabolite of many phthalate esters.
	Metric 17:	Effect Biomarker	High	The introduction cites sources in support of the use of placental growth factor and soluble-tyrosine kinase as associates of impaired vascular proliferation, which may result in increased risk of gestational hypertensive disorders. (Saito and Nakashima, 2014, not available in HERO).
	Metric 18:	Method Sensitivity	Medium	Exposure: A limit of detection is not stated, but the study does clarify that only 0.3% of values were below the limit of detection. More information may be available in cited studies (Silva et al. 2004, HERO ID: 673528; Philips et al. 2018, HERO ID:4728366)Effect: No information is provided regarding limits of detection for effect biomarkers.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
	Metric 19:	Biomarker Stability	Medium	Exposure: Urine samples were collected in 100-ml polypropylene containers, refrigerated, aliquoted and frozen at -20 deg C within 24 hours. Information regarding potential losses due to thawing are not reported.Effect: No information is provided regarding stability of blood samples used for effect biomarkers.
	Metric 20:	Sample Contamination	Medium	No information regarding potential contamination is provided.
	Metric 21:	Method Requirements	High	Exposure: Phthalic acid was assessed using high-performance liquid chromatography-tandem mass spectroscopy analysis.Effect: Effect biomarkers were measured using an immune-electrochemoluminence assay.
	Metric 22:	Matrix Adjustment	Medium	All results, including basic models, are creatinine adjusted.

Additional Comments: This prospective cohort study compares concentrations of phthalic acid in early pregnancy with various pregnancy hypertensive outcomes, including placental growth factor and soluble fms-tyrosine kinase-1 concentrations, diagnosis of gestational hypertension, pre-eclampsia, placental weight, and pregnancy-related hemodynamic functions. The study is overall thorough and high in quality, with strengths including a wide range of exposure levels, key covariate adjustment, and appropriate statistical modeling. One notable weakness is the lack of validation of model assumptions.

Overall Quality Determination High

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Participants were selected from the Generation R prospective cohort of 8879 women. Selection criteria were stated. Exclusion criteria were lack of blood pressure measurements (n= 140) or women with pre-existing hypertension (n=23), resulting in a total sample of 1233 for analysis. Demographic characteristics of those included were generally similar to those excluded.
	Metric 2:	Attrition	High	The only reported missing data were several covariate measures, which were calculated using multiple imputation. No participants are reported as being lost to follow-up.
	Metric 3:	Comparison Group	High	Baseline demographic characteristics between participants such as age, education level, ethnicity, and parity were considered as potential confounders to adjust for any potential differences by exposure level.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Concentration of phthalic acid was measured directly in spot urine samples.
	Metric 5:	Exposure Levels	Medium	The exposure range is sufficiently large, with a median of 56.99 ng/ml and an inter-quartile range of 30.62-124.05. The inter-quartile range is assumed to be 25th-75th percentile, but this is not explicitly stated. (Supplemental Table 1)
	Metric 6:	Temporality	Medium	Outcome measures are reported during early and mid-pregnancy for blood samples, and umbilical and uterine measures were taken in mid and late pregnancy. Since the exposure is measured once during early pregnancy, a clear temporality is established. However, there is no consideration of whether or not the exposure window is appropriate to precede the effect.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Effect biomarkers were measured from blood samples. Placental vascular resistance was evaluated using flow velocity waveforms, and umbilical artery pulsatility index was measured in a free-floating loop of the umbilical cord. Placental weight and diagnosis of gestational hypertensive disorders were obtained from medical records and measured using standard protocols. Biomarker and placental weight were assessed using gold-standard methodology, and there is no evidence that the methods for hemodynamic function had low validity.
	Metric 8:	Reporting Bias	High	Measured outcomes are reported in tables/figures, and discussed in the text. Effect estimates are reported with 95% confidence intervals.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	High	Potential covariates were selected via causal diagram, literature review, and previous study results (Philips et al. 2018, HERO ID:4728366). Covariates that were adjusted for included maternal age, maternal pre-pregnancy BMI, parity, ethnicity, education, maternal smoking, maternal alcohol, folic acid supplementation, gestational age at time of measurement, and creatinine.
	Metric 10:	Covariate Characterization	High	Gestational age was established during the first ultrasound visit. All other covariates were evaluated via enrollment questionnaire.
	Metric 11:	Co-exposure Counfounding	Medium	Additional exposures that the study evaluated, such as bisphenols, low and high-molecular weight phthalate metabolites, di-2-ethylhexylphthalate metabolites, and di-n-octylphthalate metabolites were all measured and controlled for in analysis of phthalic acid.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cohort study design is appropriate for assessing the association between phthalic acids and common metrics of maternal hemodynamics and gestational hypertensive disorders. Multivariable linear regression was used, as was repeated measurement regression to evaluate relationships between phthalic acid concentrations and repeatedly measured outcomes, both of which are appropriate.
	Metric 13:	Statistical Power	Medium	The number of participants is high at 1233, and cases of gestational hypertension and pre-eclampsia are not exceedingly rare (n=40 and n=24, respectively). Some subgroups - especially for pre-eclampsia - are small, but the overall number of participants is sufficiently large.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is thorough and could be reproduced given the data used.
	Metric 15:	Statistical Analysis	Low	There is no mention of assumptions for multivariable linear regression or repeated measures analysis, and are thus not assessed.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	Phthalic acid is directly measured in urine. However, phthalic acid is a general metabolite of many phthalate esters.
	Metric 17:	Effect Biomarker	High	The introduction cites sources in support of the use of placental growth factor and soluble-tyrosine kinase as associates of impaired vascular proliferation, which may result in increased risk of gestational hypertensive disorders. (Saito and Nakashima, 2014, not available in HERO).
	Metric 18:	Method Sensitivity	Medium	Exposure: A limit of detection is not stated, but the study does clarify that only 0.3% of values were below the limit of detection. More information may be available in cited studies (Silva et al. 2004, HERO ID: 673528; Philips et al. 2018, HERO ID:4728366)Effect: No information is provided regarding limits of detection for effect biomarkers.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
	Metric 19:	Biomarker Stability	Medium	Exposure: Urine samples were collected in 100-ml polypropylene containers, refrigerated, aliquoted and frozen at -20 deg C within 24 hours. Information regarding potential losses due to thawing are not reported.Effect: No information is provided regarding stability of blood samples used for effect biomarkers.
	Metric 20:	Sample Contamination	Medium	No information regarding potential contamination is provided.
	Metric 21:	Method Requirements	High	Exposure: Phthalic acid was assessed using high-performance liquid chromatography-tandem mass spectroscopy analysis.Effect: Effect biomarkers were measured using an immune-electrochemoluminence assay.
	Metric 22:	Matrix Adjustment	Medium	All results, including basic models, are creatinine adjusted.

Additional Comments: This prospective cohort study compares concentrations of phthalic acid in early pregnancy with various pregnancy hypertensive outcomes, including placental growth factor and soluble fms-tyrosine kinase-1 concentrations, diagnosis of gestational hypertension, pre-eclampsia, placental weight, and pregnancy-related hemodynamic functions. The study is overall thorough and high in quality, with strengths including a wide range of exposure levels, key covariate adjustment, and appropriate statistical modeling. One notable weakness is the lack of validation of model assumptions.

Overall Quality Determination High

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Participants were selected from the Generation R prospective cohort of 8879 women. Selection criteria were stated. Exclusion criteria were lack of blood pressure measurements (n= 140) or women with pre-existing hypertension (n=23), resulting in a total sample of 1233 for analysis. Demographic characteristics of those included were generally similar to those excluded.
	Metric 2:	Attrition	High	The only reported missing data were several covariate measures, which were calculated using multiple imputation. No participants are reported as being lost to follow-up.
	Metric 3:	Comparison Group	High	Baseline demographic characteristics between participants such as age, education level, ethnicity, and parity were considered as potential confounders to adjust for any potential differences by exposure level.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Concentration of phthalic acid was measured directly in spot urine samples.
	Metric 5:	Exposure Levels	Medium	The exposure range is sufficiently large, with a median of 56.99 ng/ml and an inter-quartile range of 30.62-124.05. The inter-quartile range is assumed to be 25th-75th percentile, but this is not explicitly stated. (Supplemental Table 1)
	Metric 6:	Temporality	Medium	Outcome measures are reported during early and mid-pregnancy for blood samples, and umbilical and uterine measures were taken in mid and late pregnancy. Since the exposure is measured once during early pregnancy, a clear temporality is established. However, there is no consideration of whether or not the exposure window is appropriate to precede the effect.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Effect biomarkers were measured from blood samples. Placental vascular resistance was evaluated using flow velocity waveforms, and umbilical artery pulsatility index was measured in a free-floating loop of the umbilical cord. Placental weight and diagnosis of gestational hypertensive disorders were obtained from medical records and measured using standard protocols. Biomarker and placental weight were assessed using gold-standard methodology, and there is no evidence that the methods for hemodynamic function had low validity.
	Metric 8:	Reporting Bias	High	Measured outcomes are reported in tables/figures, and discussed in the text. Effect estimates are reported with 95% confidence intervals.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	High	Potential covariates were selected via causal diagram, literature review, and previous study results (Philips et al. 2018, HERO ID:4728366). Covariates that were adjusted for included maternal age, maternal pre-pregnancy BMI, parity, ethnicity, education, maternal smoking, maternal alcohol, folic acid supplementation, gestational age at time of measurement, and creatinine.
	Metric 10:	Covariate Characterization	High	Gestational age was established during the first ultrasound visit. All other covariates were evaluated via enrollment questionnaire.
	Metric 11:	Co-exposure Counfounding	Medium	Additional exposures that the study evaluated, such as bisphenols, low and high-molecular weight phthalate metabolites, di-2-ethylhexylphthalate metabolites, and di-n-octylphthalate metabolites were all measured and controlled for in analysis of phthalic acid.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cohort study design is appropriate for assessing the association between phthalic acids and common metrics of maternal hemodynamics and gestational hypertensive disorders. Multivariable linear regression was used, as was repeated measurement regression to evaluate relationships between phthalic acid concentrations and repeatedly measured outcomes, both of which are appropriate.
	Metric 13:	Statistical Power	Medium	The number of participants is high at 1233, and cases of gestational hypertension and pre-eclampsia are not exceedingly rare (n=40 and n=24, respectively). Some subgroups - especially for pre-eclampsia - are small, but the overall number of participants is sufficiently large.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is thorough and could be reproduced given the data used.
	Metric 15:	Statistical Analysis	Low	There is no mention of assumptions for multivariable linear regression or repeated measures analysis, and are thus not assessed.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	Phthalic acid is directly measured in urine. However, phthalic acid is a general metabolite of many phthalate esters.
	Metric 17:	Effect Biomarker	High	The introduction cites sources in support of the use of placental growth factor and soluble-tyrosine kinase as associates of impaired vascular proliferation, which may result in increased risk of gestational hypertensive disorders. (Saito and Nakashima, 2014, not available in HERO).
	Metric 18:	Method Sensitivity	Medium	Exposure: A limit of detection is not stated, but the study does clarify that only 0.3% of values were below the limit of detection. More information may be available in cited studies (Silva et al. 2004, HERO ID: 673528; Philips et al. 2018, HERO ID:4728366)Effect: No information is provided regarding limits of detection for effect biomarkers.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
	Metric 19:	Biomarker Stability	Medium	Exposure: Urine samples were collected in 100-ml polypropylene containers, refrigerated, aliquoted and frozen at -20 deg C within 24 hours. Information regarding potential losses due to thawing are not reported.Effect: No information is provided regarding stability of blood samples used for effect biomarkers.
	Metric 20:	Sample Contamination	Medium	No information regarding potential contamination is provided.
	Metric 21:	Method Requirements	High	Exposure: Phthalic acid was assessed using high-performance liquid chromatography-tandem mass spectroscopy analysis.Effect: Effect biomarkers were measured using an immune-electrochemoluminence assay.
	Metric 22:	Matrix Adjustment	Medium	All results, including basic models, are creatinine adjusted.

Additional Comments: This prospective cohort study compares concentrations of phthalic acid in early pregnancy with various pregnancy hypertensive outcomes, including placental growth factor and soluble fms-tyrosine kinase-1 concentrations, diagnosis of gestational hypertension, pre-eclampsia, placental weight, and pregnancy-related hemodynamic functions. The study is overall thorough and high in quality, with strengths including a wide range of exposure levels, key covariate adjustment, and appropriate statistical modeling. One notable weakness is the lack of validation of model assumptions.

Overall Quality Determination High

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Participants were selected from the Generation R prospective cohort of 8879 women. Selection criteria were stated. Exclusion criteria were lack of blood pressure measurements (n= 140) or women with pre-existing hypertension (n=23), resulting in a total sample of 1233 for analysis. Demographic characteristics of those included were generally similar to those excluded.
	Metric 2:	Attrition	High	The only reported missing data were several covariate measures, which were calculated using multiple imputation. No participants are reported as being lost to follow-up.
	Metric 3:	Comparison Group	High	Baseline demographic characteristics between participants such as age, education level, ethnicity, and parity were considered as potential confounders to adjust for any potential differences by exposure level.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Concentration of phthalic acid was measured directly in spot urine samples.
	Metric 5:	Exposure Levels	Medium	The exposure range is sufficiently large, with a median of 56.99 ng/ml and an inter-quartile range of 30.62-124.05. The inter-quartile range is assumed to be 25th-75th percentile, but this is not explicitly stated. (Supplemental Table 1)
	Metric 6:	Temporality	Medium	Outcome measures are reported during early and mid-pregnancy for blood samples, and umbilical and uterine measures were taken in mid and late pregnancy. Since the exposure is measured once during early pregnancy, a clear temporality is established. However, there is no consideration of whether or not the exposure window is appropriate to precede the effect.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Effect biomarkers were measured from blood samples. Placental vascular resistance was evaluated using flow velocity waveforms, and umbilical artery pulsatility index was measured in a free-floating loop of the umbilical cord. Placental weight and diagnosis of gestational hypertensive disorders were obtained from medical records and measured using standard protocols. Biomarker and placental weight were assessed using gold-standard methodology, and there is no evidence that the methods for hemodynamic function had low validity.
	Metric 8:	Reporting Bias	High	Measured outcomes are reported in tables/figures, and discussed in the text. Effect estimates are reported with 95% confidence intervals.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	High	Potential covariates were selected via causal diagram, literature review, and previous study results (Philips et al. 2018, HERO ID:4728366). Covariates that were adjusted for included maternal age, maternal pre-pregnancy BMI, parity, ethnicity, education, maternal smoking, maternal alcohol, folic acid supplementation, gestational age at time of measurement, and creatinine.
	Metric 10:	Covariate Characterization	High	Gestational age was established during the first ultrasound visit. All other covariates were evaluated via enrollment questionnaire.
	Metric 11:	Co-exposure Counfounding	Medium	Additional exposures that the study evaluated, such as bisphenols, low and high-molecular weight phthalate metabolites, di-2-ethylhexylphthalate metabolites, and di-n-octylphthalate metabolites were all measured and controlled for in analysis of phthalic acid.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cohort study design is appropriate for assessing the association between phthalic acids and common metrics of maternal hemodynamics and gestational hypertensive disorders. Multivariable linear regression was used, as was repeated measurement regression to evaluate relationships between phthalic acid concentrations and repeatedly measured outcomes, both of which are appropriate.
	Metric 13:	Statistical Power	Medium	The number of participants is high at 1233, and cases of gestational hypertension and pre-eclampsia are not exceedingly rare (n=40 and n=24, respectively). Some subgroups - especially for pre-eclampsia - are small, but the overall number of participants is sufficiently large.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is thorough and could be reproduced given the data used.
	Metric 15:	Statistical Analysis	Low	There is no mention of assumptions for multivariable linear regression or repeated measures analysis, and are thus not assessed.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	Phthalic acid is directly measured in urine. However, phthalic acid is a general metabolite of many phthalate esters.
	Metric 17:	Effect Biomarker	High	The introduction cites sources in support of the use of placental growth factor and soluble-tyrosine kinase as associates of impaired vascular proliferation, which may result in increased risk of gestational hypertensive disorders. (Saito and Nakashima, 2014, not available in HERO).
	Metric 18:	Method Sensitivity	Medium	Exposure: A limit of detection is not stated, but the study does clarify that only 0.3% of values were below the limit of detection. More information may be available in cited studies (Silva et al. 2004, HERO ID: 673528; Philips et al. 2018, HERO ID:4728366)Effect: No information is provided regarding limits of detection for effect biomarkers.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
	Metric 19:	Biomarker Stability	Medium	Exposure: Urine samples were collected in 100-ml polypropylene containers, refrigerated, aliquoted and frozen at -20 deg C within 24 hours. Information regarding potential losses due to thawing are not reported.Effect: No information is provided regarding stability of blood samples used for effect biomarkers.
	Metric 20:	Sample Contamination	Medium	No information regarding potential contamination is provided.
	Metric 21:	Method Requirements	High	Exposure: Phthalic acid was assessed using high-performance liquid chromatography-tandem mass spectroscopy analysis.Effect: Effect biomarkers were measured using an immune-electrochemoluminence assay.
	Metric 22:	Matrix Adjustment	Medium	All results, including basic models, are creatinine adjusted.

Additional Comments: This prospective cohort study compares concentrations of phthalic acid in early pregnancy with various pregnancy hypertensive outcomes, including placental growth factor and soluble fms-tyrosine kinase-1 concentrations, diagnosis of gestational hypertension, pre-eclampsia, placental weight, and pregnancy-related hemodynamic functions. The study is overall thorough and high in quality, with strengths including a wide range of exposure levels, key covariate adjustment, and appropriate statistical modeling. One notable weakness is the lack of validation of model assumptions.

Overall Quality Determination High

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Participants were selected from the Generation R prospective cohort of 8879 women. Selection criteria were stated. Exclusion criteria were lack of blood pressure measurements (n= 140) or women with pre-existing hypertension (n=23), resulting in a total sample of 1233 for analysis. Demographic characteristics of those included were generally similar to those excluded.
	Metric 2:	Attrition	High	The only reported missing data were several covariate measures, which were calculated using multiple imputation. No participants are reported as being lost to follow-up.
	Metric 3:	Comparison Group	High	Baseline demographic characteristics between participants such as age, education level, ethnicity, and parity were considered as potential confounders to adjust for any potential differences by exposure level.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Concentration of phthalic acid was measured directly in spot urine samples.
	Metric 5:	Exposure Levels	Medium	The exposure range is sufficiently large, with a median of 56.99 ng/ml and an inter-quartile range of 30.62-124.05. The inter-quartile range is assumed to be 25th-75th percentile, but this is not explicitly stated. (Supplemental Table 1)
	Metric 6:	Temporality	Medium	Outcome measures are reported during early and mid-pregnancy for blood samples, and umbilical and uterine measures were taken in mid and late pregnancy. Since the exposure is measured once during early pregnancy, a clear temporality is established. However, there is no consideration of whether or not the exposure window is appropriate to precede the effect.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Effect biomarkers were measured from blood samples. Placental vascular resistance was evaluated using flow velocity waveforms, and umbilical artery pulsatility index was measured in a free-floating loop of the umbilical cord. Placental weight and diagnosis of gestational hypertensive disorders were obtained from medical records and measured using standard protocols. Biomarker and placental weight were assessed using gold-standard methodology, and there is no evidence that the methods for hemodynamic function had low validity.
	Metric 8:	Reporting Bias	High	Measured outcomes are reported in tables/figures, and discussed in the text. Effect estimates are reported with 95% confidence intervals.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	High	Potential covariates were selected via causal diagram, literature review, and previous study results (Philips et al. 2018, HERO ID:4728366). Covariates that were adjusted for included maternal age, maternal pre-pregnancy BMI, parity, ethnicity, education, maternal smoking, maternal alcohol, folic acid supplementation, gestational age at time of measurement, and creatinine.
	Metric 10:	Covariate Characterization	High	Gestational age was established during the first ultrasound visit. All other covariates were evaluated via enrollment questionnaire.
	Metric 11:	Co-exposure Counfounding	Medium	Additional exposures that the study evaluated, such as bisphenols, low and high-molecular weight phthalate metabolites, di-2-ethylhexylphthalate metabolites, and di-n-octylphthalate metabolites were all measured and controlled for in analysis of phthalic acid.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cohort study design is appropriate for assessing the association between phthalic acids and common metrics of maternal hemodynamics and gestational hypertensive disorders. Multivariable linear regression was used, as was repeated measurement regression to evaluate relationships between phthalic acid concentrations and repeatedly measured outcomes, both of which are appropriate.
	Metric 13:	Statistical Power	Medium	The number of participants is high at 1233, and cases of gestational hypertension and pre-eclampsia are not exceedingly rare (n=40 and n=24, respectively). Some subgroups - especially for pre-eclampsia - are small, but the overall number of participants is sufficiently large.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is thorough and could be reproduced given the data used.
	Metric 15:	Statistical Analysis	Low	There is no mention of assumptions for multivariable linear regression or repeated measures analysis, and are thus not assessed.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	Phthalic acid is directly measured in urine. However, phthalic acid is a general metabolite of many phthalate esters.
	Metric 17:	Effect Biomarker	High	The introduction cites sources in support of the use of placental growth factor and soluble-tyrosine kinase as associates of impaired vascular proliferation, which may result in increased risk of gestational hypertensive disorders. (Saito and Nakashima, 2014, not available in HERO).
	Metric 18:	Method Sensitivity	Medium	Exposure: A limit of detection is not stated, but the study does clarify that only 0.3% of values were below the limit of detection. More information may be available in cited studies (Silva et al. 2004, HERO ID: 673528; Philips et al. 2018, HERO ID:4728366)Effect: No information is provided regarding limits of detection for effect biomarkers.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
	Metric 19:	Biomarker Stability	Medium	Exposure: Urine samples were collected in 100-ml polypropylene containers, refrigerated, aliquoted and frozen at -20 deg C within 24 hours. Information regarding potential losses due to thawing are not reported.Effect: No information is provided regarding stability of blood samples used for effect biomarkers.
	Metric 20:	Sample Contamination	Medium	No information regarding potential contamination is provided.
	Metric 21:	Method Requirements	High	Exposure: Phthalic acid was assessed using high-performance liquid chromatography-tandem mass spectroscopy analysis.Effect: Effect biomarkers were measured using an immune-electrochemoluminence assay.
	Metric 22:	Matrix Adjustment	Medium	All results, including basic models, are creatinine adjusted.

Additional Comments: This prospective cohort study compares concentrations of phthalic acid in early pregnancy with various pregnancy hypertensive outcomes, including placental growth factor and soluble fms-tyrosine kinase-1 concentrations, diagnosis of gestational hypertension, pre-eclampsia, placental weight, and pregnancy-related hemodynamic functions. The study is overall thorough and high in quality, with strengths including a wide range of exposure levels, key covariate adjustment, and appropriate statistical modeling. One notable weakness is the lack of validation of model assumptions.

Overall Quality Determination High

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	Participants were selected from the Generation R prospective cohort of 8879 women. Selection criteria were stated. Exclusion criteria were lack of blood pressure measurements (n= 140) or women with pre-existing hypertension (n=23), resulting in a total sample of 1233 for analysis. Demographic characteristics of those included were generally similar to those excluded.
	Metric 2:	Attrition	High	The only reported missing data were several covariate measures, which were calculated using multiple imputation. No participants are reported as being lost to follow-up.
	Metric 3:	Comparison Group	High	Baseline demographic characteristics between participants such as age, education level, ethnicity, and parity were considered as potential confounders to adjust for any potential differences by exposure level.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Concentration of phthalic acid was measured directly in spot urine samples.
	Metric 5:	Exposure Levels	Medium	The exposure range is sufficiently large, with a median of 56.99 ng/ml and an inter-quartile range of 30.62-124.05. The inter-quartile range is assumed to be 25th-75th percentile, but this is not explicitly stated. (Supplemental Table 1)
	Metric 6:	Temporality	Medium	Outcome measures are reported during early and mid-pregnancy for blood samples, and umbilical and uterine measures were taken in mid and late pregnancy. Since the exposure is measured once during early pregnancy, a clear temporality is established. However, there is no consideration of whether or not the exposure window is appropriate to precede the effect.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Effect biomarkers were measured from blood samples. Placental vascular resistance was evaluated using flow velocity waveforms, and umbilical artery pulsatility index was measured in a free-floating loop of the umbilical cord. Placental weight and diagnosis of gestational hypertensive disorders were obtained from medical records and measured using standard protocols. Biomarker and placental weight were assessed using gold-standard methodology, and there is no evidence that the methods for hemodynamic function had low validity.
	Metric 8:	Reporting Bias	High	Measured outcomes are reported in tables/figures, and discussed in the text. Effect estimates are reported with 95% confidence intervals.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	High	Potential covariates were selected via causal diagram, literature review, and previous study results (Philips et al. 2018, HERO ID:4728366). Covariates that were adjusted for included maternal age, maternal pre-pregnancy BMI, parity, ethnicity, education, maternal smoking, maternal alcohol, folic acid supplementation, gestational age at time of measurement, and creatinine.
	Metric 10:	Covariate Characterization	High	Gestational age was established during the first ultrasound visit. All other covariates were evaluated via enrollment questionnaire.
	Metric 11:	Co-exposure Counfounding	Medium	Additional exposures that the study evaluated, such as bisphenols, low and high-molecular weight phthalate metabolites, di-2-ethylhexylphthalate metabolites, and di-n-octylphthalate metabolites were all measured and controlled for in analysis of phthalic acid.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cohort study design is appropriate for assessing the association between phthalic acids and common metrics of maternal hemodynamics and gestational hypertensive disorders. Multivariable linear regression was used, as was repeated measurement regression to evaluate relationships between phthalic acid concentrations and repeatedly measured outcomes, both of which are appropriate.
	Metric 13:	Statistical Power	Medium	The number of participants is high at 1233, and cases of gestational hypertension and pre-eclampsia are not exceedingly rare (n=40 and n=24, respectively). Some subgroups - especially for pre-eclampsia - are small, but the overall number of participants is sufficiently large.
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is thorough and could be reproduced given the data used.
	Metric 15:	Statistical Analysis	Low	There is no mention of assumptions for multivariable linear regression or repeated measures analysis, and are thus not assessed.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	Phthalic acid is directly measured in urine. However, phthalic acid is a general metabolite of many phthalate esters.
	Metric 17:	Effect Biomarker	High	The introduction cites sources in support of the use of placental growth factor and soluble-tyrosine kinase as associates of impaired vascular proliferation, which may result in increased risk of gestational hypertensive disorders. (Saito and Nakashima, 2014, not available in HERO).
	Metric 18:	Method Sensitivity	Medium	Exposure: A limit of detection is not stated, but the study does clarify that only 0.3% of values were below the limit of detection. More information may be available in cited studies (Silva et al. 2004, HERO ID: 673528; Philips et al. 2018, HERO ID:4728366)Effect: No information is provided regarding limits of detection for effect biomarkers.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Philips, E. M., Trasande, L., Kahn, L. G., Gaillard, R., Steegers, P., E.A., Jaddoe, V., V.W. (2019). Early pregnancy bisphenol and phthalate metabolite levels, maternal hemodynamics and gestational hypertensive disorders. Human Reproduction 34(2):365-373.
Health Outcome(s) Assessed:	Reproductive/Developmental
Reported Health Effect(s):	placental growth factor concentration (PIGF), soluble fms-like tyrosine kinase concentration (SFlt-1), SFlt-1: PIGF ratio, umbilical artery pulsatility index, uterine artery resistance index, uterine artery notching, placental weight, gestational hypertension, pre-eclampsia, blood pressure
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5043413

Domain	Metric		Rating	Comments
	Metric 19:	Biomarker Stability	Medium	Exposure: Urine samples were collected in 100-ml polypropylene containers, refrigerated, aliquoted and frozen at -20 deg C within 24 hours. Information regarding potential losses due to thawing are not reported.Effect: No information is provided regarding stability of blood samples used for effect biomarkers.
	Metric 20:	Sample Contamination	Medium	No information regarding potential contamination is provided.
	Metric 21:	Method Requirements	High	Exposure: Phthalic acid was assessed using high-performance liquid chromatography-tandem mass spectroscopy analysis.Effect: Effect biomarkers were measured using an immune-electrochemoluminence assay.
	Metric 22:	Matrix Adjustment	Medium	All results, including basic models, are creatinine adjusted.

Additional Comments: This prospective cohort study compares concentrations of phthalic acid in early pregnancy with various pregnancy hypertensive outcomes, including placental growth factor and soluble fms-tyrosine kinase-1 concentrations, diagnosis of gestational hypertension, pre-eclampsia, placental weight, and pregnancy-related hemodynamic functions. The study is overall thorough and high in quality, with strengths including a wide range of exposure levels, key covariate adjustment, and appropriate statistical modeling. One notable weakness is the lack of validation of model assumptions.

Overall Quality Determination High

Study Citation:	Choi, J., Eom, J., Kim, J., Lee, S., Kim, Y. (2014). Association between some endocrine-disrupting chemicals and childhood obesity in biological samples of young girls: A cross-sectional study. Environmental Toxicology and Pharmacology 38(1):51-57.
Health Outcome(s) Assessed:	Nutrition & Metabolic
Reported Health Effect(s):	Obesity
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	2510764

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	This cross-sectional study examines obesity in Korean girls ages 6-14. Controls (n=58) and cases (n=69) were enrolled in the study from a medical college in Seoul. While the BMI cutoff (85%) for obesity is provided, little information is provided relating to participant inclusion/exclusion, participation rates, or how study participants relate to the eligible study population. This makes it difficult to assess representativeness.
	Metric 2:	Attrition	High	Cases and controls were recruited from the clinical setting to participate in the study only if they had complete information for inclusion in analyses. As a result, attrition is not of concern in this study.
	Metric 3:	Comparison Group	Medium	Both cases and controls were selected from the same clinical setting in Seoul, South Korea. While the age range for controls and cases slightly differed (6-12 years for controls, 6-14 for cases), there is little additional information provided that allows for comparison between the two groups.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	High	Phthalic acid was measured in serum and urine samples. Samples were collected in the same time frame for both cases and controls and were analyzed via gas chromatography-mass spectrometry (GC/MS). GC/MS is a validated approach to measuring exposures in biomonitoring samples including serum and urine.
	Metric 5:	Exposure Levels	Low	The minimum, maximum, and mean values across case and control groups are reported, and exposure is evaluated as a linear variable. Exposure is reported in both cases and controls.
	Metric 6:	Temporality	Medium	Exposure assessment and outcome assessment were performed simultaneously. However, it is unclear whether this results in exposure falling within a relevant time period for the outcome of interest.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Obesity was determined in the clinical setting following height and weight measurement by nurses. Stadiometers were used to measure height, and electronic digital scales used to measure weight. Weight and height were used to calculate BMI for all subjects. The study reports ”childhood obesity and overweight were defined as the 85th and 95th percentiles and above the 95th percentile of the BMI based on 2007 child and adolescent physical growth standards.”

Continued on next page . . .

. . . continued from previous page

Study Citation:	Choi, J., Eom, J., Kim, J., Lee, S., Kim, Y. (2014). Association between some endocrine-disrupting chemicals and childhood obesity in biological samples of young girls: A cross-sectional study. Environmental Toxicology and Pharmacology 38(1):51-57.
Health Outcome(s) Assessed:	Nutrition & Metabolic
Reported Health Effect(s):	Obesity
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	2510764

Domain	Metric		Rating	Comments
	Metric 8:	Reporting Bias	Low	The methods section describes logistic regression in the statistical analysis section. However, only exposure distributions and p-value results of t-tests for cases and controls are reported. The results of the regression are not reported in the paper.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Low	While logistic regression is mentioned, there is no discussion of consideration of potential confounders, adjustment for those confounders, and the distribution of potential confounders was not reported for the study population (cases and controls).
	Metric 10:	Covariate Characterization	N/A	Covariates do not appear to be assessed in the analyses presented in the study.
	Metric 11:	Co-exposure Counfounding	Medium	In this cross-sectional study, other phthalates and BPA were quantified for cases and controls. There is no evidence of other co-exposures that were not accounted for but influenced the analysis.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The cross-sectional study design is acceptable for the assessment of PA exposure and prevalent obesity. Additionally, the presentation of descriptive statistics is appropriate for a cross-sectional study. While logistic regression is outlined in the methods, the results do not appear to be reported.
	Metric 13:	Statistical Power	Medium	The number of cases (n=69) and controls (n=58) is adequate to assess an association between exposure and outcome.
	Metric 14:	Reproducibility of Analyses	Low	As outlined previously, the information around logistic regression analyses is too limited to be reproducible. Variables included in models are not discussed, nor are cut-points for exposure categorization.
	Metric 15:	Statistical Analysis	High	While the methods text presents logistic regression, there is no presentation of the results of the analysis. As a result, only descriptive statistics are presented by case group. However, the set-up of the t-test for differences in the obesity and control groups is sufficiently described and appropriate for the analysis.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	Medium	Phthalic acid was measured in urine and in serum. Phthalic acid can be a metabolite of other phthalates, and thus the biomarker may be derived from multiple parent chemicals.
	Metric 17:	Effect Biomarker	N/A	BMI is used to determine obesity, thus a biomarker was not used to measure the effect.
	Metric 18:	Method Sensitivity	Medium	LOD ranges are provided for chemicals in serum and urine. Study subjects were included in the analysis if they had complete exposure information, thus it can be assumed that subjects did not have exposure levels below the LOD.
	Metric 19:	Biomarker Stability	High	Samples were stored in sealed glass containers and were stored at 20 degrees Celsius. There is no evidence of sample losses during storage.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Choi, J., Eom, J., Kim, J., Lee, S., Kim, Y. (2014). Association between some endocrine-disrupting chemicals and childhood obesity in biological samples of young girls: A cross-sectional study. Environmental Toxicology and Pharmacology 38(1):51-57.
Health Outcome(s) Assessed:	Nutrition & Metabolic
Reported Health Effect(s):	Obesity
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	2510764

Domain	Metric	Rating	Comments
	Metric 20: Sample Contamination	Medium	There is no direct statement that samples are contamination free. Samples were held in glass containers, target compounds were of appropriate grade.
	Metric 21: Method Requirements	Medium	Samples were analyzed using gas chromatography mass spectrometry (GC-MS).
	Metric 22: Matrix Adjustment	Low	The study only provides comparison of mean levels with no discussion of matrix adjustment for urinary PA levels.

Additional Comments: This cross-sectional study examines the association between obesity among young Korean girls and urinary and serum phthalic acid levels. Mean urinary and serum levels were significantly higher among obese study participants than non-obese participants. Logistic regression is discussed in the methods section but results from the analysis are not reported or discussed. Due to the nature of comparison of means, confounders were not considered in the analysis. Additionally, minimal details about sample collection make it difficult to assess the temporal relationship between exposure measurement and outcome measurement.

Overall Quality Determination Medium

Study Citation:	Sol, C. M., Santos, S., Duijts, L., Asimakopoulos, A. G., Martinez-Moral, M. P., Kannan, K., Jaddoe, V., V.W., Trasande, L. (2020). Fetal phthalates and bisphenols and childhood lipid and glucose metabolism: A population-based prospective cohort study. Environment International 144:106063.
Health Outcome(s) Assessed:	Nutrition & Metabolic
Reported Health Effect(s):	Insulin, glucose
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	6957607

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	The study population in this prospective cohort study was a subgroup of mother/child pairs with phthalate and bisphenol concentrations from the Generation R Study in Rotterdam, the Netherlands (n=1,405) from February 2004 and July 2005. Mothers were excluded if they did not have information on phthalate and bisphenol urine concentrations for at least a single time point in pregnancy (n=26). Mother-child pairs were also excluded if singleton children did not participate in postnatal studies at the age of 10 years (n=622). All key information of the study design is reported and described at all steps of the study, and the Generation R cohort is well described in the epidemiological literature.
	Metric 2:	Attrition	Medium	622 out of 1,379 children who were part of the original subsample of mothers with bisphenol and phthalate concentrations were lost to follow-up at 10 years. No reasons for loss to follow-up are reported. These children were thus excluded from further analyses. While attrition may be a concern if the exposure-outcome relationship is different among those lost to follow-up, there is no specific evidence that this is the case. In outcome-specific analyses up to 4 mother-child pairs had missing outcome data, but this is non-significant relative to the large analysis sample of 750+ individuals. Number of individuals are reported at all stages of the study.
	Metric 3:	Comparison Group	High	Key elements of the study design are reported (setting, inclusion and exclusion criteria, and methods of participant selection), and indicate that subjects were similar (recruited from the same eligible population with the same method of ascertainment and within the same time frame using the same inclusion and exclusion criteria, and were of similar age. Maternal and child characteristics were reported. Differences in characteristics of groups were considered as covariates and controlled by statistical analysis. Comparisons between participants and non-participants were also described in the available supplemental document.
Domain 2: Exposure Characterization				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Sol, C. M., Santos, S., Duijts, L., Asimakopoulos, A. G., Martinez-Moral, M. P., Kannan, K., Jaddoe, V., V.W., Trasande, L. (2020). Fetal phthalates and bisphenols and childhood lipid and glucose metabolism: A population-based prospective cohort study. Environment International 144:106063.
Health Outcome(s) Assessed:	Nutrition & Metabolic
Reported Health Effect(s):	Insulin, glucose
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	6957607

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	High	Exposure was consistently assessed using well-established methods (spot urine sample) that measured the chemical concentration of phthalic acid in urine. Exposure was measured in spot urine samples three times during pregnancy, roughly corresponding to the first, second, and third trimesters. Statistical analyses reported both trimester-specific results, and results where exposure was summed across all three time points and divided by three to account for the impact of temporal variability. Analysis was performed at the same lab for all samples using high performance liquid chromatography electrospray ionization-tandem mass spectrometry (HPLC-ESI-MS/MS) as described in Philips et al. 2018 (HERO ID: 4728366). Phthalic acid was "analyzed separately as a proxy for total phthalate exposure". Measured "urine biomarkers for exposure to phthalate metabolites". Concentrations below the LOD were divided by the square root of 2. Concentrations were converted to umol/g creatinine.
	Metric 5:	Exposure Levels	Medium	Urine concentrations were reported for the 1st, 2nd, and 3rd trimesters in the available supplemental document as median, 25th and 75th percentiles. In statistical analyses, exposure was analyzed as a continuous variable as the change in outcome per IQR increase in exposure.
	Metric 6:	Temporality	High	The study presents an appropriate temporality between maternal exposure and outcomes in 10-year old children of mothers. The interval between the exposure and the outcome is appropriate for assessing outcomes from fetal exposure.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Outcomes were assessed using well-established methods; Total cholesterol, HDL cholesterol, triglycerides, glucose, and insulin concentrations from non-fasting venous blood samples were analyzed using Cobas 8000 analyzer (c702 module) for hepatic outcomes and electrochemiluminescence immunoassay (ECLIA) on the E411 module for metabolic outcomes. All outcomes were analyzed at 10 years of age. There are some potential concerns for the use of a non-fasting blood sample, but this is unlikely to be insensitive in a way that is differential by exposure status.
	Metric 8:	Reporting Bias	High	All outcomes were reported. Association of maternal concentrations with outcomes are reported for the 1st, 2nd, and 3rd trimester, and split by sex of children, with confidence interval and/or standard deviation.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	Appropriate adjustments or explicit considerations were made for potential confounders (maternal age, ethnicity, pre-pregnancy BMI, folic acid supplementation, education level, parity, smoking habits, alcohol consumption, maternal diet, childhood BMI, child sex) in the final analyses through the use of statistical models to reduce research-specific bias. Creatinine concentration adjustments were also made.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Sol, C. M., Santos, S., Duijts, L., Asimakopoulos, A. G., Martinez-Moral, M. P., Kannan, K., Jaddoe, V., V.W., Trasande, L. (2020). Fetal phthalates and bisphenols and childhood lipid and glucose metabolism: A population-based prospective cohort study. Environment International 144:106063.
Health Outcome(s) Assessed:	Nutrition & Metabolic
Reported Health Effect(s):	Insulin, glucose
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	6957607

Domain	Metric		Rating	Comments
	Metric 10:	Covariate Characterization	Medium	Covariates were assessed using questionnaires during pregnancy and midwife and hospital records at birth. Answers to questionnaire appear to be self-reported and it was not specified if the method was validated. Maternal diet was assessed using a "previously developed food-based diet quality score" based on national dietary guidelines.
	Metric 11:	Co-exposure Counfounding	Medium	Co-exposure to phthalates and bisphenols examined in this study were considered and correlation coefficients between all measured exposures were calculated.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The Study design was appropriate and statistical analysis were appropriate. The study uses appropriate statistical methods to address the research question.
	Metric 13:	Statistical Power	Medium	The number of subjects (n=757) is likely large enough to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description is sufficient to understand what was done and be conceptually reproducible.
	Metric 15:	Statistical Analysis	High	The models calculating association coefficients are transparent and is stated which variables were included. Model assumptions were adequately described. Basic models are presented in the available supplemental document. Exposure concentrations were log-transformed to address normakity.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	Phthalic acid was directly measured in biological media.
	Metric 17:	Effect Biomarker	High	Insulin, glucose, total cholesterol, HDL cholesterol, LDL cholesterol, and triglycerides were all directly measured in biological media.
	Metric 18:	Method Sensitivity	Medium	The limit of detection was specified to be 6.68 nmol/L for phthalic acid, which is likely low enough to detect chemicals in a sufficient percentage of the samples. Limit of detections are not expected to be a concern for the biomarkers of effect.
	Metric 19:	Biomarker Stability	Medium	Urine samples were collected in "polypropylene urine collection containers, stored at 4 °C and transported within 24 h of receipt to the STAR-MDC laboratory before being distributed manually in 25-ml polypropylene vials to be frozen at − 20 °C". It is noted that phthalates have a short biological half-life of less than 24 hours. Philips et al. 2018 (HERO ID: 4728366) noted that the samples had been stored at -20 degrees Celsius for 10 years and suggested that biological activity during the storage period could not be ruled out.
	Metric 20:	Sample Contamination	High	The methodology was reported in the cited Philips et al. 2018 (HERO ID: 4728366) study. Contamination was reported to be monitored by the analysis of procedural blanks for exposure samples. Outcome biomarkers are not mentioned with regards to contamination.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Sol, C. M., Santos, S., Duijts, L., Asimakopoulos, A. G., Martinez-Moral, M. P., Kannan, K., Jaddoe, V., V.W., Trasande, L. (2020). Fetal phthalates and bisphenols and childhood lipid and glucose metabolism: A population-based prospective cohort study. Environment International 144:106063.
Health Outcome(s) Assessed:	Nutrition & Metabolic
Reported Health Effect(s):	Insulin, glucose
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	6957607

Domain	Metric		Rating	Comments
	Metric 21:	Method Requirements	High	High performance liquid chromatography electrospray ionization-tandem mass spectrometry (HPLC-ESI-MS/MS) was used for the detection and measurement of phthalate metabolites, as reported in Philips et al. 2018 (HERO ID: 4728366). Outcome biomarkers were measured on the COBAS 8000 analyzer using the c702 module or using an electrochemiluminescence immunoassay (ELCIA).
	Metric 22:	Matrix Adjustment	Medium	For the biomarker under consideration, study provides results in the main publication and in an available supplemental document for adjusted matrix concentrations (creatinine-adjusted) and reasons are given for adjustment approach.

Additional Comments: Associations between maternal urinary phthalate concentrations (in 1st, 2nd, and 3rd trimesters) and developmental metabolic parameters (serum lipids, glucose, and insulin concentrations) of children (10 years of age) were assessed in 757 mother/child pairs from a subgroup of the Generation R Study in Rotterdam, the Netherlands (2004-2005). There are no overall concerns for bias, due to the use of a prospective cohort design with extensive details on selection, exposure assessment, and outcome assessment. There was an association between maternal 3rd trimester urinary phthalic acid concentrations and increased triglycerides concentration in boys (0.20; 95% CI: 0.07-0.34). No other associations between maternal phthalic acid concentrations and developmental metabolic adaptations were reported.

Overall Quality Determination High

Study Citation:	Song, Y., Hauser, R., Hu, F. B., Franke, A. A., Liu, S., Sun, Q. (2014). Urinary concentrations of bisphenol A and phthalate metabolites and weight change: A prospective investigation in US women. International Journal of Obesity 38(12):1532-1537.
Health Outcome(s) Assessed:	Nutrition & Metabolic
Reported Health Effect(s):	Bodyweight change, BMI
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	2345937

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	Participants in this study were drawn from the Nurses' Health Study (NHS) and NHSII. The NHS and NHSII were established in 1976 and 1989, respectively, and took place in the United States. The NHS consisted of 121,700 registered nurses aged 30-55, while NHSII consisted of 18,717 registered nurses aged 32-52 years. A nested case-control study was performed among NHS and NHSII participants during the follow-up period of 2000-2008 (NHS) or 1995-2007 (NHSII), to assess type 2 diabetes. Participants in the present analysis were 977 controls from both cohorts who had a urine sample during the first-morning-void urine collection from 1996-2002.Methods of participant selection were described and participation rates were provided. Inclusion and exclusion criteria were not defined. Recruitment details and counts are specified for both parent cohorts. Overall, there are no specific concerns for selection bias.
	Metric 2:	Attrition	Medium	The study noted that urinary phthalic acid concentrations were not available for 144 of the case-control pairs from the NHS group due to "technical reasons." It is not explained what those reasons, were, and 144 represents a somewhat significant portion of the original 977, and the study does not explain if these participants were removed from the analysis. However, there is no evidence that attrition was differential by exposure or outcome.
	Metric 3:	Comparison Group	Medium	In this study, quartiles of phthalic acid were assessed in relation to linear changes in BMI, thus four separate comparison groups were created. All participants were nurses without type 2 diabetes. All relevant baseline characteristics (age, parent cohort, menopausal status, smoking) were considered as potential covariates, indicating the comparison groups created were appropriate.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	Liquid chromatography-mass spectrometry was used to measure urinary concentrations of phthalic acid for all participants. Urine samples were collected in polypropylene containers and were immediately processed upon arrival to the laboratory. Urine samples were collected from the participants in the current study from 1996-2002. While no information on LOD is provided, it appears that the study consistently assessed exposure for all participants using a high-quality method.
	Metric 5:	Exposure Levels	Medium	Exposure groups were separated into 4 quartiles, where the median concentrations for each were provided for phthalic acid.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Song, Y., Hauser, R., Hu, F. B., Franke, A. A., Liu, S., Sun, Q. (2014). Urinary concentrations of bisphenol A and phthalate metabolites and weight change: A prospective investigation in US women. International Journal of Obesity 38(12):1532-1537.
Health Outcome(s) Assessed:	Nutrition & Metabolic
Reported Health Effect(s):	Bodyweight change, BMI
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	2345937

Domain	Metric		Rating	Comments
	Metric 6:	Temporality	Medium	Bodyweights were self-reported every 2 years, and the most recent bodyweight information was from the 10-year follow-up. Baseline BMI was also measured simultaneously with exposure. Temporality is established, but it is unclear whether exposures fall within relevant exposure windows for the outcome of interest.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	High	Bodyweights were self-reported and BMI was calculated as bodyweight divided by the squared height of each participant. A correlation coefficient was calculated to examine the accuracy of self-reported bodyweights compared with measured bodyweights in 184 NHS participants; the correlation coefficient was 0.96, indicating a high level of accuracy.
	Metric 8:	Reporting Bias	Medium	Least square means (95% CI) were provided for all indicated outcomes. The number of participants in each quartile were not provided in Tables 2 and 3. Table 1 does report the number of participants in each quartile for total phthalates, but it was noted that phthalic acid was not included in the calculation of total phthalate metabolite concentrations.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Medium	Appropriate adjustments were made for potential confounders (urinary creatinine concentration, cohort origin, age, menopausal status, smoking status, alcohol consumption, physical activity,alternative healthy eating index score, and total energy intake). However, no framework or justification is provided explaining why the specific covariates were chosen.
	Metric 10:	Covariate Characterization	Medium	Covariate information was collected using self-administered questionnaires and validated food frequency questionnaires. Validation of the food frequency questionnaire was not further described.
	Metric 11:	Co-exposure Counfounding	Medium	There was direct evidence of co-exposures to other pollutants (other phthalate metabolites and BPA), which are presented as separate analyses but not adjusted for in the univariate analysis of phthalic acid.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The study design (prospective cohort) was appropriate for the research question (examining bodyweight changes over time, since urine collection). The statistical methods used were appropriate (linear regression).
	Metric 13:	Statistical Power	Medium	Statistical power calculations were not provided for the analyses reported in Table 2 and 3. Although, the number of participants in each quartile were likely adequate to detect an effect in the exposed population (n=>244 in each quartile for total phthalates; 144 total samples were missing for phthalic acid).

Continued on next page . . .

. . . continued from previous page

Study Citation:	Song, Y., Hauser, R., Hu, F. B., Franke, A. A., Liu, S., Sun, Q. (2014). Urinary concentrations of bisphenol A and phthalate metabolites and weight change: A prospective investigation in US women. International Journal of Obesity 38(12):1532-1537.
Health Outcome(s) Assessed:	Nutrition & Metabolic
Reported Health Effect(s):	Bodyweight change, BMI
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	2345937

Domain	Metric		Rating	Comments
	Metric 14:	Reproducibility of Analyses	Low	The description of the statistical analysis was sufficient; however, it is not clear how the missing concentrations for phthalic acid were handled in the analyses. It is likely that these participants were removed from the analysis, but this is not explicitly stated.
	Metric 15:	Statistical Analysis	Low	The statistical model (linear regression) was sufficient for this analysis. Phthalic acid concentrations were stated to be log-transformed to assess non-normality, and the approach for handling missing covariate data was outlined as imputing the median values of the study population.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	Phthalic acid was directly measured.
	Metric 17:	Effect Biomarker	N/A	Not applicable, no biomarkers of exposure were measured.
	Metric 18:	Method Sensitivity	Low	LOD/LOQ (value or %) are not stated. The number of detects in each quartile were not provided.
	Metric 19:	Biomarker Stability	Medium	Polypropylene containers without preservatives were used to collect the urine samples. Samples were "returned to a central biorepository via overnight courier with an icepack and were immediately processed on arrival and aliquoted into polypropylene cryovials, which were stored in the vapor phase of liquid nitrogen freezers at −130 °C". The actual time between collection and analysis was not specified and stability in the urine at these temperatures/storage conditions was not addressed.
	Metric 20:	Sample Contamination	Medium	Samples were sent overnight and immediately processed in the lab. Environmental contamination of the samples was examined in a pilot study, and it was noted that contamination was possible from the sample containers or during processing. The interclass correlation coefficient for phthalic acid was 0.82.
	Metric 21:	Method Requirements	High	Samples were analyzed using liquid chromatography-mass spectrometry.
	Metric 22:	Matrix Adjustment	Medium	Both models used in the analyses were adjusted for urinary creatinine concentrations. Unadjusted results were not provided.

Additional Comments: This prospective cohort study used data from the Nurses' Health Study (NHS) and NHSII to assess the relationship between a spot urine concentration of phthalic acid and baseline BMI as well as weight change rate over time. The study has multiple strengths, with no concerns for participant selection or outcome assessment. While the direct measurement of phthalic acid is reliable, there are some concerns related to temporality as it is unclear whether or not the exposure falls within a relevant time frame to assess weight change, especially over a 10-year time period. LOD information is also missing, which further limits confidence in the exposure assessment. However, these limitations are unlikely to have a significant enough effect on the findings of the study to downgrade from medium confidence. The study reported that higher phthalic acid concentrations were associated with a faster prospective weight gain, with a potential dose-response pattern.

Overall Quality Determination — Medium

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Nutrition & Metabolic
Reported Health Effect(s):	Serum CO2, serum triglycerides, serum methemoglobulin, serum cholesterol, serum glucose, urine pnehols
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	In this cross-sectional occupational surveillance study, workers from 9 Koppers Coal Tar plants were examined from January-October 1979. Only one of these plants (the Chicago plant) produced phthalic anhydride. Participation at the relevant phthalic anhydride plant (i.e., the Chicago plant) was approximately 46% (105/230), while participation from the total workforce across all plants was 51% (453/888). The study authors note that participation at the Chicago plant was non-representative for hourly workers relative to salaried workers. The study does not discuss recruitment or selection processes in detail but does indicate that workers could participate in the study during normal work hours without impact on their pay. Overall, details are sparse and it is unclear if those who did not participate may have been more or less exposed compared to those who were included.
	Metric 2:	Attrition	High	One participant out of 105 in the Chicago did not have a blood test taken and was excluded from analyses that included serum measures. No reason is provided for why the blood test was not taken, but there was no other indication of attrition or exclusion in the Chicago plant group or in the total included group.
	Metric 3:	Comparison Group	Low	Demographic details on sex and race are provided for all plants. The distribution of race and sex is roughly similar in the Chicago plant to the overall distribution of race and sex across the included workforce. However, these demographic differences are not controlled for in statistical analyses. There is also no discussion of age and whether groups had differences based on age.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Low	In comparisons between different plants, all Chicago plant workers were grouped together. Descriptions of the Chicago plant were very limited beyond stating that phthalic anhydride was used. A sub-group analysis of the Chicago plant separates out workers by ”type” (Phthalic, Maleic, or Coal Tar). It is unclear if phthalic anhydride workers were all actually formally exposed to phthalic anhydride, and it is not clear whether the other types of workers were not exposed. There is a large potential for exposure misclassification if not all tasks or employees were in contact with phthalic anhydride, and since exposure was estimated solely using professional judgment, exposure misclassification cannot be ruled out.
	Metric 5:	Exposure Levels	Low	The study only reports two levels of exposure, exposed and unexposed. No quantitative exposure information is available.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Nutrition & Metabolic
Reported Health Effect(s):	Serum CO2, serum triglycerides, serum methemoglobulin, serum cholesterol, serum glucose, urine pnehols
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
	Metric 6:	Temporality	Low	Work history or duration of employment was not discussed in the analysis. It is clear that the outcomes were measured after some degree of exposure, since outcomes were assessed in an occupational context. However, it is unclear how long these workers may have been exposed, and it is unclear whether the employees were free of any of the reported health outcomes prior to enrollment into the study. The temporality of exposure and outcome is uncertain.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Specific methods were not described, but description of the study suggests standard clinical methods were used (in line with NIOSH recommendations). Examination forms were provided in the appendices.
	Metric 8:	Reporting Bias	Medium	All outlined outcomes are reported in the results, however, data is provided as the proportion of individuals with ”abnormal” measurements compared against clinical standards.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Low	There was no evidence of adjustment for potential confounders, however, some were discussed in-text for medical interpretations. The distribution of sex and race is provided across exposure group, but there is no discussion of age. Covariates were not adjusted for or stratified in relation to outcomes.
	Metric 10:	Covariate Characterization	Medium	Questionnaires were provided in the appendix. It was not reported whether this was a validated questionnaire, however, there was no evidence to suggest it was an invalid instrument.
	Metric 11:	Co-exposure Counfounding	Low	Exposure to other coal tar components was described, including potential health effects. Workers in plants other than the Chicago plant were exposed to other occupational agents, and it is unclear whether those exposures contributed to the incidence of health outcomes. Within the Chicago plant, the study splits the sample into workers exposed to phthalic anhydride, maleic anhydride, or coal tar. It is not confirmed that phthalic anhydride workers were only exposed to phthalic anhydride.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The design limited the ability to determine which exposures were relevant to each health effect, however the design is sufficient to answer the study’s question as to whether there are ”abnormal” medical findings among workers at the studied coal tar plants.
	Metric 13:	Statistical Power	Low	Statistical power was not reported. A total of 105 employees completed exams from the Chicago plant. However, the analysis within the Chicago plant only identifies 14 workers who were exposed to phthalic anhydride, which may be too small of a sample size to detect an effect.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Nutrition & Metabolic
Reported Health Effect(s):	Serum CO2, serum triglycerides, serum methemoglobulin, serum cholesterol, serum glucose, urine pnehols
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what has been done and to be conceptually reproducible with access to the analytic data.
	Metric 15:	Statistical Analysis	N/A	Not applicable, no formal statistical analysis was conducted.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	N/A	Not applicable, no biomarkers of exposure were measured.
	Metric 17:	Effect Biomarker	High	All studied effect biomarkers were demonstrated to be related to adverse health outcomes and were collected from serum or urine samples.
	Metric 18:	Method Sensitivity	N/A	Limits of detection not relevant for the medical testing employed.
	Metric 19:	Biomarker Stability	Low	There is no description of the storage history and/or stability data for urine or serum samples.
	Metric 20:	Sample Contamination	Medium	There is no discussion of contamination.
	Metric 21:	Method Requirements	Medium	While no formal descriptions are provided, given the medical context of the examinations there is some confidence that accurate detection methodologies were used.
	Metric 22:	Matrix Adjustment	Medium	No discussion of matrices is described. This would be relevant for effect biomarkers measured in urine, but there is no evidence that there was no adjustment for creatinine.

Additional Comments: This occupational health surveillance study focused on several coal tar facilities in the United States. At the Chicago plant phthalic anhydride was present. The aim of the study was to compare medical findings across a wide range of health outcomes to established clinical ranges, thus no formal statistical analysis was available that compared exposed vs. unexposed. An examination of the reported data does not indicate that there were significant differences between workers exposed to phthalic anhydride in the Chicago plant and the overall population of workers, or workers at the Chicago plant not reported to be exposed to phthalic anhydride; however, this cannot be completely determined without a formal statistical analysis. There were several large concerns raised with the study, including the lack of an exposure measurement. All employees in the phthalic anhydride plant were considered exposed, however, without surveillance data or employment records, there is a large potential for exposure misclassification. There were also sparse details regarding recruitment, and a lack of consideration of potentially relevant covariates such as age. The workers in this study were all exposed to other occupational agents, thus confounding by other exposures cannot be ruled out.

Overall Quality Determination — Low

Study Citation:	Nielsen, J., Bensryd, I., Almquist, H., Dahlqvist, M., Welinder, H., Alexandersson, R., Skerfving, S. (1991). Serum IgE and lung function in workers exposed to phthalic anhydride. International Archives of Occupational and Environmental Health 63(3):199-204.
Health Outcome(s) Assessed:	Immune/Hematological
Reported Health Effect(s):	Serum antibodies (total serum IgE, specific serum IgE antibodies, specific serum IgG antibodies), symptoms of conjunctivitis, rhinitis, asthma, and non-specific bronchial hyperreactivity.
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5178100

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	Geographic setting was reported but temporal was not; inclusion/exclusion criteria were not articulated, and participation rate at different steps was not reported.
	Metric 2:	Attrition	Low	Numbers of individuals were not reported at important stages of study were not reported.
	Metric 3:	Comparison Group	Medium	There is indirect evidence that the groups were similar (authors reported that "The exposed workers and the control subjects were matched with regard to age and smoking habits" and provided limited comparative information).
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	Sampling was conducted for airborne contaminants and reported as total acid anhydrides (no data for PA alone) across 2 plants. Authors reported that controls were "not exposed to harmful chemical dust or smoke to any significant extent" but did not conduct any sampling for the control group.
	Metric 5:	Exposure Levels	Low	Reports 2 levels of exposure (exposed/unexposed)
	Metric 6:	Temporality	Low	PA's historical exposure levels were not estimated. PA's exposure sampling was for 7.9 hours of total sampling time and collected 24 samples. It is unclear whether exposures fall within relevant exposure windows for the outcome of interest.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	The outcome was assessed using self-reported symptoms, skin prick test, and serum antibodies (IgE, IgG) against phthalic anhydride-human serum albumin hapten conjugate. The authors did not report details of how the outcome assessments were performed on subjects (timing, sequence, etc.)
	Metric 8:	Reporting Bias	Medium	All outcomes were reported, but continuous data were given as median and range without SD/SE.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	Some covariates were considered by matching, e.g., "The exposed workers and the control subjects were matched with regard to age and smoking habits." All subjects were men and "Before the investigation no subject had any known lung disease." Authors evaluated atopy by history of symptoms and skin prick tests, and "There was no significant difference between the exposed workers and the control subjects with regard to smoking and history of atopy."

Continued on next page . . .

. . . continued from previous page

Study Citation:	Nielsen, J., Bensryd, I., Almquist, H., Dahlqvist, M., Welinder, H., Alexandersson, R., Skerfving, S. (1991). Serum IgE and lung function in workers exposed to phthalic anhydride. International Archives of Occupational and Environmental Health 63(3):199-204.
Health Outcome(s) Assessed:	Immune/Hematological
Reported Health Effect(s):	Serum antibodies (total serum IgE, specific serum IgE antibodies, specific serum IgG antibodies), symptoms of conjunctivitis, rhinitis, asthma, and non-specific bronchial hyperreactivity.
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5178100

Domain	Metric		Rating	Comments
	Metric 10:	Covariate Characterization	Medium	Smoking and history of atopy were assessed by interview and no information on the validity was provided. Atopy was also assessed by skin prick test.
	Metric 11:	Co-exposure Counfounding	Low	No adjustment for co-exposures was made and there is direct evidence for unbalanced provision of co-exposures. Exposed subjects had co-exposures to solvents and other acid anhydrides. Controls were ”employed in the municipal engineering department (repairmen and gardeners)” and ”were not exposed to harmful chemical dust or smoke to any significant extent.” Potential exposures to other chemicals and/or pesticides were not considered.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	Some covariates were considered by matching, e.g., ”The exposed workers and the control subjects were matched with regard to age and smoking habits.” All subjects were men and ”Before the investigation no subject had any known lung disease.” Authors evaluated atopy by history of symptoms and skin prick tests, and ”There was no significant difference between the exposed workers and the control subjects with regard to smoking and history of atopy.”
	Metric 13:	Statistical Power	Medium	Numbers of subjects were low but adequate to detect differences in some metrics.
	Metric 14:	Reproducibility of Analyses	Medium	Description sufficient to understand what was done and be conceptually reproducible
	Metric 15:	Statistical Analysis	N/A	statistical models were not applied

Additional Comments: Lung function and atopy/allergic responses were measured in 23 male workers exposed to phthalic anhydride and other acid anhydrides and compared with measurements in 18 male repairmen and gardeners employed in a municipal engineering department. Immunological outcomes included self-reported symptoms, skin prick test and serum antibodies against phthalic anhydride-human serum albumin hapten conjugate. Total acid anhydride levels in individual air samples averaged 6.6 mg/m3 (TWA) during phthalic anhydride loading. Exposed subjects had significantly higher total serum IgE (median 32 vs 15 kIU/L in controls) and IgG (median 0.21 vs 0.12 in controls) and significantly higher prevalence of non-specific bronchial hyperreactivity. No other metrics were significantly different. Exposed subjects had co-exposures to other acid anhydrides and solvents.

Overall Quality Determination — Low

Study Citation:	Nielsen, J., Welinder, H., Schutz, A., Skerfving, S. (1988). Specific serum antibodies against phthalic anhydride in occupationally exposed subjects. Journal of Allergy and Clinical Immunology 82(1):126-133.
Health Outcome(s) Assessed:	Immune/Hematological
Reported Health Effect(s):	Serum antibodies, symptoms of conjunctivitis, rhinitis, asthma
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176341

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	Geographic setting was reported but temporal was not; inclusion/exclusion criteria were not articulated, and participation rate at different steps was not reported.
	Metric 2:	Attrition	Low	Numbers of individuals were not reported at important stages of study were not reported.
	Metric 3:	Comparison Group	Medium	There is indirect evidence that the groups were similar; authors reported that control subjects with similar age distribution and smoking habits were randomly selected; Exposed groups consisted of men and women (2), while the food-processing reference group only consisted of males. limited comparative information was reported. A separate control group of 30 postal clerks was used for the analysis of specific IGM antibodies; sex of postal clerks was not specified and no comparative information of age and smoking was provided; it is not clear if any potential differences were controlled for in analysis.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	Sampling was conducted for airborne contaminants across 2 plants. Authors reported that controls (workers at a food-processing factory) "had not been exposed to any harmful chemical, smoke, or dust in their work environment " but did not conduct any sampling for the food processor or postal clerk control groups. In addition, the detailed employment records for exposed and control groups were not provided.
	Metric 5:	Exposure Levels	Low	Reports 3 qualitative levels of exposure (heavily exposed/lightly exposed/unexposed); Exposures of the 2 plants were similar so workers were "pooled". In addition, the lightly exposed areas in both plants only had 5 samples and the detected levels were <0.1 mg/m3. The exposure level of the control group is assumed to be zero, so the range of exposure in the population is limited.
	Metric 6:	Temporality	Medium	Temporality is established, but it is unclear whether exposures fall within relevant exposure windows for the outcome of interest. The average exposure duration for heavily exposed workers was 13 years (range 0-43 years) and 12 years for slightly exposed workers (range: 0.3-40 years).
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	The outcome was assessed using self-reported symptoms, skin prick test, and serum antibodies (IgE, IgG, IgM) against phthalic anhydride-human serum albumin hapten conjugate. The authors did not report details of how the outcome assessments were performed on subjects (timing, sequence, etc.)

Continued on next page . . .

. . . continued from previous page

Study Citation:	Nielsen, J., Welinder, H., Schutz, A., Skerfving, S. (1988). Specific serum antibodies against phthalic anhydride in occupationally exposed subjects. Journal of Allergy and Clinical Immunology 82(1):126-133.
Health Outcome(s) Assessed:	Immune/Hematological
Reported Health Effect(s):	Serum antibodies, symptoms of conjunctivitis, rhinitis, asthma
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176341

Domain	Metric		Rating	Comments
	Metric 8:	Reporting Bias	Medium	All outcomes were reported, but continuous data were given as median and range without SD/SE.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	Some covariates were considered by matching; authors reported that control subjects with similar age distribution and smoking habits. It was noted that there was no significant difference between the exposure groups with regard to smoking and history of atopy.
	Metric 10:	Covariate Characterization	Low	Smoking and history of atopy were assessed by interview and no information on the validity was provided. Atopy for common allergens was also assessed by skin prick test.
	Metric 11:	Co-exposure Counfounding	Low	No adjustment for co-exposures was made and there is direct evidence for unbalanced provision of co-exposures. Exposed subjects had co-exposures to solvents and other anhydrides. Controls were employed in a food-processing factory and ”had not been exposed to any harmful chemical, smoke, or dust in their work environment.” Potential exposures to other chemicals were not considered
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	Low	”For comparison of proportions, the chi-square test was used, and when expected numbers were < 5, Fisher’s exact test was used.” Association between pairs of variables were tested using Spearman’s rank-correlation test.This cross-sectional study only can show that there could be an association between PAD and lung functions, but cannot answer if only PAD caused lunch functions. There are two reasons for this unclear causality: (1) co-exposure from other chemicals in the plants were not assessed, (2) historical PAD exposure in the plants were not provided.
	Metric 13:	Statistical Power	Medium	Numbers of subjects were low but adequate to detect differences in some metrics
	Metric 14:	Reproducibility of Analyses	Medium	Description sufficient to understand what was done and be conceptually reproducible
	Metric 15:	Statistical Analysis	N/A	statistical models were not applied

Continued on next page . . .

. . . continued from previous page

Study Citation:	Nielsen, J., Welinder, H., Schutz, A., Skerfving, S. (1988). Specific serum antibodies against phthalic anhydride in occupationally exposed subjects. Journal of Allergy and Clinical Immunology 82(1):126-133.
Health Outcome(s) Assessed:	Immune/Hematological
Reported Health Effect(s):	Serum antibodies, symptoms of conjunctivitis, rhinitis, asthma
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176341

Domain	Metric	Rating	Comments
Additional Comments:	Lung function and atopy/allergic responses were measured in in 60 workers (58 men, 2 women) exposed to phthalic anhydride and other acid anhydrides and compared with measurements in 22 male workers employed in a food-processing factory; A separate control group of 30 subjects employed as postal clerks were used for the analysis of specific IgM antibodies. Exposure groups included heavy exposure (n=35), low exposure (n=25), and nonexposed (n=22). Phthalic anhydride levels in individual air samples averaged 6.6 mg/m3 (TWA) during phthalic anhydride loading. No significant difference was reported in total and specific serum levels of IgE, IgM among the exposure groups. There was no significant difference in total serum IgG among exposure groups, while there was a significant increase in specific IgG between heavy- and low-exposure groups (p=0.01). In subjects with symptoms of rhinoconjunctivitis, there was a significantly decreased total IgG compared to workers without symptoms (p = 0.01). Those with symptoms of asthma, had significantly increased specific IgG levels than subjects without symptoms of asthma (p=0.005). Exposed subjects had co-exposures to other acid anhydrides.		

Overall Quality Determination Low

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Immune/Hematological
Reported Health Effect(s):	Red blood cells, hemoglobin, hematocrit, mean corpuscular volume, mean corpuscular hemoglobin, mean corpuscular hemoglobin concentration, reticulocyte count, white blood cells, polymorphonuclear leukocytes, lymphocytes, monocytes, basophils, eosinophils, IgE, IgM, IgA, IgG
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	In this cross-sectional occupational surveillance study, workers from 9 Koppers Coal Tar plants were examined from January-October 1979. Only one of these plants (the Chicago plant) produced phthalic anhydride. Participation at the relevant phthalic anhydride plant (i.e., the Chicago plant) was approximately 46% (105/230), while participation from the total workforce across all plants was 51% (453/888). The study authors note that participation at the Chicago plant was non-representative for hourly workers relative to salaried workers. The study does not discuss recruitment or selection processes in detail but does indicate that workers could participate in the study during normal work hours without impact on their pay. Overall, details are sparse and it is unclear if those who did not participate may have been more or less exposed compared to those who were included.
	Metric 2:	Attrition	High	One participant out of 105 in the Chicago did not have a blood test taken and was excluded from analyses that included serum measures. No reason is provided for why the blood test was not taken, but there was no other indication of attrition or exclusion in the Chicago plant group or in the total included group.
	Metric 3:	Comparison Group	Low	Demographic details on sex and race are provided for all plants. The distribution of race and sex is roughly similar in the Chicago plant to the overall distribution of race and sex across the included workforce. However, these demographic differences are not controlled for in statistical analyses. There is also no discussion of age and whether groups had differences based on age.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Low	In comparisons between different plants, all Chicago plant workers were grouped together. Descriptions of the Chicago plant were very limited beyond stating that phthalic anhydride was used. A sub-group analysis of the Chicago plant separates out workers by ”type” (Phthalic, Maleic, or Coal Tar). It is unclear if phthalic anhydride workers were all actually formally exposed to phthalic anhydride, and it is not clear whether the other types of workers were not exposed. There is a large potential for exposure misclassification if not all tasks or employees were in contact with phthalic anhydride, and since exposure was estimated solely using professional judgment, exposure misclassification cannot be ruled out.
	Metric 5:	Exposure Levels	Low	The study only reports two levels of exposure, exposed and unexposed. No quantitative exposure information is available.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Immune/Hematological
Reported Health Effect(s):	Red blood cells, hemoglobin, hematocrit, mean corpuscular volume, mean corpuscular hemoglobin, mean corpuscular hemoglobin concentration, reticulocyte count, white blood cells, polymorphonuclear leukocytes, lymphocytes, monocytes, basophils, eosinophils, IgE, IgM, IgA, IgG
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
	Metric 6:	Temporality	Low	Work history or duration of employment was not discussed in the analysis. It is clear that the outcomes were measured after some degree of exposure, since outcomes were assessed in an occupational context. However, it is unclear how long these workers may have been exposed, and it is unclear whether the employees were free of any of the reported health outcomes prior to enrollment into the study. The temporality of exposure and outcome is uncertain.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Specific methods were not described, but description of the study suggests standard clinical methods were used (in line with NIOSH recommendations). Examination forms were provided in the appendices.
	Metric 8:	Reporting Bias	Medium	All outlined outcomes are reported in the results, however, data is provided as the proportion of individuals with ”abnormal” measurements compared against clinical standards.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Low	There was no evidence of adjustment for potential confounders, however, some were discussed in-text for medical interpretations. The distribution of sex and race is provided across exposure group, but there is no discussion of age. Covariates were not adjusted for or stratified in relation to outcomes.
	Metric 10:	Covariate Characterization	Medium	Questionnaires were provided in the appendix. It was not reported whether this was a validated questionnaire, however, there was no evidence to suggest it was an invalid instrument.
	Metric 11:	Co-exposure Counfounding	Low	Exposure to other coal tar components was described, including potential health effects. Workers in plants other than the Chicago plant were exposed to other occupational agents, and it is unclear whether those exposures contributed to the incidence of health outcomes. Within the Chicago plant, the study splits the sample into workers exposed to phthalic anhydride, maleic anhydride, or coal tar. It is not confirmed that phthalic anhydride workers were only exposed to phthalic anhydride.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The design limited the ability to determine which exposures were relevant to each health effect, however the design is sufficient to answer the study’s question as to whether there are ”abnormal” medical findings among workers at the studied coal tar plants.
	Metric 13:	Statistical Power	Low	Statistical power was not reported. A total of 105 employees completed exams from the Chicago plant. However, the analysis within the Chicago plant only identifies 14 workers who were exposed to phthalic anhydride, which may be too small of a sample size to detect an effect.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Immune/Hematological
Reported Health Effect(s):	Red blood cells, hemoglobin, hematocrit, mean corpuscular volume, mean corpuscular hemoglobin, mean corpuscular hemoglobin concentration, reticulocyte count, white blood cells, polymorphonuclear leukocytes, lymphocytes, monocytes, basophils, eosinophils, IgE, IgM, IgA, IgG
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what has been done and to be conceptually reproducible with access to the analytic data.
	Metric 15:	Statistical Analysis	N/A	Not applicable, no formal statistical analysis was conducted.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	N/A	Not applicable, no biomarkers of exposure were measured.
	Metric 17:	Effect Biomarker	High	All studied effect biomarkers were demonstrated to be related to adverse health outcomes and were collected from serum or urine samples.
	Metric 18:	Method Sensitivity	N/A	Limits of detection not relevant for the medical testing employed.
	Metric 19:	Biomarker Stability	Low	There is no description of the storage history and/or stability data for urine or serum samples.
	Metric 20:	Sample Contamination	Medium	There is no discussion of contamination.
	Metric 21:	Method Requirements	Medium	While no formal descriptions are provided, given the medical context of the examinations there is some confidence that accurate detection methodologies were used.
	Metric 22:	Matrix Adjustment	Medium	No discussion of matrices is described. This would be relevant for effect biomarkers measured in urine, but there is no evidence that there was no adjustment for creatinine.

Additional Comments: This occupational health surveillance study focused on several coal tar facilities in the United States. At the Chicago plant phthalic anhydride was present. The aim of the study was to compare medical findings across a wide range of health outcomes to established clinical ranges, thus no formal statistical analysis was available that compared exposed vs. unexposed. An examination of the reported data does not indicate that there were significant differences between workers exposed to phthalic anhydride in the Chicago plant and the overall population of workers, or workers at the Chicago plant not reported to be exposed to phthalic anhydride; however, this cannot be completely determined without a formal statistical analysis. There were several large concerns raised with the study, including the lack of an exposure measurement. All employees in the phthalic anhydride plant were considered exposed, however, without surveillance data or employment records, there is a large potential for exposure misclassification. There were also sparse details regarding recruitment, and a lack of consideration of potentially relevant covariates such as age. The workers in this study were all exposed to other occupational agents, thus confounding by other exposures cannot be ruled out.

Overall Quality Determination — Low

Study Citation:	Wernfors, M., Nielsen, J., Schütz, A., Skerfving, S. (1986). Phthalic anhydride-induced occupational asthma. International Archives of Allergy and Applied Immunology 79(1):77-82.
Health Outcome(s) Assessed:	Immune/Hematological
Reported Health Effect(s):	Irritation of the upper airways, rhinitis, asthma, PA-induced asthma, chronic productive bronchitis, spirometry
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176303

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	This occupational cohort study examined workers from two plants that utilized flaked phthalic anhydride. Two separate groups were recruited - current employees, and former employees. There is no discussion of how the currently employed were recruited or selected, and it is unclear whether the 48 current employees represent a large proportion of the total eligible population and whether or not the selected employees were similar to the eligible population. More details were provided for former employees, as the study specifies that inquiries were mailed to 85 former employees, and that 70 (82%) replied. It is not specified whether they contacted all former employees, and there is no other description of the total eligible population.Four plants (A, B, C, and D) are mentioned as part of the study, but exposure is only described for plants A and B, and no further description of plants is provided. It is unclear whether participants were recruited from all plants, or just A and B. The study later states that asthma was reported in Plant A by 11 of 42 former and 5 of 28 present employees, and in Plant B (4/23 and 1/12). These numbers imply that 8 present employees and 5 former employees may have come from Plant C and D to reach the total numbers of 48 current employees and 70 former employees, but this information is not explicitly spelled out, and the differences in numbers may be due to other factors. Due to a lack of detailed information, selection bias cannot be ruled out.
	Metric 2:	Attrition	High	No attrition was reported. Results tables do not indicate any loss or exclusion.
	Metric 3:	Comparison Group	Low	The study authors do not address whether groups were similar, and covariates were not addressed in the statistical analysis. No demographic information (beyond smoking) was provided.
Domain 2: Exposure Characterization				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Wernfors, M., Nielsen, J., Schütz, A., Skerfving, S. (1986). Phthalic anhydride-induced occupational asthma. International Archives of Allergy and Applied Immunology 79(1):77-82.
Health Outcome(s) Assessed:	Immune/Hematological
Reported Health Effect(s):	Irritation of the upper airways, rhinitis, asthma, PA-induced asthma, chronic productive bronchitis, spirometry
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176303

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	Low	PA levels were determined in plants A and B by use of two different methods. Some plant A personal samples were collected in glass fiber filters and PA was determined using UV spectrophotometry. The LOD for this method was 0.1 mg/m3 per 200L air sample. For other samples in plant A, and those in plant B, personal samples were collected using battery-operated pumps and respirable particle samplers in some cases. In the second method was determined using HPLC, and the LOD was 0.03 mg/m3 per 20L air sample. The methods for exposure assessment in plants C and D are not described, and it is unclear whether participants were even recruited from those two plants. However, estimates previously calculated above in Metric 1 imply that the total number of participants from Plant C and D may have been only 13, a relatively small proportion of the total sample. It is unclear whether these exposure levels were detected among the actual participants of the present study, and it is unclear whether the exposure levels are applicable to former employees. There is potential for exposure misclassification due to the inconsistency of exposure measurements and the uncertainty of whether these exposure measurements apply to the total sample population.
	Metric 5:	Exposure Levels	Low	PA measurements were reported by plant and work operation. Participants were all treated as ”exposed”. Some comparisons were made with former employees, however, it’s not clear if they were assigned a separate exposure level.
	Metric 6:	Temporality	Medium	The study authors describe a ”latency period between start of employment and onset of respiratory symptoms” which ”ranged from 1 month to 16 years.” Temporality was established, however, it was not entirely clear if these were all incident cases.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Questionnaires were used to assess the presence of relevant conditions (e.g., chronic bronchitis, asthma, etc.). Spirometry was performed using a Vitalograph, with specific pathological spirometric values defined. A subset of participants were subjected to skin-prick testing with 15 common allergens - these specific allergens are not provided, but the source of the allergens is specified to be Allergologisk Laboratorium (Copenhagen, Denmark). Diagnostic definitions for positive skin-prick tests are provided.
	Metric 8:	Reporting Bias	Medium	Outcomes mentioned in the abstract, introduction, and methods were mostly addressed. Study authors provided spirometry results only for asthmatic employees. Most results were simple counts or proportions. Ranges were provided in some cases.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Wernfors, M., Nielsen, J., Schütz, A., Skerfving, S. (1986). Phthalic anhydride-induced occupational asthma. International Archives of Allergy and Applied Immunology 79(1):77-82.
Health Outcome(s) Assessed:	Immune/Hematological
Reported Health Effect(s):	Irritation of the upper airways, rhinitis, asthma, PA-induced asthma, chronic productive bronchitis, spirometry
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176303

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	Low	Covariates were largely not addressed. There was some qualitative consideration of smoking in the analysis, but no formal stratification. The range of employment duration is provided, but it is not clear how the range relates to exposure and outcome groups. Other potential covariates, such as age, are not described. Thus, considerations were not made for confounders and the distribution of primary covariates and potential co-founders is not reported.
	Metric 10:	Covariate Characterization	Medium	Smoking was assessed through survey and questionnaire. Duration of employment was likely assessed via company records.
	Metric 11:	Co-exposure Counfounding	Medium	There is some potential for co-exposures in this setting, particularly with other anhydrides (i.e., maelic and trimellitic), however, the study authors describe the use of these materials as infrequent.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	Low	Prevalence and severity of symptoms were analyzed in an occupational cohort. The design of the study limited understanding dose-response relationships and clear comparisons with unexposed or less exposed individuals.
	Metric 13:	Statistical Power	Medium	Statistical power was not reported, however, the number of current and former employees was likely sufficient to detect an effect.
	Metric 14:	Reproducibility of Analyses	Low	The analysis was not described in detail, as it is unclear which exact Plants participants originated from and how exactly exposure was determined.
	Metric 15:	Statistical Analysis	High	Chi-square tests were used to compare different groups (i.e., current and former employees, asthmatic vs non-asthmatic, etc.). No concerns were identified.

Additional Comments: This study utilized current and former employees of a phthalic anhydride resin plant to investigate the effects of PAD on respiratory and immune endpoints. There were a few concerns regarding this study, including the participation rate and lack of detail regarding recruitment of current workers. Additionally, there was some uncertainty on exposure classification. Exposure values were only measured in two of four plants, and exposure was assumed for all employees. The analytical design also limits the ability to assess differences in outcomes by exposure level.

Overall Quality Determination — Low

Study Citation:	Nielsen, J., Welinder, H., Schutz, A., Skerfving, S. (1988). Specific serum antibodies against phthalic anhydride in occupationally exposed subjects. Journal of Allergy and Clinical Immunology 82(1):126-133.
Health Outcome(s) Assessed:	Sensitization
Reported Health Effect(s):	skin prick test and PA-HSA-specific immunoglobulins
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176341

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	Geographic setting was reported but temporal was not; inclusion/exclusion criteria were not articulated, and participation rate at different steps was not reported.
	Metric 2:	Attrition	Low	Numbers of individuals were not reported at important stages of study were not reported.
	Metric 3:	Comparison Group	Medium	There is indirect evidence that the groups were similar; authors reported that control subjects with similar age distribution and smoking habits were randomly selected; Exposed groups consisted of men and women (2), while the reference group only consisted of males. Limited comparative information was reported.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	Sampling was conducted for airborne contaminants across 2 plants. Authors reported that controls (workers at a food-processing factory) "had not been exposed to any harmful chemical, smoke, or dust in their work environment " but did not conduct any sampling for the control group. In addition, the detailed employment records for exposed and control groups were not provided.
	Metric 5:	Exposure Levels	Low	Reports 3 qualitative levels of exposure (heavily exposed/lightly exposed/unexposed); Exposures of the 2 plants were similar so workers were "pooled". In addition, the lightly exposed areas in both plants only had 5 samples and the detected levels were <0.1 mg/m3. The exposure level of the control group is assumed to be zero, so the range of exposure in the population is limited.
	Metric 6:	Temporality	Medium	Temporality is established, but it is unclear whether exposures fall within relevant exposure windows for the outcome of interest. The average exposure duration for heavily exposed workers was 13 years (range 0-43 years) and 12 years for slightly exposed workers (range: 0.3-40 years).
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Sensitization was assessed using a skin prick test and serum antibodies (IgE, IgG, IgM) against phthalic anhydride-human serum albumin hapten conjugate. The authors did not report details of how the outcome assessments were performed on subjects (timing, sequence, etc.), however, details were provided on RAST and ELISA assays and appeared standard.
	Metric 8:	Reporting Bias	High	All outcomes were reported as incidence and % of exposure group.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Nielsen, J., Welinder, H., Schutz, A., Skerfving, S. (1988). Specific serum antibodies against phthalic anhydride in occupationally exposed subjects. Journal of Allergy and Clinical Immunology 82(1):126-133.
Health Outcome(s) Assessed:	Sensitization
Reported Health Effect(s):	skin prick test and PA-HSA-specific immunoglobulins
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176341

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	High	Some covariates were considered by matching; authors reported that control subjects with similar age distribution and smoking habits. It was noted that there was no significant difference between the exposure groups with regard to smoking and history of atopy.
	Metric 10:	Covariate Characterization	Low	Smoking and history of atopy were assessed by interview and no information on the validity was provided. Atopy for common allergens was also assessed by skin prick test.
	Metric 11:	Co-exposure Counfounding	Low	No adjustment for co-exposures was made and there is direct evidence for unbalanced provision of co-exposures. Exposed subjects had co-exposures to solvents and other anhydrides. Controls were employed in a food-processing factory and "had not been exposed to any harmful chemical, smoke, or dust in their work environment." Potential exposures to other chemicals and/or pesticides were not considered.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	Low	"For comparison of proportions, the chi-square test was used, and when expected numbers were < 5, Fisher's exact test was used." Association between pairs of variables were tested using Spearman's rank-correlation test. This cross-sectional study only can show that there could be an association between PAD and lung functions, but cannot answer if only PAD caused lunch functions. There are two reasons for this unclear causality: (1) co-exposure from other chemicals in the plants were not assessed, (2) historical PAD exposure in the plants were not provided.
	Metric 13:	Statistical Power	Medium	Numbers of subjects were low but adequate to detect differences in some metrics; there were no significant differences in respiratory outcomes (chronic bronchitis).
	Metric 14:	Reproducibility of Analyses	Medium	Description sufficient to understand what was done and be conceptually reproducible.
	Metric 15:	Statistical Analysis	N/A	Statistical models were not applied.

Additional Comments: Lung function and atopy/allergic responses were measured in 60 workers (58 men, 2 women) exposed to phthalic anhydride and other acid anhydrides and compared with measurements in 22 male workers employed in a food-processing factory. Exposure groups included heavy exposure (n=35), low exposure (n=25), and non-exposed (n=22). Respiratory outcomes, including chronic bronchitis, included self-reported symptoms. Phthalic anhydride levels in individual air samples averaged 6.6 mg/m3 (TWA) during phthalic anhydride loading. Six (17%) of heavily exposed subjects had chronic bronchitis and one (4%) of low exposed workers had chronic bronchitis; the difference was not statistically significant. There was no difference between the exposure groups regarding total serum levels of IgE, IgG, and IgM, nor specific IgE and IgM against PAH. There was a significant difference of specific IgG against PAH between heavy and low exposure groups.

Overall Quality Determination Medium

Study Citation:	Wernfors, M., Nielsen, J., Schütz, A., Skerfving, S. (1986). Phthalic anhydride-induced occupational asthma. International Archives of Allergy and Applied Immunology 79(1):77-82.
Health Outcome(s) Assessed:	Sensitization
Reported Health Effect(s):	Irritation of the upper airways, rhinitis, asthma, PA-induced asthma, chronic productive bronchitis, spirometry
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176303

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	This occupational cohort study examined workers from two plants that utilized flaked phthalic anhydride. Two separate groups were recruited - current employees, and former employees. There is no discussion of how the currently employed were recruited or selected, and it is unclear whether the 48 current employees represent a large proportion of the total eligible population and whether or not the selected employees were similar to the eligible population. More details were provided for former employees, as the study specifies that inquiries were mailed to 85 former employees, and that 70 (82%) replied. It is not specified whether they contacted all former employees, and there is no other description of the total eligible population.Four plants (A, B, C, and D) are mentioned as part of the study, but exposure is only described for plants A and B, and no further description of plants is provided. It is unclear whether participants were recruited from all plants, or just A and B. The study later states that asthma was reported in Plant A by 11 of 42 former and 5 of 28 present employees, and in Plant B (4/23 and 1/12). These numbers imply that 8 present employees and 5 former employees may have come from Plant C and D to reach the total numbers of 48 current employees and 70 former employees, but this information is not explicitly spelled out, and the differences in numbers may be due to other factors. Due to a lack of detailed information, selection bias cannot be ruled out.
	Metric 2:	Attrition	High	No attrition was reported. Results tables do not indicate any loss or exclusion.
	Metric 3:	Comparison Group	Low	The study authors do not address whether groups were similar, and covariates were not addressed in the statistical analysis. No demographic information (beyond smoking) was provided.
Domain 2: Exposure Characterization				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Wernfors, M., Nielsen, J., Schütz, A., Skerfving, S. (1986). Phthalic anhydride-induced occupational asthma. International Archives of Allergy and Applied Immunology 79(1):77-82.
Health Outcome(s) Assessed:	Sensitization
Reported Health Effect(s):	Irritation of the upper airways, rhinitis, asthma, PA-induced asthma, chronic productive bronchitis, spirometry
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176303

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	Low	PA levels were determined in plants A and B by use of two different methods. Some plant A personal samples were collected in glass fiber filters and PA was determined using UV spectrophotometry. The LOD for this method was 0.1 mg/m3 per 200L air sample. For other samples in plant A, and those in plant B, personal samples were collected using battery-operated pumps and respirable particle samplers in some cases. In the second method was determined using HPLC, and the LOD was 0.03 mg/m3 per 20L air sample. The methods for exposure assessment in plants C and D are not described, and it is unclear whether participants were even recruited from those two plants. However, estimates previously calculated above in Metric 1 imply that the total number of participants from Plant C and D may have been only 13, a relatively small proportion of the total sample. It is unclear whether these exposure levels were detected among the actual participants of the present study, and it is unclear whether the exposure levels are applicable to former employees. There is potential for exposure misclassification due to the inconsistency of exposure measurements and the uncertainty of whether these exposure measurements apply to the total sample population.
	Metric 5:	Exposure Levels	Low	PA measurements were reported by plant and work operation. Participants were all treated as ”exposed”. Some comparisons were made with former employees, however, it’s not clear if they were assigned a separate exposure level.
	Metric 6:	Temporality	Medium	The study authors describe a ”latency period between start of employment and onset of respiratory symptoms” which ”ranged from 1 month to 16 years.” Temporality was established, however, it was not entirely clear if these were all incident cases.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Questionnaires were used to assess the presence of relevant conditions (e.g., chronic bronchitis, asthma, etc.). Spirometry was performed using a Vitalograph, with specific pathological spirometric values defined. A subset of participants were subjected to skin-prick testing with 15 common allergens - these specific allergens are not provided, but the source of the allergens is specified to be Allergologisk Laboratorium (Copenhagen, Denmark). Diagnostic definitions for positive skin-prick tests are provided.
	Metric 8:	Reporting Bias	Medium	Outcomes mentioned in the abstract, introduction, and methods were mostly addressed. Study authors provided spirometry results only for asthmatic employees. Most results were simple counts or proportions. Ranges were provided in some cases.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Wernfors, M., Nielsen, J., Schütz, A., Skerfving, S. (1986). Phthalic anhydride-induced occupational asthma. International Archives of Allergy and Applied Immunology 79(1):77-82.
Health Outcome(s) Assessed:	Sensitization
Reported Health Effect(s):	Irritation of the upper airways, rhinitis, asthma, PA-induced asthma, chronic productive bronchitis, spirometry
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176303

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	Low	Covariates were largely not addressed. There was some qualitative consideration of smoking in the analysis, but no formal stratification. The range of employment duration is provided, but it is not clear how the range relates to exposure and outcome groups. Other potential covariates, such as age, are not described. Thus, considerations were not made for confounders and the distribution of primary covariates and potential co-founders is not reported.
	Metric 10:	Covariate Characterization	Medium	Smoking was assessed through survey and questionnaire. Duration of employment was likely assessed via company records.
	Metric 11:	Co-exposure Counfounding	Medium	There is some potential for co-exposures in this setting, particularly with other anhydrides (i.e., maelic and trimellitic), however, the study authors describe the use of these materials as infrequent.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	Low	Prevalence and severity of symptoms were analyzed in an occupational cohort. The design of the study limited understanding dose-response relationships and clear comparisons with unexposed or less exposed individuals.
	Metric 13:	Statistical Power	Medium	Statistical power was not reported, however, the number of current and former employees was likely sufficient to detect an effect.
	Metric 14:	Reproducibility of Analyses	Low	The analysis was not described in detail, as it is unclear which exact Plants participants originated from and how exactly exposure was determined.
	Metric 15:	Statistical Analysis	High	Chi-square tests were used to compare different groups (i.e., current and former employees, asthmatic vs non-asthmatic, etc.). No concerns were identified.

Additional Comments: This study utilized current and former employees of a phthalic anhydride resin plant to investigate the effects of PAD on respiratory and immune endpoints. There were a few concerns regarding this study, including the participation rate and lack of detail regarding recruitment of current workers. Additionally, there was some uncertainty on exposure classification. Exposure values were only measured in two of four plants, and exposure was assumed for all employees. The analytical design also limits the ability to assess differences in outcomes by exposure level.

Overall Quality Determination — Low

Study Citation:	Nielsen, J., Welinder, H., Schutz, A., Skerfving, S. (1988). Specific serum antibodies against phthalic anhydride in occupationally exposed subjects. Journal of Allergy and Clinical Immunology 82(1):126-133.
Health Outcome(s) Assessed:	Ocular & Sensory
Reported Health Effect(s):	conjunctivitis
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176341

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	Geographic setting was reported but temporal was not; inclusion/exclusion criteria were not articulated, and participation rate at different steps was not reported.
	Metric 2:	Attrition	Low	Numbers of individuals were not reported at important stages of study were not reported.
	Metric 3:	Comparison Group	Medium	There is indirect evidence that the groups were similar; authors reported that control subjects with similar age distribution and smoking habits were randomly selected; Exposed groups consisted of men and women (2), while the reference group only consisted of males. Limited comparative information was reported.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	Sampling was conducted for airborne contaminants across 2 plants. Authors reported that controls (workers at a food-processing factory) "had not been exposed to any harmful chemical, smoke, or dust in their work environment " but did not conduct any sampling for the control group. In addition, the detailed employment records for exposed and control groups were not provided.
	Metric 5:	Exposure Levels	Low	Reports 3 qualitative levels of exposure (heavily exposed/lightly exposed/unexposed); Exposures of the 2 plants were similar so workers were "pooled". In addition, the lightly exposed areas in both plants only had 5 samples and the detected levels were <0.1 mg/m3. The exposure level of the control group is assumed to be zero, so the range of exposure in the population is limited.
	Metric 6:	Temporality	Medium	Temporality is established, but it is unclear whether exposures fall within relevant exposure windows for the outcome of interest. The average exposure duration for heavily exposed workers was 13 years (range 0-43 years) and 12 years for slightly exposed workers (range: 0.3-40 years).
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Low	Conjunctivitis was assessed using self-reported symptoms using a form administered and recorded by a physician.
	Metric 8:	Reporting Bias	High	All outcomes were reported as incidence and % of exposure group.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Nielsen, J., Welinder, H., Schutz, A., Skerfving, S. (1988). Specific serum antibodies against phthalic anhydride in occupationally exposed subjects. Journal of Allergy and Clinical Immunology 82(1):126-133.
Health Outcome(s) Assessed:	Ocular & Sensory
Reported Health Effect(s):	conjunctivitis
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176341

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	High	Some covariates were considered by matching; authors reported that control subjects with similar age distribution and smoking habits. It was noted that there was no significant difference between the exposure groups with regard to smoking and history of atopy.
	Metric 10:	Covariate Characterization	Low	Smoking and history of atopy were assessed by interview and no information on the validity was provided. Atopy for common allergens was also assessed by skin prick test.
	Metric 11:	Co-exposure Counfounding	Low	No adjustment for co-exposures was made and there is direct evidence for unbalanced provision of co-exposures. Exposed subjects had co-exposures to solvents and other anhydrides. Controls were employed in a food-processing factory and "had not been exposed to any harmful chemical, smoke, or dust in their work environment." Potential exposures to other chemicals and/or pesticides were not considered.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	Low	"For comparison of proportions, the chi-square test was used, and when expected numbers were < 5, Fisher's exact test was used." Association between pairs of variables were tested using Spearman's rank-correlation test. This cross-sectional study only can show that there could be an association between PAD and lung functions, but cannot answer if only PAD caused lunch functions. There are two reasons for this unclear causality: (1) co-exposure from other chemicals in the plants were not assessed, (2) historical PAD exposure in the plants were not provided.
	Metric 13:	Statistical Power	Medium	Numbers of subjects were low but adequate to detect differences in some metrics; there were no significant differences in respiratory outcomes (chronic bronchitis).
	Metric 14:	Reproducibility of Analyses	Medium	Description sufficient to understand what was done and be conceptually reproducible.
	Metric 15:	Statistical Analysis	N/A	Statistical models were not applied.

Additional Comments: Conjunctivitis, lung function and atopy/allergic responses were measured in 60 workers (58 men, 2 women) exposed to phthalic anhydride and other acid anhydrides and compared with measurements in 22 male workers employed in a food-processing factory. Exposure groups included heavy exposure (n=35), low exposure (n=25), and non-exposed (n=22). Respiratory outcomes, including chronic bronchitis, included self-reported symptoms. Phthalic anhydride levels in individual air samples averaged 6.6 mg/m3 (TWA) during phthalic anhydride loading. Six (17%) of heavily exposed subjects had chronic bronchitis and one (4%) of low exposed workers had chronic bronchitis; the difference was not statistically significant. Sixteen (46%) of heavily exposed subjects and 5 (%) of low exposed subjects had conjunctivitis. Since there is no non-exposed subjects with conjunctivitis in this study, difference between exposed and non-exposed groups could not be compared.

Overall Quality Determination Low

Study Citation:	Nielsen, J., Bensryd, I., Almquist, H., Dahlqvist, M., Welinder, H., Alexandersson, R., Skerfving, S. (1991). Serum IgE and lung function in workers exposed to phthalic anhydride. International Archives of Occupational and Environmental Health 63(3):199-204.
Health Outcome(s) Assessed:	Irritation
Reported Health Effect(s):	Self-reported conjunctivitis, rhinitis, rhino-conjunctivitis, dry cough, non-specific bronchial hyperactivity, and chronic bronchitis.
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5178100

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	Geographic setting was reported but temporal was not; inclusion/exclusion criteria were not articulated, and participation rate at different steps was not reported.
	Metric 2:	Attrition	Low	Numbers of individuals were not reported at important stages of study were not reported.
	Metric 3:	Comparison Group	Medium	There is indirect evidence that the groups were similar (authors reported that "The exposed workers and the control subjects were matched with regard to age and smoking habits" and provided limited comparative information).
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	Sampling was conducted for airborne contaminants and reported as total acid anhydrides (no data for PA alone) across 2 plants. Authors reported that controls were "not exposed to harmful chemical dust or smoke to any significant extent" but did not conduct any sampling for the control group.
	Metric 5:	Exposure Levels	Low	Reports 2 levels of exposure (exposed/unexposed)
	Metric 6:	Temporality	Low	PA's historical exposure levels were not estimated. PA's exposure sampling was for 7.9 hours of total sampling time and collected 24 samples. It is unclear whether exposures fall within relevant exposure windows for the outcome of interest.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Low	The outcome was assessed using self-reported symptoms.
	Metric 8:	Reporting Bias	High	All outcomes were reported.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	Some covariates were considered by matching, e.g., "The exposed workers and the control subjects were matched with regard to age and smoking habits." All subjects were men and "Before the investigation no subject had any known lung disease." Authors evaluated atopy by history of symptoms and skin prick tests, and "There was no significant difference between the exposed workers and the control subjects with regard to smoking and history of atopy."
	Metric 10:	Covariate Characterization	Low	Smoking and history of atopy were assessed by interview and no information on the validity was provided. Atopy was also assessed by skin prick test.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Nielsen, J., Bensryd, I., Almquist, H., Dahlqvist, M., Welinder, H., Alexandersson, R., Skerfving, S. (1991). Serum IgE and lung function in workers exposed to phthalic anhydride. International Archives of Occupational and Environmental Health 63(3):199-204.
Health Outcome(s) Assessed:	Irritation
Reported Health Effect(s):	Self-reported conjunctivitis, rhinitis, rhino-conjunctivitis, dry cough, non-specific bronchial hyperactivity, and chronic bronchitis.
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5178100

Domain	Metric		Rating	Comments
	Metric 11:	Co-exposure Counfounding	Low	No adjustment for co-exposures was made and there is direct evidence for unbalanced provision of co-exposures. Exposed subjects had co-exposures to solvents and other acid anhydrides. Controls were ”employed in the municipal engineering department (repairmen and gardeners)” and ”were not exposed to harmful chemical dust or smoke to any significant extent.” Potential exposures to other chemicals and/or pesticides were not considered.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	Low	Study design was appropriate to evaluate the association between exposure and health outcomes in a cross-sectional study. For example, the manuscript showed that ”For comparison of distributions between different groups (exposed workers and controls; groups with different symptoms), the Mann-Whitney U-test was used. For comparison of symptoms in the exposed group and control group, Fisher’s test was applied. All stated P-values involved two-tailed analysis; differences were considered to be statistically significant at P < 0 05.”However, the statistical analyses did not put co-exposure pollutants into consideration to adjust the association between exposure and health outcomes.
	Metric 13:	Statistical Power	Medium	Numbers of subjects were low but adequate to detect differences in some metrics.
	Metric 14:	Reproducibility of Analyses	Medium	Description sufficient to understand what was done and be conceptually reproducible
	Metric 15:	Statistical Analysis	N/A	statistical models were not applied

Additional Comments: Self-reported symptoms were evaluated in 23 male workers exposed to phthalic anhydride and other acid anhydrides and compared with measurements in 18 male repairmen and gardeners employed in a municipal engineering department. Total acid anhydride levels in individual air samples averaged 6.6 mg/m3 (TWA) during phthalic anhydride loading. Exposed subjects had significantly higher prevalence of work-related conjunctivitis (48% vs 6% in controls), total conjunctivitis (48% vs 17%), and work-related rhino-conjunctivitis (22% vs 0%). No differences in the prevalence of rhinitis, dry cough, or chronic bronchitis were noted. Exposed subjects had co-exposures to other acid anhydrides and solvents.

Overall Quality Determination — Low

Study Citation:	Nielsen, J., Welinder, H., Schutz, A., Skerfving, S. (1988). Specific serum antibodies against phthalic anhydride in occupationally exposed subjects. Journal of Allergy and Clinical Immunology 82(1):126-133.
Health Outcome(s) Assessed:	Irritation
Reported Health Effect(s):	conjunctivitis, rhinitis, rhino-conjunctivitis
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176341

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	Geographic setting was reported but temporal was not; inclusion/exclusion criteria were not articulated, and participation rate at different steps was not reported.
	Metric 2:	Attrition	Low	Numbers of individuals were not reported at important stages of study were not reported.
	Metric 3:	Comparison Group	Medium	There is indirect evidence that the groups were similar; authors reported that control subjects with similar age distribution and smoking habits were randomly selected; Exposed groups consisted of men and women (2), while the reference group only consisted of males. limited comparative information was reported.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Medium	Sampling was conducted for airborne contaminants across 2 plants. Authors reported that controls (workers at a food-processing factory) "had not been exposed to any harmful chemical, smoke, or dust in their work environment " but did not conduct any sampling for the control group. In addition, the detailed employment records for exposed and control groups were not provided.
	Metric 5:	Exposure Levels	Low	Reports 3 qualitative levels of exposure (heavily exposed/lightly exposed/unexposed); Exposures of the 2 plants were similar so workers were "pooled". In addition, the lightly exposed areas in both plants only had 5 samples and the detected levels were <0.1 mg/m3. The exposure level of the control group is assumed to be zero, so the range of exposure in the population is limited.
	Metric 6:	Temporality	Medium	Temporality is established, but it is unclear whether exposures fall within relevant exposure windows for the outcome of interest. The average exposure duration for heavily exposed workers was 13 years (range 0-43 years) and 12 years for slightly exposed workers (range: 0.3-40 years).
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Low	The outcome was assessed using self-reported symptoms.
	Metric 8:	Reporting Bias	High	all outcomes were reported as incidence and % of exposure group.
Domain 4: Potential Confounding / Variability Control				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Nielsen, J., Welinder, H., Schutz, A., Skerfving, S. (1988). Specific serum antibodies against phthalic anhydride in occupationally exposed subjects. Journal of Allergy and Clinical Immunology 82(1):126-133.
Health Outcome(s) Assessed:	Irritation
Reported Health Effect(s):	conjunctivitis, rhinitis, rhino-conjunctivitis
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5176341

Domain	Metric		Rating	Comments
	Metric 9:	Covariate Adjustment	High	Some covariates were considered by matching; authors reported that control subjects with similar age distribution and smoking habits. It was noted that there was no significant difference between the exposure groups with regard to smoking and history of atopy.
	Metric 10:	Covariate Characterization	Low	Smoking and history of atopy were assessed by interview and no information on the validity was provided. Atopy for common allergens was also assessed by skin prick test.
	Metric 11:	Co-exposure Counfounding	Low	No adjustment for co-exposures was made and there is direct evidence for unbalanced provision of co-exposures. Exposed subjects had co-exposures to solvents and other anhydrides. Controls were employed in a food-processing factory and "had not been exposed to any harmful chemical, smoke, or dust in their work environment." Potential exposures to other chemicals and/or pesticides were not considered
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	Low	"For comparison of proportions, the chi-square test was used, and when expected numbers were < 5, Fisher's exact test was used." Association between pairs of variables were tested using Spearman's rank-correlation test.This cross-sectional study only can show that there could be an association between PAD and lung functions, but cannot answer if only PAD caused lunch functions. There are two reasons for this unclear causality: (1) co-exposure from other chemicals in the plants were not assessed, (2) historical PAD exposure in the plants were not provided.
	Metric 13:	Statistical Power	Medium	Numbers of subjects were low but adequate to detect differences in some metrics; though there were no significant differences in irritation outcomes.
	Metric 14:	Reproducibility of Analyses	Medium	Description sufficient to understand what was done and be conceptually reproducible
	Metric 15:	Statistical Analysis	N/A	statistical models were not applied

Additional Comments: Self-reported symptoms were measured in 60 workers (58 men, 2 women) exposed to phthalic anhydride and other acid anhydrides and compared with measurements in 22 male workers employed in a food-processing factory. Exposure groups included heavy exposure (n=35), low exposure (n=25), and nonexposed (n=22). Phthalic anhydride levels in individual air samples averaged 6.6 mg/m3 (TWA) during phthalic anhydride loading. Rhinitis was reported in 14 (40%) of heavy-exposed subjects and 5 (20%) of low-exposed subjects; rhinoconjunctivitis was seen in 16 (46%) and 3 (12%) in heavy- and low-exposed subjects, respectively; conjunctivitis was seen in 6 (17%) and 5 (20%) in heavy- and low-exposed subjects, respectively; there were no significant differences in the prevalence of atopy based on self-reported history and skinprick test.

Overall Quality Determination — Low

Study Citation:	Sol, C. M., Santos, S., Duijts, L., Asimakopoulos, A. G., Martinez-Moral, M. P., Kannan, K., Jaddoe, V., V.W., Trasande, L. (2020). Fetal phthalates and bisphenols and childhood lipid and glucose metabolism: A population-based prospective cohort study. Environment International 144:106063.
Health Outcome(s) Assessed:	Hepatic/Liver
Reported Health Effect(s):	Total cholesterol, HDL cholesterol, LDL cholesterol, triglycerides
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	6957607

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	High	The study population in this prospective cohort study was a subgroup of mother/child pairs with phthalate and bisphenol concentrations from the Generation R Study in Rotterdam, the Netherlands (n=1,405) from February 2004 and July 2005. Mothers were excluded if they did not have information on phthalate and bisphenol urine concentrations for at least a single time point in pregnancy (n=26). Mother-child pairs were also excluded if singleton children did not participate in postnatal studies at the age of 10 years (n=622). All key information of the study design is reported and described at all steps of the study, and the Generation R cohort is well described in the epidemiological literature.
	Metric 2:	Attrition	Medium	622 out of 1,379 children who were part of the original subsample of mothers with bisphenol and phthalate concentrations were lost to follow-up at 10 years. No reasons for loss to follow-up are reported. These children were thus excluded from further analyses. While attrition may be a concern if the exposure-outcome relationship is different among those lost to follow-up, there is no specific evidence that this is the case. In outcome-specific analyses up to 4 mother-child pairs had missing outcome data, but this is non-significant relative to the large analysis sample of 750+ individuals. Number of individuals are reported at all stages of the study.
	Metric 3:	Comparison Group	High	Key elements of the study design are reported (setting, inclusion and exclusion criteria, and methods of participant selection), and indicate that subjects were similar (recruited from the same eligible population with the same method of ascertainment and within the same time frame using the same inclusion and exclusion criteria, and were of similar age. Maternal and child characteristics were reported. Differences in characteristics of groups were considered as covariates and controlled by statistical analysis. Comparisons between participants and non-participants were also described in the available supplemental document.
Domain 2: Exposure Characterization				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Sol, C. M., Santos, S., Duijts, L., Asimakopoulos, A. G., Martinez-Moral, M. P., Kannan, K., Jaddoe, V., V.W., Trasande, L. (2020). Fetal phthalates and bisphenols and childhood lipid and glucose metabolism: A population-based prospective cohort study. Environment International 144:106063.
Health Outcome(s) Assessed:	Hepatic/Liver
Reported Health Effect(s):	Total cholesterol, HDL cholesterol, LDL cholesterol, triglycerides
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	6957607

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	High	Exposure was consistently assessed using well-established methods (spot urine sample) that measured the chemical concentration of phthalic acid in urine. Exposure was measured in spot urine samples three times during pregnancy, roughly corresponding to the first, second, and third trimesters. Statistical analyses reported both trimester-specific results, and results where exposure was summed across all three time points and divided by three to account for the impact of temporal variability. Analysis was performed at the same lab for all samples using high performance liquid chromatography electrospray ionization-tandem mass spectrometry (HPLC-ESI-MS/MS) as described in Philips et al. 2018 (HERO ID: 4728366). Phthalic acid was "analyzed separately as a proxy for total phthalate exposure". Measured "urine biomarkers for exposure to phthalate metabolites". Concentrations below the LOD were divided by the square root of 2. Concentrations were converted to umol/g creatinine.
	Metric 5:	Exposure Levels	Medium	Urine concentrations were reported for the 1st, 2nd, and 3rd trimesters in the available supplemental document as median, 25th and 75th percentiles. In statistical analyses, exposure was analyzed as a continuous variable as the change in outcome per IQR increase in exposure.
	Metric 6:	Temporality	High	The study presents an appropriate temporality between maternal exposure and outcomes in 10-year old children of mothers. The interval between the exposure and the outcome is appropriate for assessing outcomes from fetal exposure.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Outcomes were assessed using well-established methods; Total cholesterol, HDL cholesterol, triglycerides, glucose, and insulin concentrations from non-fasting venous blood samples were analyzed using Cobas 8000 analyzer (c702 module) for hepatic outcomes and electrochemiluminescence immunoassay (ECLIA) on the E411 module for metabolic outcomes. All outcomes were analyzed at 10 years of age. There are some potential concerns for the use of a non-fasting blood sample, but this is unlikely to be insensitive in a way that is differential by exposure status.
	Metric 8:	Reporting Bias	High	All outcomes were reported. Association of maternal concentrations with outcomes are reported for the 1st, 2nd, and 3rd trimester, and split by sex of children, with confidence interval and/or standard deviation.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	High	Appropriate adjustments or explicit considerations were made for potential confounders (maternal age, ethnicity, pre-pregnancy BMI, folic acid supplementation, education level, parity, smoking habits, alcohol consumption, maternal diet, childhood BMI, child sex) in the final analyses through the use of statistical models to reduce research-specific bias. Creatinine concentration adjustments were also made.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Sol, C. M., Santos, S., Duijts, L., Asimakopoulos, A. G., Martinez-Moral, M. P., Kannan, K., Jaddoe, V., V.W., Trasande, L. (2020). Fetal phthalates and bisphenols and childhood lipid and glucose metabolism: A population-based prospective cohort study. Environment International 144:106063.
Health Outcome(s) Assessed:	Hepatic/Liver
Reported Health Effect(s):	Total cholesterol, HDL cholesterol, LDL cholesterol, triglycerides
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	6957607

Domain	Metric		Rating	Comments
	Metric 10:	Covariate Characterization	Medium	Covariates were assessed using questionnaires during pregnancy and midwife and hospital records at birth. Answers to questionnaire appear to be self-reported and it was not specified if the method was validated. Maternal diet was assessed using a "previously developed food-based diet quality score" based on national dietary guidelines.
	Metric 11:	Co-exposure Counfounding	Medium	Co-exposure to phthalates and bisphenols examined in this study were considered and correlation coefficients between all measured exposures were calculated.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The Study design was appropriate and statistical analysis were appropriate. The study uses appropriate statistical methods to address the research question.
	Metric 13:	Statistical Power	Medium	The number of subjects (n=757) is likely large enough to detect an effect.
	Metric 14:	Reproducibility of Analyses	Medium	The description is sufficient to understand what was done and be conceptually reproducible.
	Metric 15:	Statistical Analysis	High	The models calculating association coefficients are transparent and is stated which variables were included. Model assumptions were adequately described. Basic models are presented in the available supplemental document. Exposure concentrations were log-transformed to address normakity.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	High	Phthalic acid was directly measured in biological media.
	Metric 17:	Effect Biomarker	High	Insulin, glucose, total cholesterol, HDL cholesterol, LDL cholesterol, and triglycerides were all directly measured in biological media.
	Metric 18:	Method Sensitivity	Medium	The limit of detection was specified to be 6.68 nmol/L for phthalic acid, which is likely low enough to detect chemicals in a sufficient percentage of the samples. Limit of detections are not expected to be a concern for the biomarkers of effect.
	Metric 19:	Biomarker Stability	Medium	Urine samples were collected in "polypropylene urine collection containers, stored at 4 °C and transported within 24 h of receipt to the STAR-MDC laboratory before being distributed manually in 25-ml polypropylene vials to be frozen at − 20 °C". It is noted that phthalates have a short biological half-life of less than 24 hours. Philips et al. 2018 (HERO ID: 4728366) noted that the samples had been stored at -20 degrees Celsius for 10 years and suggested that biological activity during the storage period could not be ruled out.
	Metric 20:	Sample Contamination	High	The methodology was reported in the cited Philips et al. 2018 (HERO ID: 4728366) study. Contamination was reported to be monitored by the analysis of procedural blanks for exposure samples. Outcome biomarkers are not mentioned with regards to contamination.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Sol, C. M., Santos, S., Duijts, L., Asimakopoulos, A. G., Martinez-Moral, M. P., Kannan, K., Jaddoe, V., V.W., Trasande, L. (2020). Fetal phthalates and bisphenols and childhood lipid and glucose metabolism: A population-based prospective cohort study. Environment International 144:106063.
Health Outcome(s) Assessed:	Hepatic/Liver
Reported Health Effect(s):	Total cholesterol, HDL cholesterol, LDL cholesterol, triglycerides
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	6957607

Domain	Metric		Rating	Comments
	Metric 21:	Method Requirements	High	High performance liquid chromatography electrospray ionization-tandem mass spectrometry (HPLC-ESI-MS/MS) was used for the detection and measurement of phthalate metabolites, as reported in Philips et al. 2018 (HERO ID: 4728366). Outcome biomarkers were measured on the COBAS 8000 analyzer using the c702 module or using an electrochemiluminescence immunoassay (ELCIA).
	Metric 22:	Matrix Adjustment	Medium	For the biomarker under consideration, study provides results in the main publication and in an available supplemental document for adjusted matrix concentrations (creatinine-adjusted) and reasons are given for adjustment approach.

Additional Comments: Associations between maternal urinary phthalate concentrations (in 1st, 2nd, and 3rd trimesters) and developmental metabolic parameters (serum lipids, glucose, and insulin concentrations) of children (10 years of age) were assessed in 757 mother/child pairs from a subgroup of the Generation R Study in Rotterdam, the Netherlands (2004-2005). There are no overall concerns for bias, due to the use of a prospective cohort design with extensive details on selection, exposure assessment, and outcome assessment. There was an association between maternal 3rd trimester urinary phthalic acid concentrations and increased triglycerides concentration in boys (0.20; 95% CI: 0.07-0.34). No other associations between maternal phthalic acid concentrations and developmental metabolic adaptations were reported.

Overall Quality Determination — High

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Hepatic/Liver
Reported Health Effect(s):	Lactate dehydrogenase (LDH), indirect bilirubin, albumin, gamma-glutamyl transferase (GGT), alkaline phosphatase, direct bilirubin
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Low	In this cross-sectional occupational surveillance study, workers from 9 Koppers Coal Tar plants were examined from January-October 1979. Only one of these plants (the Chicago plant) produced phthalic anhydride. Participation at the relevant phthalic anhydride plant (i.e., the Chicago plant) was approximately 46% (105/230), while participation from the total workforce across all plants was 51% (453/888). The study authors note that participation at the Chicago plant was non-representative for hourly workers relative to salaried workers. The study does not discuss recruitment or selection processes in detail but does indicate that workers could participate in the study during normal work hours without impact on their pay. Overall, details are sparse and it is unclear if those who did not participate may have been more or less exposed compared to those who were included.
	Metric 2:	Attrition	High	One participant out of 105 in the Chicago did not have a blood test taken and was excluded from analyses that included serum measures. No reason is provided for why the blood test was not taken, but there was no other indication of attrition or exclusion in the Chicago plant group or in the total included group.
	Metric 3:	Comparison Group	Low	Demographic details on sex and race are provided for all plants. The distribution of race and sex is roughly similar in the Chicago plant to the overall distribution of race and sex across the included workforce. However, these demographic differences are not controlled for in statistical analyses. There is also no discussion of age and whether groups had differences based on age.
Domain 2: Exposure Characterization				
	Metric 4:	Measurement of Exposure	Low	In comparisons between different plants, all Chicago plant workers were grouped together. Descriptions of the Chicago plant were very limited beyond stating that phthalic anhydride was used. A sub-group analysis of the Chicago plant separates out workers by ”type” (Phthalic, Maleic, or Coal Tar). It is unclear if phthalic anhydride workers were all actually formally exposed to phthalic anhydride, and it is not clear whether the other types of workers were not exposed. There is a large potential for exposure misclassification if not all tasks or employees were in contact with phthalic anhydride, and since exposure was estimated solely using professional judgment, exposure misclassification cannot be ruled out.
	Metric 5:	Exposure Levels	Low	The study only reports two levels of exposure, exposed and unexposed. No quantitative exposure information is available.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Hepatic/Liver
Reported Health Effect(s):	Lactate dehydrogenase (LDH), indirect bilirubin, albumin, gamma-glutamyl transferase (GGT), alkaline phosphatase, direct bilirubin
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
	Metric 6:	Temporality	Low	Work history or duration of employment was not discussed in the analysis. It is clear that the outcomes were measured after some degree of exposure, since outcomes were assessed in an occupational context. However, it is unclear how long these workers may have been exposed, and it is unclear whether the employees were free of any of the reported health outcomes prior to enrollment into the study. The temporality of exposure and outcome is uncertain.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Specific methods were not described, but description of the study suggests standard clinical methods were used (in line with NIOSH recommendations). Examination forms were provided in the appendices.
	Metric 8:	Reporting Bias	Medium	All outlined outcomes are reported in the results, however, data is provided as the proportion of individuals with ”abnormal” measurements compared against clinical standards.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Low	There was no evidence of adjustment for potential confounders, however, some were discussed in-text for medical interpretations. The distribution of sex and race is provided across exposure group, but there is no discussion of age. Covariates were not adjusted for or stratified in relation to outcomes.
	Metric 10:	Covariate Characterization	Medium	Questionnaires were provided in the appendix. It was not reported whether this was a validated questionnaire, however, there was no evidence to suggest it was an invalid instrument.
	Metric 11:	Co-exposure Counfounding	Low	Exposure to other coal tar components was described, including potential health effects. Workers in plants other than the Chicago plant were exposed to other occupational agents, and it is unclear whether those exposures contributed to the incidence of health outcomes. Within the Chicago plant, the study splits the sample into workers exposed to phthalic anhydride, maleic anhydride, or coal tar. It is not confirmed that phthalic anhydride workers were only exposed to phthalic anhydride.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	The design limited the ability to determine which exposures were relevant to each health effect, however the design is sufficient to answer the study’s question as to whether there are ”abnormal” medical findings among workers at the studied coal tar plants.
	Metric 13:	Statistical Power	Low	Statistical power was not reported. A total of 105 employees completed exams from the Chicago plant. However, the analysis within the Chicago plant only identifies 14 workers who were exposed to phthalic anhydride, which may be too small of a sample size to detect an effect.

Continued on next page . . .

. . . continued from previous page

Study Citation:	TOMA, (1981). 1979 Cross-sectional health study of workers at nine Koppers coal tar plants combined report.
Health Outcome(s) Assessed:	Hepatic/Liver
Reported Health Effect(s):	Lactate dehydrogenase (LDH), indirect bilirubin, albumin, gamma-glutamyl transferase (GGT), alkaline phosphatase, direct bilirubin
Chemical:	Phthalic anhydride- Parent compound
HERO ID:	5299399 Linked HERO ID(s): 5299399, 1481371

Domain	Metric		Rating	Comments
	Metric 14:	Reproducibility of Analyses	Medium	The description of the analysis is sufficient to understand precisely what has been done and to be conceptually reproducible with access to the analytic data.
	Metric 15:	Statistical Analysis	N/A	Not applicable, no formal statistical analysis was conducted.
Domain 6: Other (if applicable) Considerations for Biomarker Selection and Measurement (Lakind et al. 2014)				
	Metric 16:	Use of Biomarker of Exposure	N/A	Not applicable, no biomarkers of exposure were measured.
	Metric 17:	Effect Biomarker	High	All studied effect biomarkers were demonstrated to be related to adverse health outcomes and were collected from serum or urine samples.
	Metric 18:	Method Sensitivity	N/A	Limits of detection not relevant for the medical testing employed.
	Metric 19:	Biomarker Stability	Low	There is no description of the storage history and/or stability data for urine or serum samples.
	Metric 20:	Sample Contamination	Medium	There is no discussion of contamination.
	Metric 21:	Method Requirements	Medium	While no formal descriptions are provided, given the medical context of the examinations there is some confidence that accurate detection methodologies were used.
	Metric 22:	Matrix Adjustment	Medium	No discussion of matrices is described. This would be relevant for effect biomarkers measured in urine, but there is no evidence that there was no adjustment for creatinine.

Additional Comments: This occupational health surveillance study focused on several coal tar facilities in the United States. At the Chicago plant phthalic anhydride was present. The aim of the study was to compare medical findings across a wide range of health outcomes to established clinical ranges, thus no formal statistical analysis was available that compared exposed vs. unexposed. An examination of the reported data does not indicate that there were significant differences between workers exposed to phthalic anhydride in the Chicago plant and the overall population of workers, or workers at the Chicago plant not reported to be exposed to phthalic anhydride; however, this cannot be completely determined without a formal statistical analysis. There were several large concerns raised with the study, including the lack of an exposure measurement. All employees in the phthalic anhydride plant were considered exposed, however, without surveillance data or employment records, there is a large potential for exposure misclassification. There were also sparse details regarding recruitment, and a lack of consideration of potentially relevant covariates such as age. The workers in this study were all exposed to other occupational agents, thus confounding by other exposures cannot be ruled out.

Overall Quality Determination Low

Study Citation:	Barker, R. D., Tongeren, van, M. J., Harris, J. M., Gardiner, K., Venables, K. M., Taylor, Newman, A. J. (1998). Risk factors for sensitisation and respiratory symptoms among workers exposed to acid anhydrides: A cohort study. Occupational and Environmental Medicine 55(10):684-691.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Respiratory symptoms
Chemical:	Phthalic anhydride- Mixture: Acid anhydrides (phthalic anhydride, maleic anhydride, trimellitic anhyride)
HERO ID:	831017

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	Participants in this occupational cohort study were selected from four UK factories where acid anhydrides were used. Potential participants were identified with factory personnel and medical records. Participants were eligible if they had started to work in an area where acid anhydrides were used and continued to work there for more than a month after 1 January 1960 or 1 January 1979, depending on the factory. The target population consisted of 506 workers of which 401 completed questionnaires and 378 completed skin prick testing. Exclusion criteria were death (n=26), refusal to take part (n=35), and unable to trace (n=44). Selection criteria were stated. The study does not present a comparison of exposure, outcome, or demographic characteristics of those included vs. those excluded from the study, but there is no evidence to suggest that exclusion was related to exposure or outcome status.
	Metric 2:	Attrition	Medium	23 participants who completed questionnaires did not complete skin prick testing due to ”difficulties arranging a meeting” and one person refused to complete the testing due to known severe reactions to ”common inhalant aeroallergens.” These individuals were removed from the analysis of phthalic acid. There appears to be some additional participants excluded from the analyses, due to an expected analytic sample size of 378 and some analyses only including 374 or 366 workers. While detailed explanation is not provided for why measurements were unavailable, there is no evidence that this missingness would be related to both outcome and exposure, and the missingness represents a small proportion of the total sample. It is possible exclusions were made for participants who had job titles where an exposure value could not be applied, as analyses comparing full shift exposure had the lowest sample sizes.
	Metric 3:	Comparison Group	Low	Study authors note that age range was 18-81. They add that workers with respiratory symptoms were older and all sensitized or reported respiratory symptoms were men. Age and sex are not adjusted for in the analyses despite being different across the participants. No further demographic information was provided. The study also performs a case-control analysis, where those with respiratory symptoms were considered to be cases. Cases were stated to be matched with four controls who worked in the same factory at the same date as the cases’ symptoms began, with no adjustment for age.
Domain 2: Exposure Characterization				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Barker, R. D., Tongeren, van, M. J., Harris, J. M., Gardiner, K., Venables, K. M., Taylor, Newman, A. J. (1998). Risk factors for sensitisation and respiratory symptoms among workers exposed to acid anhydrides: A cohort study. Occupational and Environmental Medicine 55(10):684-691.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Respiratory symptoms
Chemical:	Phthalic anhydride- Mixture: Acid anhydrides (phthalic anhydride, maleic anhydride, trimellitic anhyride)
HERO ID:	831017

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	Medium	Phthalic anhydride exposure was measured by job group where participants were asked to wear an air sampling pump for the duration of a full workshift and during specific tasks. Workers were grouped by job title and workers were randomly selected to have their workshift exposure levels measured. Some exposure levels were established retrospectively based on past records and estimates. A job-time-exposure matrix used to identify full shift exposure. For each participant, the highest mean full shift exposure value linked with their job title was used in analyses. Further exposure assessment details are available in HEROID 831008 and 831018.
	Metric 5:	Exposure Levels	Medium	The exposure range for all acid anhydrides is likely to be sufficiently large, where full shift exposure to acid anhydride ranged from less than 10 ug/m^3 to greater than 100 ug/m^3. Three separate levels of exposure were used in statistical analyses (<10, 10-<100, and >=100 ug/m^3).
	Metric 6:	Temporality	High	Health outcomes were measured roughly at similar time points to exposure estimates. Health outcomes excluded those who reported that symptoms occurred before starting work at the factory. While the exposure assessment and health outcomes were measured roughly at the same point in time, workers were required to work at the factory for more than a month. There is thus high confidence that exposure preceded the appearance of outcomes. There are no concerns for temporality with sensitization outcomes, as exposure is required to have occurred before the outcome can appear.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Self-administered questionnaires were used to assess the presence of respiratory symptoms. While there is no discussion of validation, the questions are all reported in the study. Participants were stated to have respiratory symptoms if they experienced chest tightness, difficulty breathing, or wheeze/whistling in the chest.A subset of participants were subjected to skin-prick testing with common allergens (cat fur, mixed grass pollen, dematophagiodes pteronyssinus, and human serum albumin conjugated to acid anhydrides. A diagnostic definition for a positive skin-prick test for skin sensitization is provided.
	Metric 8:	Reporting Bias	High	Measured outcomes are reported in tables and discussed in the text. Effect estimates are reported with 95% confidence intervals. In case-control analyses, case and control numbers are reported.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Low	Adjustments were made for some potential confounders (smoking and atopy). Despite differences in age and sex in those with and without symptoms, authors did not adjust for these measures. No framework or justification is provided explaining why the specific covariates were chosen.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Barker, R. D., Tongeren, van, M. J., Harris, J. M., Gardiner, K., Venables, K. M., Taylor, Newman, A. J. (1998). Risk factors for sensitisation and respiratory symptoms among workers exposed to acid anhydrides: A cohort study. Occupational and Environmental Medicine 55(10):684-691.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Respiratory symptoms
Chemical:	Phthalic anhydride- Mixture: Acid anhydrides (phthalic anhydride, maleic anhydride, trimellitic anhyride)
HERO ID:	831017

Domain	Metric		Rating	Comments
	Metric 10:	Covariate Characterization	Medium	Covariate information was collected using self-administered questionnaires. While no validation is discussed, there is no evidence to suggest the questionnaire was inappropriate.
	Metric 11:	Co-exposure Counfounding	Low	The primary analyses used a measure of ”acid anhydrides”. No analysis only considered a single exposure. It is unclear exactly how the ”acid anhydrides” measure is used, but the description of the analysis implies that the full shift exposure value was assigned as the highest value of either PA, MA, or TMA. There were different provisions of acid anhydrides at different factories, with Factory 2 only reporting TMA and Factory 4 not reporting MA. Thus, there was an unbalanced provision of co-exposures across study groups, and these were not accounted for by only evaluating the highest exposure in each participant.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	Prevalence and severity of symptoms were analyzed in an occupational cohort and logistic regressions and chi squared tests were used to investigate the relation between exposure and outcomes.
	Metric 13:	Statistical Power	Medium	The number of participants (n=401) is likely sufficiently large to guarantee a high enough statistical power. The subgroups that were analyzed were also likely to be sufficiently large.
	Metric 14:	Reproducibility of Analyses	Low	The description of the analysis is somewhat sufficient to understand what has been done and to be conceptually reproducible with access to the analytic data. The study’s exact calculation of ”exposure to acid anhydride” is unclear, although it can be inferred that the study likely assigned participants exposure level to be either the value for PA, MA, or TMA (whichever one is the highest). There are discrepancies between the final analytic sample size and they actual number of participants that appear in the analysis.
	Metric 15:	Statistical Analysis	High	Statistical methods for risk estimates are clearly described and appropriate using ORs and 95% CIs. Logistic regression models were created to assess relationships between actual exposure to acid anhydrides and sensitization, as well as respiratory symptoms. Model assumptions are not likely to be violated.

Additional Comments: This occupational cohort examined numerous employees from multiple factories to examine the association between acid anhydride exposure and health outcomes, including respiratory symptoms and sensitization. The key limitations include not adjusting for key covariates, the lack of clarity regarding how exactly the exposure value of ”acid anhydrides” was determined, and inconsistent numbers of participants used across analyses. However, there are minimal concerns with the overall exposure assessment, participant recruitment, and outcome assessment. The utility of this study for a hazard determination for phthalic anhydride is limited by the use of the ”acid anhydride” exposure in statistical analysis, which cannot be traced specifically back to phthalic anhyride.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Barker, R. D., Tongeren, van, M. J., Harris, J. M., Gardiner, K., Venables, K. M., Taylor, Newman, A. J. (1998). Risk factors for sensitisation and respiratory symptoms among workers exposed to acid anhydrides: A cohort study. Occupational and Environmental Medicine 55(10):684-691.
Health Outcome(s) Assessed:	Lung/Respiratory
Reported Health Effect(s):	Respiratory symptoms
Chemical:	Phthalic anhydride- Mixture: Acid anhydrides (phthalic anhydride, maleic anhydride, trimellitic anhyride)
HERO ID:	831017

Domain	Metric	Rating	Comments
Overall Quality Determination		**Medium**	

Study Citation:	Barker, R. D., Tongeren, van, M. J., Harris, J. M., Gardiner, K., Venables, K. M., Taylor, Newman, A. J. (1998). Risk factors for sensitisation and respiratory symptoms among workers exposed to acid anhydrides: A cohort study. Occupational and Environmental Medicine 55(10):684-691.
Health Outcome(s) Assessed:	Sensitization
Reported Health Effect(s):	Skin sensitization
Chemical:	Phthalic anhydride- Mixture: Acid anhydrides (phthalic anhydride, maleic anhydride, trimellitic anhyride)
HERO ID:	831017

Domain	Metric		Rating	Comments
Domain 1: Study Participation				
	Metric 1:	Participant Selection	Medium	Participants in this occupational cohort study were selected from four UK factories where acid anhydrides were used. Potential participants were identified with factory personnel and medical records. Participants were eligible if they had started to work in an area where acid anhydrides were used and continued to work there for more than a month after 1 January 1960 or 1 January 1979, depending on the factory. The target population consisted of 506 workers of which 401 completed questionnaires and 378 completed skin prick testing. Exclusion criteria were death (n=26), refusal to take part (n=35), and unable to trace (n=44). Selection criteria were stated. The study does not present a comparison of exposure, outcome, or demographic characteristics of those included vs. those excluded from the study, but there is no evidence to suggest that exclusion was related to exposure or outcome status.
	Metric 2:	Attrition	Medium	23 participants who completed questionnaires did not complete skin prick testing due to ”difficulties arranging a meeting” and one person refused to complete the testing due to known severe reactions to ”common inhalant aeroallergens.” These individuals were removed from the analysis of phthalic acid. There appears to be some additional participants excluded from the analyses, due to an expected analytic sample size of 378 and some analyses only including 374 or 366 workers. While detailed explanation is not provided for why measurements were unavailable, there is no evidence that this missingness would be related to both outcome and exposure, and the missingness represents a small proportion of the total sample. It is possible exclusions were made for participants who had job titles where an exposure value could not be applied, as analyses comparing full shift exposure had the lowest sample sizes.
	Metric 3:	Comparison Group	Low	Study authors note that age range was 18-81. They add that workers with respiratory symptoms were older and all sensitized or reported respiratory symptoms were men. Age and sex are not adjusted for in the analyses despite being different across the participants. No further demographic information was provided. The study also performs a case-control analysis, where those with respiratory symptoms were considered to be cases. Cases were stated to be matched with four controls who worked in the same factory at the same date as the cases’ symptoms began, with no adjustment for age.
Domain 2: Exposure Characterization				

Continued on next page . . .

. . . continued from previous page

Study Citation:	Barker, R. D., Tongeren, van, M. J., Harris, J. M., Gardiner, K., Venables, K. M., Taylor, Newman, A. J. (1998). Risk factors for sensitisation and respiratory symptoms among workers exposed to acid anhydrides: A cohort study. Occupational and Environmental Medicine 55(10):684-691.
Health Outcome(s) Assessed:	Sensitization
Reported Health Effect(s):	Skin sensitization
Chemical:	Phthalic anhydride- Mixture: Acid anhydrides (phthalic anhydride, maleic anhydride, trimellitic anhyride)
HERO ID:	831017

Domain	Metric		Rating	Comments
	Metric 4:	Measurement of Exposure	Medium	Phthalic anhydride exposure was measured by job group where participants were asked to wear an air sampling pump for the duration of a full workshift and during specific tasks. Workers were grouped by job title and workers were randomly selected to have their workshift exposure levels measured. Some exposure levels were established retrospectively based on past records and estimates. A job-time-exposure matrix used to identify full shift exposure. For each participant, the highest mean full shift exposure value linked with their job title was used in analyses. Further exposure assessment details are available in HEROID 831008 and 831018.
	Metric 5:	Exposure Levels	Medium	The exposure range for all acid anhydrides is likely to be sufficiently large, where full shift exposure to acid anhydride ranged from less than 10 ug/m^3 to greater than 100 ug/m^3. Three separate levels of exposure were used in statistical analyses (<10, 10-<100, and >=100 ug/m^3).
	Metric 6:	Temporality	High	Health outcomes were measured roughly at similar time points to exposure estimates. Health outcomes excluded those who reported that symptoms occurred before starting work at the factory. While the exposure assessment and health outcomes were measured roughly at the same point in time, workers were required to work at the factory for more than a month. There is thus high confidence that exposure preceded the appearance of outcomes. There are no concerns for temporality with sensitization outcomes, as exposure is required to have occurred before the outcome can appear.
Domain 3: Outcome Assessment				
	Metric 7:	Outcome Measurement or Characterization	Medium	Self-administered questionnaires were used to assess the presence of respiratory symptoms. While there is no discussion of validation, the questions are all reported in the study. Participants were stated to have respiratory symptoms if they experienced chest tightness, difficulty breathing, or wheeze/whistling in the chest.A subset of participants were subjected to skin-prick testing with common allergens (cat fur, mixed grass pollen, dematophagiodes pteronyssinus, and human serum albumin conjugated to acid anhydrides. A diagnostic definition for a positive skin-prick test for skin sensitization is provided.
	Metric 8:	Reporting Bias	High	Measured outcomes are reported in tables and discussed in the text. Effect estimates are reported with 95% confidence intervals. In case-control analyses, case and control numbers are reported.
Domain 4: Potential Confounding / Variability Control				
	Metric 9:	Covariate Adjustment	Low	Adjustments were made for some potential confounders (smoking and atopy). Despite differences in age and sex in those with and without symptoms, authors did not adjust for these measures. No framework or justification is provided explaining why the specific covariates were chosen.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Barker, R. D., Tongeren, van, M. J., Harris, J. M., Gardiner, K., Venables, K. M., Taylor, Newman, A. J. (1998). Risk factors for sensitisation and respiratory symptoms among workers exposed to acid anhydrides: A cohort study. Occupational and Environmental Medicine 55(10):684-691.
Health Outcome(s) Assessed:	Sensitization
Reported Health Effect(s):	Skin sensitization
Chemical:	Phthalic anhydride- Mixture: Acid anhydrides (phthalic anhydride, maleic anhydride, trimellitic anhyride)
HERO ID:	831017

Domain	Metric		Rating	Comments
	Metric 10:	Covariate Characterization	Medium	Covariate information was collected using self-administered questionnaires. While no validation is discussed, there is no evidence to suggest the questionnaire was inappropriate.
	Metric 11:	Co-exposure Counfounding	Low	The primary analyses used a measure of ”acid anhydrides”. No analysis only considered a single exposure. It is unclear exactly how the ”acid anhydrides” measure is used, but the description of the analysis implies that the full shift exposure value was assigned as the highest value of either PA, MA, or TMA. There were different provisions of acid anhydrides at different factories, with Factory 2 only reporting TMA and Factory 4 not reporting MA. Thus, there was an unbalanced provision of co-exposures across study groups, and these were not accounted for by only evaluating the highest exposure in each participant.
Domain 5: Analysis				
	Metric 12:	Study Design and Methods	High	Prevalence and severity of symptoms were analyzed in an occupational cohort and logistic regressions and chi squared tests were used to investigate the relation between exposure and outcomes.
	Metric 13:	Statistical Power	Medium	The number of participants (n=401) is likely sufficiently large to guarantee a high enough statistical power. The subgroups that were analyzed were also likely to be sufficiently large.
	Metric 14:	Reproducibility of Analyses	Low	The description of the analysis is somewhat sufficient to understand what has been done and to be conceptually reproducible with access to the analytic data. The study’s exact calculation of ”exposure to acid anhydride” is unclear, although it can be inferred that the study likely assigned participants exposure level to be either the value for PA, MA, or TMA (whichever one is the highest). There are discrepancies between the final analytic sample size and they actual number of participants that appear in the analysis.
	Metric 15:	Statistical Analysis	High	Statistical methods for risk estimates are clearly described and appropriate using ORs and 95% CIs. Logistic regression models were created to assess relationships between actual exposure to acid anhydrides and sensitization, as well as respiratory symptoms. Model assumptions are not likely to be violated.

Additional Comments: This occupational cohort examined numerous employees from multiple factories to examine the association between acid anhydride exposure and health outcomes, including respiratory symptoms and sensitization. The key limitations include not adjusting for key covariates, the lack of clarity regarding how exactly the exposure value of ”acid anhydrides” was determined, and inconsistent numbers of participants used across analyses. However, there are minimal concerns with the overall exposure assessment, participant recruitment, and outcome assessment. The utility of this study for a hazard determination for phthalic anhydride is limited by the use of the ”acid anhydride” exposure in statistical analysis, which cannot be traced specifically back to phthalic anhyride.

Continued on next page . . .

. . . continued from previous page

Study Citation:	Barker, R. D., Tongeren, van, M. J., Harris, J. M., Gardiner, K., Venables, K. M., Taylor, Newman, A. J. (1998). Risk factors for sensitisation and respiratory symptoms among workers exposed to acid anhydrides: A cohort study. Occupational and Environmental Medicine 55(10):684-691.
Health Outcome(s) Assessed:	Sensitization
Reported Health Effect(s):	Skin sensitization
Chemical:	Phthalic anhydride- Mixture: Acid anhydrides (phthalic anhydride, maleic anhydride, trimellitic anhyride)
HERO ID:	831017

Domain	Metric	Rating	Comments
Overall Quality Determination		**Medium**	