C T F A

THE COSMETIC, TOILETRY, AND FRAGRANCE ASSOCIATION

November 30, 2000

DEC 1 2000

E. EDWARD KAVANAUGH
PRESIDENT

Dr. Mary S. Wolfe
Executive Secretary
National Toxicology Program
P.O. Box 12233, A3-07
Research Triangle Park, NC 27709

RE: Review of Nominations for Listing in the 10th *Report on Carcinogens* (65 Federal Register 65352): Non-Asbestiform Talc

Dear Dr. Wolfe,

The Cosmetic, Toiletry, and Fragrance Association[1] (CTFA) appreciates the opportunity to provide comments on the proposed listing of non-asbestiform talc in the 10th Report on Carcinogens. Non-asbestiform talc is used within the personal care products industry, and thus the review for listing is of significant interest to CTFA members. The basis of the nomination is addressed in the documents included in this submission, which show conclusively that the listing of non-asbestiform talc is not scientifically justified.

Enclosed are comments on the proposal prepared by the following individuals:

- Dr. Donald Dungworth, Professor of Veterinary Pathology, Emeritus, University of California, Davis; the emphasis of Dr. Dungworth's comments is on the animal inhalation studies. Contact information for Dr. Dungworth can be found with his submission.

- Dr. Jay Goodman, Professor of Pharmacology and Toxicology, Michigan State University; Dr. Goodman was a member of NTP's Board of Scientific Counselors reviewing the NTP chronic bioassay with non-asbestiform talc, and his comments address that study. Contact information for Dr. Goodman can be found with his submission.

[1] CTFA is the U.S. national trade association representing the personal care products industry. CTFA is comprised of over 300 active members that produce the vast majority of the cosmetics distributed in the U.S. and that also produce many over-the-counter drugs designed for dermal application. The association also has over 300 associate members that provide raw ingredients and suppliers and services to the industry. Many of CTFA's members are international companies that do business in foreign countries as well.

1101 17TH ST., N.W., SUITE 300 WASHINGTON, D.C. 20036-4702
202.331.1770 FAX 202.331.1969
http://www.ctfa.org

- Dr. Kenneth Rothman, Professor, Department of Epidemiology and Medicine, Boston University; Dr. Harris Pastides, Dean, School of Public Health, University of South Carolina; and Dr. Jonathan Samet, Chairman, Department of Epidemiology, Johns Hopkins University; their review is of the epidemiology studies relating to talc and ovarian cancer. The contact information for the authors of this document is as follows: Kenneth Rothman, Dr.P.H., MPH, Boston University, 715 Albany Street B6, Boston, MA 02215, 617-638-8089, krothman@bu.edu; Harris Pastides, Ph.D., MPH, School of Public Health, 109 Health Sciences Building, University of South Carolina, Columbia, SC 29208, 803-777-5032 (phone), 803-777-4783 (fax), hpastides@sph.sc.edu; Jonathan Samet, M.D., M.S., Department of Epidemiology, Johns Hopkins School of Hygiene and Public Health, 615 North Wolfe Street, Suite 6039, Baltimore, MD 21205, 410-955-3286 (phone), 410-955-0863 (fax), jsamet@jhsph.edu.

- Dr. Samuel Shapiro, MB, FRCP(E), Emeritus Director, Slone Epidemiology Unit, Boston University School of Public Health, and Visiting Professor of Epidemiology, Columbia University School of Public Health; review of the epidemiology studies relating to talc and ovarian cancer. Contact information for Dr. Shapiro is Division of Epidemiology, The Joseph L. Mailman School of Public Health, Columbia University, 600 W. 168th Street, PH18, New York, New York, 10032.

- Joshua Muscat, MPH, Research Scientist, American Health Foundation; review of the epidemiology studies relating to talc and ovarian cancer. Contact information is American Health Foundation, 1 Dana Road, Valhalla, New York, 10595, 914-789-7353 (phone); 914-592-6317 (fax); jmuscat2@earthlink.net.

CTFA appreciates the opportunity to submit information on the proposed listing.

Sincerely,

[redacted]

Gerald McEwen, Jr., Ph.D., J.D.
Vice President-Science

Enclosures

D.L.DUNGWORTH, B.V.SC., PH.D., M.R.C.V.S., DIP.A.C.V.P.
PROFESSOR OF VETERINARY PATHOLOGY, EMERITUS
UNIVERSITY OF CALIFORNIA, DAVIS

6260 CAPE GEORGE ROAD
PORT TOWNSEND, WA 98368

FAX 360-385-4966
PHONE 360-385-5006
E-MAIL DLD1@OLYPEN.COM

November 29, 2000

Board of Scientific Counselors
National Toxicology Program
National Institute of Environmental Health Sciences
P.O. Box 12233
Research Triangle Park, NC 27709

Dear Board Members,

My name is Dr. Donald L. Dungworth and I am an Emeritus Professor of Veterinary Pathology at the University of California at Davis. My primary research area has been the study of Pulmonary Pathology/Toxicology.

The enclosed document was prepared in response to the NTP proposal to list non-asbestiform talc in the 10th Report on Carcinogens.

Thank you for consideration of these comments. Please feel free to contact me if additional information is needed.

Sincerely,

[redacted]

Donald L. Dungworth
Emeritus Professor

ANALYSIS OF NTP DRAFT REPORT ON TALC, ASBESTIFORM AND NON-ASBESTIFORM

D. L. Dungworth

Emeritus Professor of Veterinary Pathology

University of California at Davis

A. Executive Summary

Based on the analysis of findings concerning non-asbestiform talc, the following conclusions are justified:

- Rigorous evaluation of the animal studies does not support the conclusion that, based on animal studies, non-asbestiform talc can reasonably be anticipated to be a human carcinogen because neither the requirement for activity in multiple species, nor multiple sites, nor multiple routes has been satisfied adequately.

- Non-asbestiform talc induces pulmonary tumors in female rats at considerable lung burdens and when associated with persistent, severe, chronic active inflammation and accompanying epithelial proliferation.

- The carcinogenic effect in the rat lung has a threshold related to the precondition for considerable particle-induced inflammation and epithelial proliferation.

- Even if a mechanism for tumor induction by poorly soluble, non-genotoxic particles were to exist in human lungs analogous to that in rats (and evidence thus far runs counter to this), tumors would arise at exposure concentration levels well above those that cause clinically relevant pneumoconiosis. The NTP draft document (pp. 71/72) contains a statement agreeing with this conclusion.

- The statistically significant (P< 0.01) increase in adrenal pheochromocytomas in male F344 rats, which have a very high incidence in controls, does not provide biologically plausible evidence for a true carcinogenic effect given the susceptibility of chromaffin cells in the F344 rat to neural, humoral and growth factor stimuli.

B. Introduction

This analysis will be limited to experimental findings with non-asbestiform talc, and the emphasis will be on animal inhalation studies

Evidence will be reviewed according to requirements for a substance to be reasonably anticipated to be a human carcinogen from studies in experimental animals i.e. an increased incidence of malignant and/or a combination of malignant and benign tumors: (1) in multiple species or at multiple sites, or (2) by multiple routes of exposure, or (3) to an unusual degree with regard to incidence, site or type of tumor or age at onset. Furthermore, the conclusions regarding carcinogenicity are to be based on scientific judgement, with consideration given to all relevant information.

The requirement for causing tumors in multiple species or by multiple exposure routes has not been met by non-asbestiform talc, nor has that for an unusual degree regarding incidence, site or type of tumor or age at onset. The conclusions with regard to non-asbestiform talc, therefore, must be based on interpretation of the chronic inhalation study in which high lung burdens of talc in rats, but not mice, were associated with an increase in primary lung tumors in female rats. The frequency of benign and malignant pheochromocytomas also increased 20 to 27% from a high control level in both male and female rats that also had evidence of impairments in pulmonary function. Analysis of the NTP experimental inhalation study, other relevant *in vivo* studies and *in vitro* studies follows.

C. Inhalation Studies (NTP 1993 and Wehner et al. 1977)

1. Pulmonary Carcinogenesis Endpoint

Female F344 rats exposed for up to 122 weeks (6 hours/day, 5 days/week) to 0, 6 or 18 mg/m^3 air concentration of non-asbestiform talc developed a significant increase in bronchiolo-alveolar adenomas or carcinomas at the 18 mg/m^3 level. Fourteen of 50 (28%) of exposed females had tumors compared to 1 of 50 (2%) of controls. No tumors were observed in female rats exposed to 6 mg/m^3 air concentration of talc. Male rats did not have a significant increase in lung tumors at either 6 or 18 mg/m^3. Failure to detect tumors in male rats was probably related mainly to their shorter life span (113 weeks in males compared to 122 weeks in females), although a lesser degree of inflammation in male rats (NTP 1993) might also have played some part. Pulmonary tumors induced in rats by various particles (carbon black, titanium oxide and diesel exhaust) are mostly found after 25 to 30 months on study (Dungworth et al. 1994, Nikula 2000). This is especially true for bronchiolo-alveolar carcinomas. The largest number of tumors is typically found in rats terminated at the end of a 30-month study. Male rats surviving only up to 26 months (113 weeks), as in the talc study, are therefore unlikely to reveal a significant increase in tumors, especially if the stimulus is weak. Where the survival of male rats is sufficiently long, for example as reported for a diesel exhaust study (Mauderley et al. 1987), both male and female rats have significant increases in lung tumors over controls.

In the same NTP study, male and female B6C3F1 mice exposed to 0, 6 or 18 mg/m^3 of non-asbestiform talc for up to 104 weeks, using the same regimen as in rats, did not have any increases in frequency of lung tumors. It is worth noting here that it has been estimated that the rats and mice in the NTP study were exposed to between 2,000 and 20,000 times the levels of airborne talc compared to humans using the product normally (Zazenski et al. 1995).

With regard to non-neoplastic pulmonary changes, both rats and mice in the NTP study developed lesions associated with accumulation of particles (see below), but they were more severe and complex in the rats.

In a follow up report to the NTP (1993) document, Boorman and Seely (1995) reported that the lifetime inhalation exposure of the rats and mice caused no ovarian lesions and no talc particles were detected in the ovaries of a 20% random sampling of exposed female rats using polarized light.

Apart from the NTP study, the hamster is the third species that has been exposed to airborne talc (baby-powder) (Wehner et al. 1977). The calculated respirable fraction of talc was around 8 mg/m^3 and exposure ranged up to 150 minutes/day, five days/week for up to 300 days, unless hamsters died sooner. Surviving hamsters were kept to the end of their natural life span. Evaluation of pulmonary changes was confounded by a variety of incidental lesions. No significant neoplastic or non-neoplastic changes were found over the course of the study.

2. Significance of Rat Lung Carcinogenesis

In the NTP study, female rats exposed to levels of non-asbestiform talc causing a high lung burden developed pulmonary tumors. Non-asbestiform talc is not a pulmonary carcinogen for mice. The positive pulmonary carcinogenesis by high lung burdens of non-asbestiform talc in rats places it together with other low solubility, low toxigenic and non-genotoxic particles that have the same effect, i.e. carbon black and titanium dioxide. These particles have been referred to as "nuisance dusts", particles not otherwise classified (PNOC - ACGIH 1994), or poorly soluble particulate materials (PSPs - ILSI, 2000). This type of particle does not cause tumors in mice (NTP 1993, Heinrich et al. 1995, Nikula 2000). The high lung burdens in rats are associated with chronic active inflammation, epithelial hyperplasia/metaplasia, and increased tumor development after 24 to

30 months. The high particle burden associated with the inflammatory and proliferative events in rats has been referred to as "overload" by those studying kinetics of particle deposition, clearance and retention (Oberdorster 1995). This has led to the general concept that there is a link between high particle burden (overload), persistent inflammation, epithelial hyperplasia and neoplasia which, if not specific to the rat, occurs to a much greater extent in the rat than in any of the other species examined. The most recent consideration of the relevance of response to particle "overload" in the rat lung to human risk assessment was an ILSI Risk Science Institute Workshop (ILSI 2000). With respect to interpretation of the carcinogenic response of the rat lung to non-asbestiform talc (and similar particles) there are several questions:

a. What is known of the mechanisms of particle-induced carcinogenesis in the rat lung and does a threshold exist?
b. What is the sensitivity of the rat to particle-induced carcinogenesis compared to other species?
c. What is the probable relevance for humans of the particle-induced tumors in rats?

These three questions will be addressed in turn.

(a) <u>What is known of the mechanisms of particle-induced carcinogenesis in the rat lung and does a threshold exist?</u>

Pulmonary tumors in rats arise as a late event out of a background of persistent particle-induced inflammation. This chronic active inflammation is associated with epithelial hyperplasia and metaplasia, as well as other hallmarks of chronic active inflammation, including fibrosis (Dungworth et al. 1994). The concept that persistent inflammation and epithelial hyperplasia/metaplasia provide the initial steps in the carcinogenic process is generally accepted (Oberdorster 1995, Driscoll 1996, ILSI 2000). The key event in this sequence is the severity of

particle-induced inflammation. The rat is much more sensitive to particle induced inflammation than mice, hamsters, or monkeys (Heinrich et al. 1986, Dungworth et al. 1994, Donaldson 1999, Nikula 2000). This is currently believed to play a major role in the failure of mice and hamsters to develop lung tumors when exposed analogously to rats.

As just stated, the key initiating event in the rat is the particle-induced inflammation which, when persistent and severe at high lung burdens, can eventually give rise to tumors. Particle "overload" can act as a surrogate for the chain of events in the rat because the degree of overload is correlated with the severity of inflammation. Overload as determined by particle kinetics, can occur in mice and hamsters (Oberdorster 1995, ILSI 2000), but has no bearing on carcinogenicity because of the lack of good correlation between "overload" and severity of inflammation in these two species. It is therefore somewhat misleading to equate particle-induced carcinogenesis simply to "overload". The crucial factors are the cellular and subcellular events actually taking place within the pulmonary parenchyma.

An important implication of particle-induced inflammation being the initial event in the carcinogenic process in the rat lung, especially by non-genotoxic (not directly genotoxic) particles such as non-asbestiform talc or carbon black, is that tumors will not develop if there is little or no inflammation. In other words, there is a threshold below which tumors would not be expected to occur. The best evidence to date linking inflammation and mutational events, which are believed to be necessary for tumorigenesis, is from Driscoll et al. (1996). Using the hprt clonal selection assay on isolated alveolar type II cells from rats exposed to carbon black, they found significantly increased mutant frequencies caused by a 13-week exposure (6 hours/day, 5 days/week) to 7.1 and 52.8 mg/m^3. No significant mutant frequency was caused by exposure to 1.1 mg/m^3 carbon black, and at this concentration there was no significant pulmonary inflammation, thus

supporting the likelihood of absence of mutational events in the absence of significant inflammation.

b. What is the sensitivity of the rat to particle-induced carcinogenesis compared to other species?

The greatly reduced sensitivity of mice and hamsters to particle-induced inflammation compared to rats, and the correlation with absence of demonstrated carcinogenesis in mice and hamsters, was dealt with in the preceding section. Among rodent species tested, therefore, particle-induced carcinogenesis is either a rat specific phenomenon or it is only manifest in the rat across the concentration ranges tested. The upper ranges tested have often been well above realistic levels.

Carcinogenicity of poorly soluble particles has not been tested in non-human primates, but short-term tests do reveal differences in particle response between rats and primates. Comparison of lungs from Cynomolgus monkeys and F344 rats exposed to the same concentrations of shale dusts for 2 years revealed chronic active inflammation associated with intra-alveolar particles in rats but more interstitialization and much less alveolar-related inflammation in monkeys (MacFarland and Coate, 1982). Comparison of Cynomolgus monkeys and Sprague-Dawley rats exposed to coke dust for 2 years produced similar findings (Klonne et al.1987). Nikula et al. (1997) carried out more detailed studies on intrapulmonary particles, including morphometric analyses, in comparing responses of male Cynomolgus monkeys and F344 rats to 2 years of exposure to coal dust, diesel exhaust or a combination of both. Their findings confirmed and amplified the earlier studies with Cynomolgus monkeys. Rats retained a greater proportion of particulate material within terminal airspaces and this was associated with significant alveolar inflammation, epithelial hyperplasia and septal fibrosis. Monkeys, in contrast, retained a greater proportion of particulate

material in the pulmonary interstitium and did not have significant alveolar inflammation and its sequelae.

These consistent findings in monkeys are important for two main reasons. First, they indicate that comparing pulmonary responses in rats and monkeys on the basis of accumulated lung burden of particles is erroneous. The relative compartmentalization of the retained lung burden is crucial to interpretation. Interstitialization of poorly soluble, low toxigenic particles takes them out of the inflammatory pathway - as evidenced by the findings in Cynomolgus monkeys. When such particles are interstitialized in rats, they also do not incite an inflammatory response (Dungworth et al. 1994). Even if humans are predicted to accumulate higher lung burdens of talc than rats and mice (Pickrell et al. 1989, also cited on p.65 of the NTP draft document), this does not imply greater ensuing inflammation and probability of tumor development because interstitialization has not been taken into account. Second, the findings show that the response in a non-human primate is more like that in mice and hamsters than rats.

Various reasons have been adduced to account for the greater sensitivity of rats to particle-induced tumors in the lungs (Donaldson 2000). Emphasis has been on a greater proinflammatory and proliferative milieu. Factors that are may be just as important, and which do not appear to have been explored, are: I) that proliferating epithelium in an inflammatory environment in the rat lung might be more prone than that of other species to errors in DNA replication, and 2) the epithelial cells might be less able to repair DNA errors arising either spontaneously or as a result of substances such as reactive oxygen species.

c. What is the probable relevance for humans of particle-induced lung tumors in rats?

There is a strong link between particle-induced inflammation and eventual tumor development in the rat. The evidence obtained thus far indicates much less inflammation in monkeys than rats. It is therefore reasonable to presume that there is much less drive toward carcinogenesis in monkeys than rats. Evidence that humans respond more like monkeys than rats has been provided by Green (2000). He has concluded that, with a particle such as coal dust, the chronic active inflammation and epithelial hyperplasia in terminal air spaces, which pave the way for carcinogenesis in the rat, do not occur to a significant degree in humans.

Support for the conclusion that the rat is not predictive for humans at overload exposures to poorly soluble, low toxigenic and non-genotoxic particles is also provided by epidemiologic studies. Moderate to high lung burdens of particles such as coal dust, titanium dioxide, carbon black and talc have occurred following occupational exposures. No conclusive evidence of an increase in lung cancer caused by the particles alone has been found (IARC 1987, 1989, 1996, 1997, NTP draft 2000). Watson and Valberg (1996) also concluded that the rat is not predictive for humans in dealing with this type of particle at overload exposures. All the evidence points to the conclusion that if a mechanism for tumor induction in human lungs by poorly soluble, non-genotoxic particles exists, as in the rat (i.e. out of persistent inflammation and epithelial proliferation), it would be manifest at exposure concentration levels well above those needed for clinically-significant pneumoconiosis to become apparent.

3. Significance of Increased Frequency of Adrenal Medullary Pheochromocytomas in the NTP (1993) Study

Increased frequencies of benign and malignant pheochromocytomas occurred in both male and female rats exposed to 18 mg/m^3 of non-asbestiform talc. The increase was from 53% (control) to 79% (treated) in males and from 27% (control) to 47% (treated) in females. The increase was significant in males at the $P<0.01$ level and in females at the $P<0.05$ level. Goodman (1995) has presented evidence for why only the increase in males should be considered statistically significant, and this seems to be the most sensible approach. Calculating significance on the basis of combined benign and malignant pheochromocytomas is realistic because of the difficulty in histologic differentiation between benign and malignant proliferations (Tischler and DeLellis 1998). The biological relevance, however, has even more bearing on the interpretation than the question of whether a statistically significant increase has been achieved, and whether it is directly related to talc.

Hyperplastic and neoplastic proliferation in the rat adrenal medulla is frequent in some strains of rats commonly used in carcinogenicity studies, the F344 included. This often causes a problem in interpretation of true carcinogenic hazard (see reviews by Tischler and DeLellis 1988, Tischler and Coupland 1994, Tischler 1996).

Dietary influences and a wide variety of exogenous agents can increase the incidence of pheochromocytomas, particularly in strains like the F344 with a high spontaneous (control) incidence. The evidence indicates that various humoral and neural (cholinergic) stimuli can increase the rate of proliferation of chromaffin cells of the adrenal medulla and this facilitates development of neoplastic transformation (Tischler and Coupland 1994, Tischler 1996). The F344 rat, with control incidences in the range of 30 to 50% in males and 15 to 30% in females, appears to be especially vulnerable to small variations in proliferative stimuli from

neural, hormonal or growth factor stimuli. It is therefore not surprising that rats with impaired pulmonary function caused by talc pneumoconiosis at the 18 mg/m^3 exposure level have increases in pheochromocytomas. Another plausible, but unsubstantiated mechanism could be the influence of fibroblast growth factor (FGF) associated with the progressively fibrosing lung inflammation and which has been shown to stimulate proliferation of chromaffin cells (see Tischler 1996).

It is therefore highly improbable that an increase of pheochromocytomas in the proliferatively labile chromaffin cells of the F344 rat can be considered relevant for human carcinogenesis, especially when they are associated with chronic pulmonary inflammation and functional impairment.

D. Non-inhalation Studies

The most significant of these quoted in the NTP draft is the failure of intraperitoneal injection of talc in rats to cause a significant increase in mesotheliomas (Pott et al. 1974, 1976). This should be interpreted in light of the fact that intraperitoneal injection of dusts is a highly sensitive system for detecting carcinogenic potency (Pott and Roller 1996, Roller and Pott 1998). The results of the study by Stanton et al. (1981) are not interpretable because of uncertainties about the nature of the various grades of talc and the possibility of confounding by "solid state" carcinogenesis. The study by Bischoff and Bryson (1976) is also unconvincing in showing any biologically plausible tumor response to the intrathoracic injection of USP grade talc.

E. Genotoxicity Studies

The studies reviewed by the NTP Talc draft document and by NTP (1993) are consistent in their not demonstrating any genotoxic effects of non-asbestiform

talc in prokaryotic, eukaryotic or mammalian systems *in vitro*. The conclusion that non -asbestiform talc is not directly genotoxic is not therefore in dispute.

F. Conclusions

Based on the foregoing analysis, the following conclusions are justified:

- Rigorous evaluation of the animal studies does not support the NTP conclusion that, based on animal studies, non-asbestiform talc can reasonably be anticipated to be a human carcinogen because neither the requirement for activity in multiple species, nor multiple sites, nor multiple routes has been satisfied adequately.

- Non-asbestiform talc induces pulmonary tumors in female rats at considerable lung burdens and when associated with persistent, severe, chronic active inflammation and accompanying epithelial proliferation.

- The carcinogenic effect in the rat lung has a threshold related to the precondition for considerable particle-induced inflammation and epithelial proliferation.

- Even if a mechanism for tumor induction by poorly soluble, non-genotoxic particles were to exist in human lungs analogous to that in rats (and evidence thus far runs counter to this), tumors would arise at exposure concentration levels well above those that cause clinically relevant pneumoconiosis. The NTP draft document (pp. 71/72) contains a statement agreeing with this conclusion.

- The statistically significant ($P < 0.01$) increase in adrenal pheochromocytomas in male F344 rats, which have a very high incidence in controls, does not provide biologically plausible evidence for a true carcinogenic effect given the

susceptibility of chromaffin cells in the F344 rat to neural, humoral and growth factor stimuli.

G. References

ACGIH (1994) American Conference of Governmental Industrial Hygienists: 1993-1994. Threshold limit values for chemical substances and physical agents and biological exposure indices. ACGIH, Cincinnati, OH

Bischoff F, Bryson G (1976) Talc at the rodent intrathoracic, intraperitoneal, and subcutaneous sites [Abstract]. Proc Amer Assoc Cancer Res 17:1

Boorman GA, Seely JC (1995) The lack of an ovarian effect of lifetime talc exposure in F344/N rats and B6C3F1 mice. Regul Toxicol Pharmacol 21:242-243

Donaldson K (2000) Nonneoplastic lung responses induced in experimental animals by exposure to poorly soluble nonfibrous particles. Inhal Toxicol 12: 121-139

Driscoll KE, Carter JM, Howard BW et al. (1996) Pulmonary inflammatory, chemokine, and mutagenic responses in rats after subchronic inhalation of carbon black. Toxicol Appl Pharmacol 136:372-380

Dungworth DL, Mohr U, Heinrich U, Ernst H, Kittel B (1994) Pathologic effects of inhaled particles in rat lungs: associations between inflammatory and neoplastic processes. *In* Toxic and Carcinogenic Effects of Solid Particles in the Respiratory Tract, Mohr U, Dungworth DL, Mauderley JL, Oberdorster G (Eds), ILSI Press, Washington, DC, pp 75-98

Goodman JI (1995) An analysis of the National Toxicology Program's (NTP) tecnnical report (NTP TR 421) on the toxicology and carcinogenesis of talc. Regul Toxicol Pharmacol 21:244-249

Green FHY (2000) Pulmonary responses to inhaled poorly soluble particulate in the human. Inhal Toxicol 12:59-95

Heinrich U, Muhle H, Takenaka S et al. (1986) Chronic effects on the respiratory tract of hamsters, mice and rats after long-term inhalation of high concentratons of filtered and unfiltered diesel engine emissions. J Appl Toxicol 6:383-395

Heinrich U, Fuhst R, Rittinghausen S et el. (1995) Chronic inhalation exposure of Wistar rats and two different strains of mice to diesel engine exhaust, carbon black, and titanium dioxide. Inhal Toxicol 7:533-556

IARC (1987) Talc *In* Silica and Some Silicates, (42) IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Humans, Lyon, France, pp 185-224

IARC (1989) Titanium dioxide *In* Some Organic Solvents, Resin Monomers and Related Compounds, Pigments and Occupational Exposures in Paint Manufacture and Painting, (47) IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Humans, Lyon, France, pp 307-326

IARC (1996) Carbon black *In* Printing Processes and Printing Inks, Carbon Black and Some Nitro Compounds, (65) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Lyon, France, pp 149-262

IARC (1997) Coal dust *In* Silica, Some Silicates, Coal Dust and para-Aramid Fibrils, (68) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Lyon, France, pp 337-406

ILSI (2000) The relevance of the rat lung response to particle overload for human risk assessment: a workshop consensus report. Inhal Toxicol 12:1-17

Klonne DR, Burns JM, Halder CA et el. (1987) Two-year inhalation toxicity of petroleum coke in rats and monkeys. Amer J Indust Med 11:375-389

MacFarland HN, Coate WB (1982) Long-term inhalaton studies with raw and processed shale dusts. Ann Occup Hyg 26:213-226

Mauderley JL, Jones RK, Griffith WC et al. (1987) Diesel exhaust is a pulmonary carcinogen in rats exposed chronically by inhalation. Fundam Appl Toxicol 9: 208-221

Nikula KJ (2000) Rat lung tumors induced by exposure to selected poorly soluble nonfibrous particles. Inhal Toxicol 12:97-119

Nikula KJ, Avila KJ, Griffith WC, Mauderley JL (1997) Lung tissue responses and sites of particle retention differ between rats and Cynomolgus monkeys exposed chronically to diesel exhaust and coal dust. Fundam Appl Toxicol 37:37-53

Oberdorster G (1995) The NTP talc inhalation study: a critical appraisal focused on lung particle overload. Regul Toxicol Pharmacol 21:233-241

Pickrell JA, Snipes MB, Benson JM et al. (1989) Talc deposition and effects after 20 days of repeated inhalation exposure of rats and mice to talc. Environ Res 49: 233-245

Pott F, Dolgner R, Friedrichs KH, Huth F (1976) Oncogenic effect of fibrous dusts. Animal experimentation and its relation to human carcinogenesis. Ann Anat Pathol 21:237-246

Pott F, Huth F, Friedrichs KH (1974) Tumorigenic effect of fibrous dusts in experimental animals. Environ Health Perspect 9:313-315

Pott F, Roller M (1996) Carcinogenicity of synthetic fibres in experimental animals: its significance for workers. J Occup Health Safety 12:333-339

Roller M, Pott F (1998) Carcinogenicity of man-made fibres in experimental animals and its relevance for classificaton of insulation wools. Eur J Oncol 3: 231-239

Stanton MF, Layard M, Tegeris A, et al.(1981) Relation of particle dimension to carcinogenicity in amphibole asbestoses and other fibrous minerals. J Nat Cancer Inst 67:965-975

Watson AY, Valberg PA (1996) Particle-induced lung tumors in rats: evidence for species specificity in mechanisms. Inhal Toxicol 8 (suppl): 227-257

Wehner AP, Zwicker GM, Cannon WC, Carlton WW (1977) Inhalation of talc baby powder by hamsters. Fd Cosmet Toxicol 15:121-129

Tischler AS (1996) Cell proliferation in the adult adrenal medulla: chromaffin cells as a model for indirect carcinogenesis. *In* Endocrine System, Second Edn, ILSI Monographs on Pathology of Laboratory Animals, Jones TC, Capen CC, Mohr U (Eds), Springer Verlag, Berlin, Heidelberg , New York, pp 405-411

Tischler AS, Coupland RE (1994) Changes in structure and function of the adrenal medulla. *In* Pathobioogy of the Aging Rat, Vol 2, Mohr U, Dungworth DL, Capen CC (Eds), ILSI Press, Washington DC, pp 245-268

Tischler AS, DeLellis RA (1988) The rat adrenal medulla. II. Proliferative lesions. J Amer Coll Toxicol 7:23-44

Zazenski R, Ashton WH, Briggs D et al. (1995) Talc: occurrence, characterization, and consumer applications. Regul Toxicol Pharmacol 21: 218-229

MICHIGAN STATE UNIVERSITY

DEPARTMENT OF PHARMACOLOGY AND TOXICOLOGY
B440 LIFE SCIENCES BUILDING

EAST LANSING • MICHIGAN • 48824-1317
PHONE: (517) 353-7146
FAX:(517) 353-8915

November 22, 2000

Board of Scientific Counselors
National Toxicology Program
National Institute of Environmental
Health Sciences
P.O. Box 12233
Research Triangle Park, NC 27709

Re: Proposal to list non-asbestiform talc in the 10th Report on Carcinogens

Dear Board Members:

I was a member of the National Toxicology Program's Board of Scientific Counselors (NTP Board), and a member of the Board's Technical Reports Review Subcommittee, when the report "NTP Toxicology and Carcinogenicity Studies of Nonasbestiform, Cosmetic Grade Talc" (the NTP Talc Report) was reviewed on June 23-24, 1992. My purpose in writing to you is to request that you reject the proposal captioned above.

In my view, the appropriate conclusion to be drawn from the NTP Talc Report is that the maximum tolerated dose (MTD) was exceeded in the female rats exposed to the high dose of talc (18 mg/m^3 for 6 hours daily, 5 days/week), based upon marked, chronic lung toxicity. Therefore, the data obtained in the female rats at this dose are not appropriate for human risk assessment. The results of the NTP bioassay indicated clear evidence of carcinogenic activity of talc only in female rats exposed to the dose that, in my view, exceeded the MTD. These remarks are not intended to connote that the talc bioassay is invalid. The good news is that under the conditions of this study mice did not develop tumors, male rats did not develop lung tumors, and female rats exposed to 6 mg/m^3 for 6 hours daily, 5 days/week did not develop lung tumors. Furthermore, the pheochromocytomas observed in the rats, when viewed in light of the high spontaneous incidence of this tumor, do not appear to be treatment-related. Thus, it is not appropriate to employ the results of the NTP talc bioassay as a basis for a cancer risk assessment involving a linear extrapolation to estimate a theoretical risk to humans that might be exposed to relatively low doses of talc. Importantly, I believe that the NTP data should be regarded as supporting the safety of talc under typical conditions of use.

A thorough discussion of the NTP Talc Report and the basis for the views presented above appear in the following publication: Goodman, Jay I. (1995) An analysis of the National Toxicology Program's (NTP) Technical Report (NTP TR 421) on the Toxicology and Carcinogenesis Studies of Talc, Regulatory Toxicology and Pharmacology 21: 244-249. It is my opinion that the views I expressed 5 years ago remain valid today.

Thank you for considering my views, and please do not hesitate to contact me if you would like additional information.

Sincerely,

[redacted]

Jay I. Goodman, Ph.D.
Professor

REGULATORY TOXICOLOGY AND PHARMACOLOGY 21, 244–249 (1995)

An Analysis of the National Toxicology Program's (NTP) Technical Report (NTP TR 421) on the Toxicology and Carcinogenesis Studies of Talc[1]

JAY I. GOODMAN[2]

Department of Pharmacology and Toxicology, Institute for Environmental Toxicology, B-440 Life Sciences Building, Michigan State University, East Lansing, Michigan 48824

Received October 1, 1994

The NTP toxicology and carcinogenicity studies of nonasbestiform, cosmetic-grade talc (the NTP Talc Report) were conducted by exposing male and female F344/N rats and B6C3F1 mice to target aerosol concentrations of 0, 6, and 18 mg/m³ talc for 6 hr daily, 5 days per week. Based on results of the high dose, the Report concluded that talc caused lung tumors in female rats and pheochromocytomas in male and female rats, and there was no evidence of carcinogenic activity in mice. A thorough evaluation of lung toxicity revealed that talc-induced lung tumors occurred only in the group of animals that exhibited the most profound degree of chronic toxicity. However, these data were presented as empirical observations rather than discussed in a manner that would relate them to the risk assessment implications of the bioassay, i.e., relevant data were collected but not "used." In addition, the evaluation of the pheochromocytomas was inadequate because it failed to place sufficient emphasis on the spontaneous incidence of this tumor in rats. These deficiencies caused the author to vote against the conclusions presented in the Talc Report when it was reviewed by the NTP Board of Scientific Counselors. The appropriate conclusions are (1) the data do not indicate that the pheochromocytomas were treatment-related; (2) the maximum tolerated dose (MTD) was exceeded in the female rats exposed to the high dose; and (3) talc is not expected to cause lung tumors under conditions of exposure that fail to result in marked chronic lung toxicity.

INTRODUCTION

The NTP toxicology and carcinogenicity studies of nonasbestiform, cosmetic-grade talc (the NTP Talc Report) were conducted by exposing male and female F344/N rats and B6C3F1 mice to target aerosol concentrations of 0, 6, and 18 mg/m³ talc for 6 hr daily, 5 days per week (NTP, 1993). A unique aspect of these studies was the inclusion of a very thorough evaluation of lung toxicity, in addition to the standard gross pathology and histopathology assessments. The NTP Talc Report was reviewed by the NTP Board of Scientific Counselors (The Board), Technical Reports Review Subcommittee on June 23–24, 1992. Jay I. Goodman participated in this review as a member of the Board and a member of the Board's Technical Reports Review Subcommittee. The Board concurred with the recommendations of the NTP staff and voted to approve (by a vote of seven yes votes to one no vote (J. I. Goodman)) the following conclusions: "Under conditions of these inhalation studies there was some evidence of carcinogenic activity of talc in male F344/N rats based on an increased incidence of benign and malignant pheochromocytomas of the adrenal gland. There was clear evidence of carcinogenic activity of talc in female F344/N rats based on increased incidences of alveolar/bronchiolar adenomas and carcinomas of the lung and benign and malignant pheochromocytomas of the adrenal gland. There was no evidence of carcinogenic activity of talc in male or female B6C3F1 mice exposed to 6 or 18 mg/m³."

The specific aims of this paper are to focus on the studies involving rats, the species in which there was deemed to be a carcinogenic effect, and (1) to provide an overview of the talc bioassay (NTP, 1993) with an emphasis on the rationale for the author's disagreement with the conclusions reached by the other members of the Board regarding the carcinogenicity of talc and (2) to provide a realistic perspective regarding the significance of the talc bioassay.

SYNOPSIS OF THE CARCINOGENICITY STUDY OF TALC

Groups of 50 male and 49 or 50 female F344/N rats were exposed to aerosols containing 0, 6, or 18 mg/m³ talc, 6 hr a day 5 days per week, until mortality in any

[1] Presented, in part, at the International Society of Regulatory Toxicology and Pharmacology/U.S. FDA Workshop on Talc: Consumer Uses and Health Perspectives, NIH, Bethesda, MD, January 31–February 1, 1994.

[2] To whom correspondence should be addressed. Fax: (517) 353-8915.

0273-2300/95 $6.00

TABLE 1
Summary of the Lifetime and 2-Year Carcinogenicity Studies of Talc[a]

	Male F344/N rats	Female F344/N rats
Exposure levels	0, 6, or 18 mg/m³	0, 6, or 18 mg/m³
Body weights	High-dose group slightly lower than controls	High-dose group slightly lower than controls
Survival rates	9/50, 14/50, 16/50	11/50, 13/49, 9/50
Neoplastic effects	Adrenal medulla: benign or malignant pheochromocytoma (26/49, 32/48, 37/47)	Lung: alveolar/bronchiolar adenoma or carcinoma (1/50, 0/48, 13/50) Adrenal medulla: benign or malignant pheochromocytoma (18/48, 14/47, 23/49)
Level of evidence of carcinogenic activity	Some evidence	Clear evidence

[a] National Toxicology Program, 1993.

exposure group reached 80% (113 weeks for males and 122 weeks for females). The survival of male and female rats exposed to talc was similar to that of the controls, and the mean body weights of the animals in the high-dose group were only slightly lower (not more than approximately 10% lower) than those of controls after Week 65 (NTP, 1993). The results were interpreted as indicating some evidence of carcinogenicity based on benign and malignant pheochromocytomas in male and female rats and clear evidence of carcinogenicity based on alveolar/bronchiolar adenomas and carcinomas of the lung in female rats. The study is summarized in Table 1.

Groups of 47 to 49 male and 48 to 50 female B6C3F1 mice were exposed to aerosols containing 0, 6, or 18 mg/m³ talc, 6 hr a day 5 days per week, for up to 103 or 104 weeks. Final mean body weights and survival of the mice exposed to talc were similar to those of the controls (NTP, 1993). The results were interpreted as indicating no evidence of carcinogenic activity in mice.

INTERIM EVALUATIONS IN RATS

In a parallel series of ancillary studies, additional groups of animals were exposed to talc, as described above, and examined for interim pathology evaluations or pulmonary function tests after 6, 11, 18, and 24 months and lung biochemistry and cytology studies after 24 months. These are outlined below.

PARAMETERS ASSESSED IN THE ANCILLARY STUDIES

- Lung talc burden
- Pulmonary function
- Lung biochemistry; bronchoalveolar lavage fluid was analyzed to determine:
  - Cell injury—lactate dehydrogenase.
  - Chronic inflammation—polymorphonuclear leukocytes, pulmonary macrophages, alkaline phosphatase, and protein.
  - Lysosomal activation—β-glucuronidase and acid phosphatase.
  - Response to oxidant injury—increased glutathione reductase.
- Lung collagen metabolism, protein synthesis, and proteinase activity.

LUNG TALC BURDEN, LUNG WEIGHTS, AND RESPIRATORY FUNCTION

The lung talc burden at the 6-, 11-, 18-, and 24-month interim sacrifice times was similar in male and

TABLE 2
Lung Talc Burden (Normalized to Exposure Concentration) of Rats[a]

	Male		Female	
	6 mg/m³	18 mg/m³	6 mg/m³	18 mg/m³
6-Month interim	0.439 ± 0.040[b]	0.602 ± 0.013[c]	0.406 ± 0.032	0.464 ± 0.007[c]
12-Month interim	0.731 ± 0.098	1.165 ± 0.113[c]	0.785 ± 0.043	0.787 ± 0.187
18-Month interim	1.22 ± 0.12	1.53 ± 0.05	1.28 ± 0.06	1.35 ± 0.04
24-Month interim	1.74 ± 0.21	1.34 ± 0.19	1.52 ± 0.15	1.63 ± 0.13

[a] Table F3, National Toxicology Program, 1993.
[b] Units are presented as mg talc/g control lung/mg/m³.
[c] Significantly different ($P \leq 0.05$) from the 6 mg/m³ group by Dunn's or Shirley's test.

TABLE 3
Total Lung Capacity of Rats[a]

	0 mg/m³	6 mg/m³	18 mg/m³
		Male	
6-Month interim	19.86 ± 0.54[b]	19.48 ± 0.46	19.25 ± 0.39
11-Month interim	20.06 ± 0.32	18.44 ± 0.39**	17.67 ± 0.45**
18-Month interim	20.30 ± 0.45	18.87 ± 0.41**	16.34 ± 0.52**
24-Month interim	20.50 ± 0.83	20.20 ± 0.28	16.47 ± 1.53
		Female	
6-Month interim	14.20 ± 0.25	14.56 ± 0.27	13.80 ± 0.27
11-Month interim	13.29 ± 0.21	12.91 ± 0.17	12.06 ± 0.26**
18-Month interim	13.34 ± 0.26	12.68 ± 0.28**	11.43 ± 0.31**
24-Month interim	14.85 ± 0.31	13.73 ± 0.34**	11.50 ± 1.07**

[a] Table F10, National Toxicology Program, 1993.
[b] Units are presented as ml; ratio is (dosed group mean/control group mean) × 100.
* $P \leq 0.01$.
** Significantly different ($P \leq 0.05$) from the control by Dunn's and Shirley's test.

female rats exposed to 6 or 18 mg/m³ talc, respectively (Table 2).

The absolute and relative lung weights of male rats exposed to 18 mg/m³ were significantly greater than those of controls at the 6-, 11-, and 18-month interim evaluations and at the end of the lifetime study, while those of female rats exposed to 18 mg/m³ were significantly greater at the 11-, 18-, and 24-month interim evaluations and at the end of the lifetime study. In both males and females, there was a concentration-related impairment of respiratory function which increased in severity with increasing duration of exposure (NTP, 1993). While decreases in total lung capacity were noted in males and females, statistically significant reductions at the 24-month interim were only seen in females and this occurred at both the 6 and 18 mg/m³ doses (Table 3).

BRONCHOALVEOLAR LAVAGE AND LUNG BIOCHEMISTRY

Following the completion of the pulmonary function tests at the 24-month interim evaluation a thorough assessment of biochemical and cellular parameters that could provide insight regarding lung toxicity was made. Bronchoalveolar lavage was performed on the remaining rats in these groups and the lavage fluid was evaluated for enzymes and protein (Table 4) and cell content (Table 5). In addition, lung collagen metabolism and protein synthesis were assessed (Table 6) and proteinase activity in both lavage fluid and lung homogenate supernatant fluid was evaluated (Table 7, males; and Table 8, females). While dose-related increases in these parameters (indicative of toxicity) were noted in both males and females, there was a larger number of parameters that showed statistically significant increases in the females and, in general, the magnitude of the increases was higher in the females.

ASSESSMENT OF THE NTP TALC REPORT

The NTP Talc Report noted that both *in vitro* and *in vivo* studies indicate that talc is not genotoxic and it included a thorough evaluation of lung toxicity (NTP, 1993). However, these data were presented as empirical observations rather than discussed in a manner that would relate them to the risk assessment implications of the bioassay. In other words, relevant data were collected but not used. In addition, the evaluation of the pheochromocytomas was inadequate because it failed to place sufficient emphasis on the spontaneous incidence of this tumor in rats. These deficiencies caused the author to vote against the conclusions presented in the Talc Report when it was reviewed by the NTP Board of Scientific Counselors. The rationale for this decision in the context of a more realistic perspective on the Report is presented below.

The carcinogen bioassay is a qualitative test; it is not a risk assessment. The use of biological information is required in order to place results of the bioassay into proper perspective and to take a rational approach toward risk assessment (Goodman, 1990, 1994). This notion was endorsed in a recent review of the NTP by the NTP Board of Scientific Counselors where it was stated that "Mechanistic studies should pervade the activities of the NTP. . . . With this emphasis the NTP should be able to strengthen its interpretation of test results in the context of human health. Such interpretation should become a part of reports issued by the NTP" (NTP, 1992; Goodman, 1994). It is appropriate to consider the NTP Talc Report in this context.

Talc-induced lung tumors were not detected in male rats, female mice, or male mice. In rats, the principal toxic lesions associated with inhalation exposure to talc included chronic granulomatous inflammation, alveolar epithelial hyperplasia, squamous metaplasia and squamous cysts, and interstitial fibrosis of the lung. These lesions were accompanied by impaired pulmonary function. While the talc burden in the lungs of males and females was similar (Table 2), the degree of chronic toxicity and inflammation was substantially higher in the females (Tables 3–8; NTP, 1993). In mice, inhalation exposure to talc produced some chronic inflammation. In contrast to rats, alveolar epithelial hyperplasia, squamous metaplasia, and interstitial fibrosis were not observed. Overall, markedly less talc-induced lung toxicity was produced in mice than in rats (NTP, 1993). These observations need to be juxtaposed with the findings regarding lung tumors. It is apparent that an increase in lung tumors was seen only in the test animals that clearly exhibited the highest degree of chronic lung toxicity, the female rats exposed to the high dose (18 mg/m³) of talc. The amount of toxicity leads the author to

TABLE 4
Bronchoalveolar Lavage Fluid Enzymes of Rats at the 24-Month Interim Evaluation[a]

	0 mg/m³	6 mg/m³	18 mg/m³
Male			
β-Glucuronidase[b]	1.09 ± 0.40	18.86 ± 3.20*	89.24 ± 14.24**
Lactate dehydrogenase	1634 ± 545	3193 ± 606	8262 ± 380*
Alkaline phosphatase	364.7 ± 147	572.8 ± 86.8	1604.7 ± 143*
Glutathione reductase	103.08 ± 16.43	99.35 ± 19.79	110.99 ± 0.55*
Total protein[c]	1.78 ± 0.40	3.12 ± 0.64	5.79 ± 0.55*
Female			
β-Glucuronidase[b]	3.33 ± 0.97	41.05 ± 4.39**	154.16 ± 17.21**
Lactate dehydrogenase	1655 ± 266	3906 ± 444*	14E3 ± 1E3**
Alkaline phosphatase	427.8 ± 30.9	853.6 ± 79.7**	2594.7 ± 221**
Glutathione reductase	100.6 ± 1.7	135.2 ± 22.4	460.0 ± 44.8*
Total protein[c]	1.20 ± 0.22	4.30 ± 0.36**	12.96 ± 0.28**

[a] Table F4, National Toxicology Program, 1993.
[b] Units presented as mIU/g control lung.
[c] Units presented as mg/g control lung.
* Significantly different ($P \leq 0.05$) from the control group by Dunn's and Shirley's test.
** $P \leq 0.01$.

conclude that the MTD was exceeded in the female rats exposed to the 18 mg/m³ dose.

The overall incidences of pheochromocytomas (benign, malignant, or complex) of the adrenal medulla that occurred in the 18 mg/m³ exposed rats were significantly greater than those of the controls (males, $P = 0.006$; females, $P = 0.024$) (NTP, 1993). However, the spontaneous incidence of these tumors was 53% in males and 27% in females. In the author's view, the criterion for statistical significance should be $P < 0.01$. This is consistent with the "guideline" indicating that a compound should be considered carcinogenic if the highest dose produces an increase in a common tumor that is significant at the 1% ($P < 0.01$) level or an increase in a rare tumor that is significant at the 5% ($P < 0.05$) level; otherwise, the compound is regarded as a noncarcinogen (Haseman, 1983; Haseman *et al.*, 1986). A rare neoplasm was defined as a neoplasm occurring with a frequency of less than 1%. Therefore, there is no statistically significant increase in pheochromocytomas in talc-treated female rats and talc should not be deemed carcinogenic at this site in these animals.

There has been a pronounced increase in the spontaneous occurrence of pheochromocytomas in male F344 rats, in studies conducted by the NTP, over the past 10 years (Rao *et al.*, 1990; NTP, 1993). In addition, the high rate of pheochromocytomas in the control female rats in the talc bioassay (13/48, 27%) is in sharp contrast to reports indicating that the spontaneous incidence in control female F344 rats, in NTP studies, has been less than 10% (Rao *et al.*, 1990), e.g., in the range of 5.2–7.1% between 1980 and 1983 (Haseman and Rao, 1992). Unfortunately, the NTP failed to make a comparison of the spontaneous pheochromocytomas observed in the talc

TABLE 5
Bronchoalveolar Lavage Fluid Cell Populations of Rats at the 24-Month Interim Evaluation[a]

	0 mg/m³	6 mg/m³	18 mg/m³
Male			
Polymorphonuclear cells[b]	0.333 ± 0.167	24.417 ± 2.557*	32.50 ± 3.000*
Lymphocytes	0.000 ± 0.000	0.500 ± 0.258	0.500 ± 0.500
Macrophages	93.67 ± 3.3.72	70.25 ± 2.53*	62.75 ± 1.75*
Epithelial cells	6.00 ± 3.61	4.83 ± 1.41	4.25 ± 1.75
Female			
Polymorphonuclear cells	0.625 ± 0.315	25.778 ± 2.673**	37.000 ± 1.528**
Lymphocytes	0.000 ± 0.000	0.722 ± 0.188*	1.333 ± 0.667*
Macrophages	91.38 ± 1.75	71.22 ± 2.95**	57.33 ± 4.67**
Epithelial cells	8.00 ± 2.01	2.28 ± 0.50*	4.33 ± 2.60

[a] Table F5, National Toxicology Program, 1993.
[b] Units presented as percentage of total cells.
* Significantly different ($P \leq 0.05$) from the control group by Dunn's and Shirley's test.
** $P \leq 0.01$.

TABLE 6
Lung Collagen Metabolism and Protein Synthesis in Rats at the 24-Month Interim Evaluation[a]

	0 mg/m³	6 mg/m³	18 mg/m³
Male			
Lavage fluid collagenous peptides[b]	39.79 ± 5.07	46.99 ± 6.51	79.21 ± 13.73
Total lung collagen[c]	13.87 ± 0.60	15.98 ± 0.39*	18.88 ± 3.35*
Collagen production[d]	1.58 ± 0.17	1.60 ± 0.17	1.63 ± 0.22
Collagen degradation[e]	31.67 ± 1.72	27.74 ± 1.42	9.18 ± 2.38**
Noncollagenous protein synthesis[f]	142.1 ± 14.5	199.8 ± 22.1*	312.2 ± 10.6**
Female			
Lavage fluid collagenous peptides	78.27 ± 11.64	115.36 ± 8.61*	174.71 ± 13.56**
Total lung collagen	14.32 ± 0.66	19.95 ± 1.58*	36.47 ± 3.39**
Collagen production	0.982 ± 0.185	1.804 ± 0.144*	2.264 ± 0.347**
Collagen degradation	14.41 ± 2.44	21.59 ± 4.99	9.88 ± 1.63
Noncollagenous protein synthesis	173.9 ± 34.5	325.8 ± 90.9	564.3 ± 107*

[a] Table F7, National Toxicology Program, 1993.
[b] Units are presented as μg/g control lung.
[c] Units are presented as mg/g control lung.
[d] Units are presented as percentage new protein.
[e] Units are presented as percentage new collagen.
[f] Units are presented as dpm × 10^3/g control lung.
* Significantly different ($P \leq 0.05$) from the control group by Dunn's and Shirley's test.
** $P \leq 0.01$.

bioassay with the historical spontaneous incidence of pheochromocytomas in rats. This would have been particularly appropriate (as the author indicated during the Board's review of the NTP Talc Report, June 23–24, 1992).

The marked increase in the spontaneous incidence of pheochromocytomas in female rats coupled with the lack of a statistically significant increase of this tumor in the treated animals (when the appropriate P value is employed) and the trend toward an increasing spontaneous incidence in males indicate that the data presented in the NTP Talc Report do not warrant a conclusion that talc causes pheochromocytomas in rats.

CONCLUSIONS

The appropriate conclusion to be drawn from the NTP Talc Report is that the MTD was exceeded in the female rats exposed to the high dose and talc is not expected to cause lung tumors under conditions of exposure that fail to result in marked chronic lung toxicity. Clear evidence of carcinogenic activity of talc was

TABLE 7
Proteinase Activity in Lavage Fluid and Lung Homogenate Supernatant Fluid of Male Rats at the 24-Month Interim Evaluation[a]

	0 mg/m³	6 mg/m³	18 mg/m³
Lavage fluid			
Acid proteinase	0.994 ± 0.329[b]	1.866 ± 0.174	4.307 ± 0.218*
Cathepsin D	0.147 ± 0.147	0.599 ± 0.150	2.420 ± 0.147**
Cathepsin B	0.924 ± 0.415	1.267 ± 0.094	1.887 ± 0.365
Homogenate supernatant fluid			
Acid proteinase	10.92 ± 0.64	17.51 ± 0.90*	25.13 ± 1.50**
Cathepsin D	8.53 ± 0.91	14.04 ± 0.62*	21.03 ± 1.56**
Cathepsin B	2.39 ± 0.41	3.48 ± 0.37	4.10 ± 0.06*
Neutral proteinase	0.715 ± 0.168	2.417 ± 0.304*	4.505[c]
PMN elastase cathepsin G	0.490 ± 0.218	1.936 ± 0.242*	4.457 ± 0.377**
Macrophage elastase collagenase	0.225 ± 0.099	0.482 ± 0.077	0.000[c]

[a] Table F8, National Toxicology Program, 1993.
[b] Units are presented as mg/hr/mg control lung.
[c] $n = 1$; no statistic calculated.
* Significantly different ($P \leq 0.05$) from the control group by Dunn's and Shirley's test.
** $P \leq 0.01$.

TABLE 8
Proteinase Activity in Lavage Fluid and Lung Homogenate Supernatant Fluid of Female Rats at the 24-Month Interim Evaluation[a]

	0 mg/m³	6 mg/m³	18 mg/m³
Lavage fluid			
Acid proteinase	1.52 ± 0.12[b]	8.46 ± 0.33[*]	6.05 ± 0.73[**]
Cathepsin D	0.015 ± 0.015	1.310 ± 0.292[*]	4.043 ± 0.578[**]
Cathepsin B	1.61 ± 0.26	2.15 ± 0.22	2.01 ± 0.17
Homogenate supernatant fluid			
Acid proteinase	14.04 ± 0.95	29.43 ± 1.18[**]	38.61 ± 1.81[**]
Cathepsin D	10.05 ± 0.63	22.97 ± 1.07[**]	30.25 ± 1.60[**]
Cathepsin B	3.99 ± 0.58	6.46 ± 0.60[*]	8.37 ± 0.42[**]
Neutral proteinase	0.648 ± 0.087	5.040 ± 0.418[**]	12.293 ± 1.598[**]
PMN elastase cathepsin G	0.785 ± 0.142	4.351 ± 0.261[**]	10.313 ± 2.694[**]
Macrophage elastase collagenase	0.054 ± 0.037	0.068 ± 0.175[*]	2.012 ± 1.126[*]

[a] Table F8, National Toxicology Program, 1993.
[b] Units are presented as mg/hr/mg control lung.
[*] Significantly different ($P \leq 0.05$) from the control group by Dunn's and Shirley's test.
[**] $P \leq 0.01$.

seen only in female rats exposed to the high dose of talc and only under circumstances in which there was evidence of marked chronic lung toxicity. These remarks are not intended to connote that the talc bioassay is invalid. The good news is that under the conditions of this study mice did not develop tumors, male rats did not develop lung tumors, and female rats exposed to the 6 mg/m³ dose did not develop lung tumors. Furthermore, it is doubtful that the pheochromocytomas observed in the talc-treated rats were treatment related. Thus, at the high dose, talc exhibited equivocal evidence of carcinogenic activity toward the adrenal medulla of male rats.

There is a considerable debate regarding the proper criteria for a MTD (Kociba, 1987). However, in a practical sense the chief concern with the MTD is how the resulting data are used in the risk assessment process (McConnell, 1989). Any high dose, no matter how high, that permits animals to live long enough to develop tumors is not necessarily an appropriate high dose. The recent review of the NTP addressed this issue and concluded that "the implicit assumptions underlying extrapolation from the MTD . . . do not appear to be valid. Therefore, both the criteria for selection of the high dose used and the default criteria that are employed for extrapolation from high dose to low dose must be reevaluated in a critical manner" (NTP, 1992; Goodman, 1994). A consideration of the data presented in the NTP Talc Report leads the author to conclude that it would not be appropriate to use the lung tumor end point in female rats as the basis for a risk assessment involving a linear extrapolation to estimate a theoretical risk to humans that might be exposed to relatively low doses of talc.

REFERENCES

Goodman, J. I. (1990). The carcinogen bioassay in perspective: A requirement for incorporation of biological information. *In Vitro Toxicol.* **3**, 1–7.

Goodman, J. I. (1994). A rational approach to risk assessment requires the use of biological information: An analysis of the National Toxicology Program (NTP), final report of the Advisory Review by the NTP Board of Scientific Counselors. *Regul. Toxicol. Pharmacol.* **19**, 51–59.

Haseman, J. K. (1983). A reexamination of false-positive rates for carcinogenicity studies. *Fundam. Appl. Toxicol.* **3**, 334–339.

Haseman, J. K., Winbush, J. S., and O'Donnell, M. W. (1986). Use of dual control groups to estimate false positive rates in laboratory animal carcinogenicity studies. *Fundam. Appl. Toxicol.* **7**, 573–584.

Haseman, J. K., and Rao, G. N. (1992). Effects of corn oil, time-related changes, and inter-laboratory variability on tumor occurrence in control Fischer 344 (F344/N) rats. *Toxicol. Pathol.* **20**, 52–60.

Kociba, R. J. (1987). Issues in biochemical applications to risk assessment: How should the MTD be selected for chronic bioassays? *Environ. Health Perspect.* **76**, 169–174.

McConnell, E. (1989). The maximum tolerated dose: The debate. *J. Am. Coll. Toxicol.* **8**, 1115–1120.

National Toxicology Program (NTP) (1992). Final report of the advisory review by the NTP Board of Scientific Counselors. *Fed. Regist.* **57**, 31721–31730.

National Toxicology Program (NTP) (1993). NTP technical report on the toxicology and carcinogenesis studies of talc in F344/N and B6C3F1 mice. NTP TR 421. U.S. Department of Health and Human Services, National Institutes of Health.

Rao, G. N., Haseman, J. K., Grumbein, S., Crawford, D. D., and Eustis, S. L. (1990). Growth, body weight, survival, and tumor trends in F344/N rats during an eleven-year period. *Toxicolog. Pathol.* **18**, 61–70.

Sontag, J. M., Page, N. P., and Saffiotti, U. (1976). Guidelines for carcinogen bioassay in small rodents. National Institutes of Health (NIH), DHEW Publ. No. 76-801, Washington, DC.

Interpretation of Epidemiologic Studies on Talc and Ovarian Cancer

prepared by

Kenneth J. Rothman
Harris Pastides
Jonathan Samet

November 28, 2000

Executive Summary

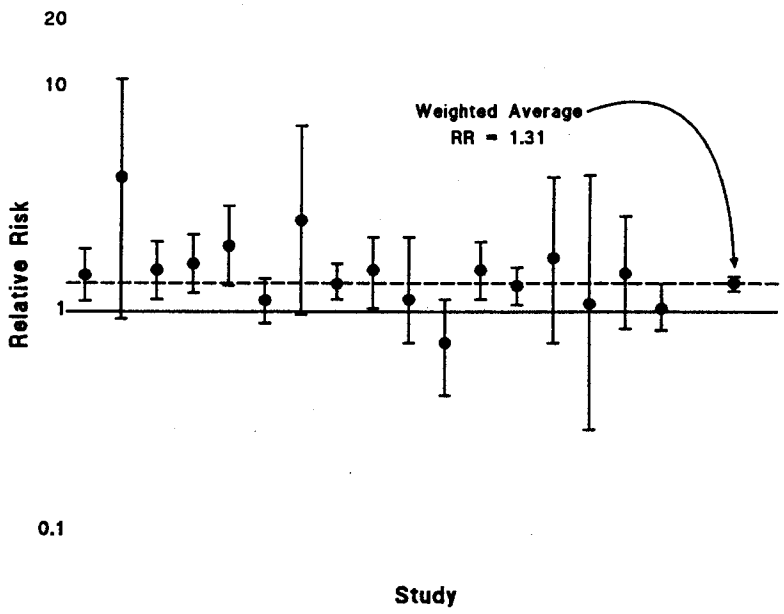
A weighted average of the results from epidemiologic studies to date measuring the relation between talc and ovarian cancer risk gives an overall relative risk of 1.31, with a 95% confidence interval of 1.21–1.41. Bias and causation are competing explanations for the weak positive association observed. This weak association could be an underestimate of a stronger association if there are errors in measuring talc exposure that apply uniformly to all study subjects (nondifferential misclassification). On the other hand, nondifferential misclassification does not bias an association that is null to begin with, so postulating nondifferential misclassification cannot shed light on whether the association results from a causal relation or not. Most of the published studies are interview-based case-control studies, subject to recall bias, which can readily give rise to associations of this magnitude. The evidence from these studies regarding recall bias is mixed. Uncontrolled confounding can also easily explain associations this weak; although no single confounding factor would seem to account for the overall effect, the combined effect of several such unidentified confounders could do so. In considering these competing explanations of bias and causation, the evidence in favor of a causal explanation is only the overall weak association of a relative risk of 1.31. The lack of a plausible biologic mechanism, on the other hand, weighs against a causal interpretation. Also weighing against a causal explanation is the dose-response pattern among talc users, which is an inverse trend for both duration of use and frequency of use. A causal relation would predict a positive trend, not an inverse trend. Based on these considerations, we suggest that the evidence to date does not indicate that talc can be "reasonably anticipated to be a human carcinogen."

Introduction

In this document we offer an interpretation of the epidemiologic literature with respect to the causal hypothesis that talc exposure causes an increase in the occurrence of ovarian cancer. Overall, we identified 23 epidemiologic studies conducted since 1980 that have examined consumer talc exposure with respect to subsequent risk for ovarian cancer.[1-23] The search methodology is described in the appendix. Sixteen of these were case-control studies reporting new data with effect estimates for talc exposure,[2-5,7,10,11,13-15,17-19,21-23] and one was a cohort study reporting an effect estimate.[9] One study examined occupational exposure to talc in women, but there were few exposed women in this study[16]; the other studies did not report quantitative effect estimates. The importance of this comparatively small set of epidemiologic studies is underscored by the paucity of relevant animal research on this question.

Most of these published reports come from epidemiologic studies in which talc was not the primary focus. Perhaps for this reason, talc exposure information was often crude. In only a few of these studies was there any attempt to categorize talc exposure by frequency of use or duration of use. For the 17 studies that reported some epidemiologic measure of effect, it was usually a relative risk estimate for ovarian cancer given that there was some exposure to talc, compared with no exposure or minimal exposure. These results are depicted graphically in figure 1. The findings on balance indicate a slight positive association between talc exposure and ovarian cancer, with an overall weighted relative risk of 1.31, and a 95% confidence interval of 1.21–1.41.

Figure 1
Study-specific Relative Risk Estimates for Ovarian Cancer Among Talc Users, and Overall Weighted Average of Study Results



Issues Affecting Causal Inference

Inferring a causal relation from a pattern of epidemiologic results follows no recipe, but certain principles can be applied. To begin with, what alternative explanations might be offered to explain a pattern of positive findings? If an uncontrolled confounding factor or a study-related bias could explain the results, a causal inference is less reasonable. Second, is there a plausible biologic mechanism? For example, environmental tobacco smoke shows a weak association with lung cancer in numerous epidemiologic studies of never smokers, but the plausibility of the relation, based on the known constituents of the smoke and their effect in higher concentrations, among active smokers, makes a causal inference more reasonable. Third, is there a consistent dose-response trend in the data? With rare exception, every causal relation in epidemiologic research shows a progressive relation between various measures of increasing exposure. In this discussion paper, we address the following issues that we believe are potentially relevant to causal inference regarding talc and ovarian cancer:

1. Exposure misclassification
2. Recall bias
3. Confounding
4. Dose-response trends
5. Biologic mechanism

Below we discuss briefly the import of each of these topics with respect to the interpretation of the epidemiologic literature of talc and ovarian cancer. We omit discussion of the role of chance in explaining any of the findings, because the combined weight of the 17 studies in figure 1 indicates that chance alone is an unlikely explanation for the overall weighted average of relative risks from the studies of 1.31. Other possible issues, such as selection biases and reverse causation might be relevant, but appear less important to us in interpreting these results, so we have omitted them in the interests of brevity. (Reverse causation, for example, could occur if preclinical ovarian cancer prompted women to use talc; while this situation is possible in some instances, we do not think it is a realistic explanation for the observed effects.)

Exposure Misclassification

Nearly all the studies were case-control studies. It is commonly believed that the validity of case-control studies is worse than that of cohort studies, but this view is mistaken. The validity of a study depends on the specifics of the study design, the nature of the data, and the nature of the hypothesis that the study addresses. For example, a cohort study that examines the long-term risk of cancer among coffee drinkers after a one-time dietary assessment of coffee consumption would suffer from weak exposure assessment. Although the exposure information might be accurate for the time at which it was collected, the exposure status of cohort members will change with time and the initial measure might be only poorly correlated with a more meaningful measure of coffee consumption. The effect of having a poor measure of exposure will be considerable nondifferential misclassification, a type of error that introduces a bias into study results that tends to drive effect estimates towards the null condition of no effect. In contrast, it may be possible to get more detailed exposure information from study subjects in a case-control study, which might thus avoid some of the bias that would result from a cohort study.

Much like coffee consumption, talc exposure is likely to vary over time as women age and their reasons for deciding to use talc change. Consequently a single baseline assessment of talc exposure at the start of follow-up in a cohort may lead to effect estimates that are biased toward the null. If talc habits are steady over time, a single baseline assessment becomes more informative. Furthermore, if talc use influences cancer risk with a long induction period, talc assessment at the start of a cohort study is more meaningful than an assessment of coffee drinking on heart disease risk, which is thought to have only a short-term effect.

Case-control studies also suffer from exposure misclassification, but the potential exists to extract more detailed history of exposure from the subject interview. In most of these studies, the exposure metric is based on interview information. It is subject to inaccuracies from recall error, as well as inaccuracies reflecting the nature of the questions asked and their relation to any biologically relevant measure of talc exposure. Ideally one would wish to have a measure of talc dose within the upper reproductive tract. The actual measures obtained by interview, however, are likely to be only modestly correlated with a hypothetically ideal measure. The result of this inevitable non-differential misclassification would be to bias any real effect towards the null. Nevertheless, one cannot draw the conclusion that the overall slight positive relation between talc exposure and ovarian cancer must be an underestimate of a larger effect because of nondifferential misclassification. Non-differential misclassification does not introduce any bias toward the null if the association is null to begin with, so to draw the conclusion that the overall effect estimate from the 17 studies is an underestimate, one must already know or assume that there is an even stronger positive relation in the data. Thus, the prospect of non-differential misclassification in measuring talc exposure does not provide any help by itself in assessing whether talc is related to ovarian cancer.

Recall Bias

Cohort studies do not suffer from recall bias, but recall bias is an issue for case-control studies that obtain exposure information from subject interviews. Such was the case for all the case-control studies whose effects are summarized in figure 1. Recall bias can readily introduce enough bias to produce the modestly-sized overall effect (RR = 1.3) that emerges from these studies. As an example, one of us reported an association between Bendectin and congenital heart disease in 1979, with a RR of 1.6.[24] One possibility for that positive relation was recall bias, a strong consideration in light of the study design that produced the finding (the study was not designed to evaluate Bendectin, which was only an incidental finding). To resolve the issue, a second study was undertaken, this time aimed at evaluating an effect of Bendectin by eliminating recall bias using a different design.[25] The second study found a RR of 1.0, prompting the conclusion that the RR of 1.6 reported in the earlier study was due to recall bias. The amount of recall bias for Bendectin in the 1979 study amounted to an apparent effect that was much stronger than the overall effect estimate for talc and ovarian cancer in the combined studies in figure 1.

We believe that there is mixed evidence for recall bias in these studies. We base this interpretation on the few studies that examined the effect of talc separately among women who had a tubal ligation and those who did not. If recall bias were the explanation for the full effect seen in the published literature, we would predict that the effect of talc exposure would appear to be about the same for women who have a tubal ligation and those who did not, because tubal ligation is unlikely to affect recall bias. In contrast, it would likely affect any biologic action of

talc. Only three studies give information relevant to this question. In those studies, the evidence is mixed. In one study the effect of talc is greater among women who have not had a tubal ligation,[22] and in a second, talc use appeared to have no adverse effect among women who had either a hysterectomy or a tubal ligation.[23] In the third study,[2] however, there was little difference in the effect of talc for women with and without tubal ligation or hysterectomy and the effect for both groups was near null. Thus, the overall evidence on the possibility of recall bias is equivocal, with no clear answer as to whether recall bias can be eliminated as an explanation.

Confounding

Although there are some strong risk factors for ovarian cancer, for any of them to be confounding to an extent that could account for the positive relations that have been reported, they would have to be strongly correlated with talc use. Family history, ethnicity, obesity and some reproductive risk factors are positively associated with the risk of ovarian cancer, but the magnitude of these associations does not appear high enough to introduce enough confounding, even jointly, to explain completely the positive association. Of course, it remains possible that yet unidentified risk factors for ovarian cancer could be important confounders, and several such factors in the aggregate could give risk to an overall association as weak as the one between talc and ovarian cancer.

Dose-response trends

A nearly constant feature of causal relations in epidemiology and in the pathogenesis of cancer in particular is a monotonically increasing relation between measures of exposure and disease risk. Even when disease risk increases through a threshold phenomenon, progressive dose-response trends are observed because the exposure measure varies and smooths the step relation of a threshold into a gradual climb in risk. In contrast, many biases would not produce a monotonic dose-response relation. For example, Horwitz and Feinstein advanced a theory of "detection-bias" as a non-causal alternative to the theory that exogenous estrogens cause endometrial cancer.[26] According to this theory, administration of estrogens would provoke genital bleeding among some women, leading to a work up and to the diagnosis of pre-existing endometrial cancers, accounting for the observed association. This theory, however, predicted that the increase in endometrial cancer risk would be greatest for short-term users of exogenous estrogens and would decline toward no effect for longer-term users. In actual fact this inverse dose-response trend was not observed, undermining the detection bias theory.

Exposure to talc can be characterized by the age at which use started, the number of years of use, and the frequency of use (e.g., number of times per day or per week). Among the talc studies, several reported on either frequency of talc use or duration of talc use, or both. We combined the findings from these studies into a meta-regression,[27] an analysis that combines dose-specific information from various studies into a single weighted regression analysis. Each data point in a meta-regression represents one effect estimate at a given dose level; the data points are weighted by the precision of each estimate, back-calculated from the confidence interval for that estimate.

In figure 2 we show the data points and meta-regression line for frequency of talc use, and in figure 3 for duration of talc use. These regression analyses confirm the picture that one obtains from reading the individual studies (table 1): the dose-response relation across dose levels above zero for talc exposure is not increasing, but instead declines. Although

misclassification could flatten a dose-response curve, it would not produce an inverse dose-response curve. Thus, the observed pattern, whether based on individual studies or from the combined meta-regression analysis, is not consistent with a causal interpretation for talc exposure. Instead it suggests that some as yet unidentified bias accounts for the overall modest relation between talc exposure and ovarian cancer.

Figure 2
Trend in Relative Risk by Frequency of Talc Use Among Users

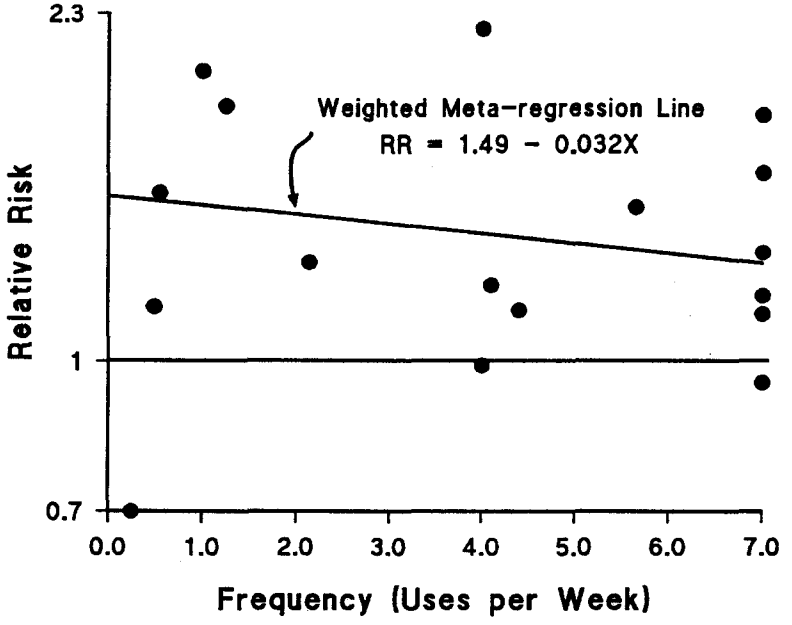


Figure 3
Trend in Relative Risk by Duration of Talc Use Among Users

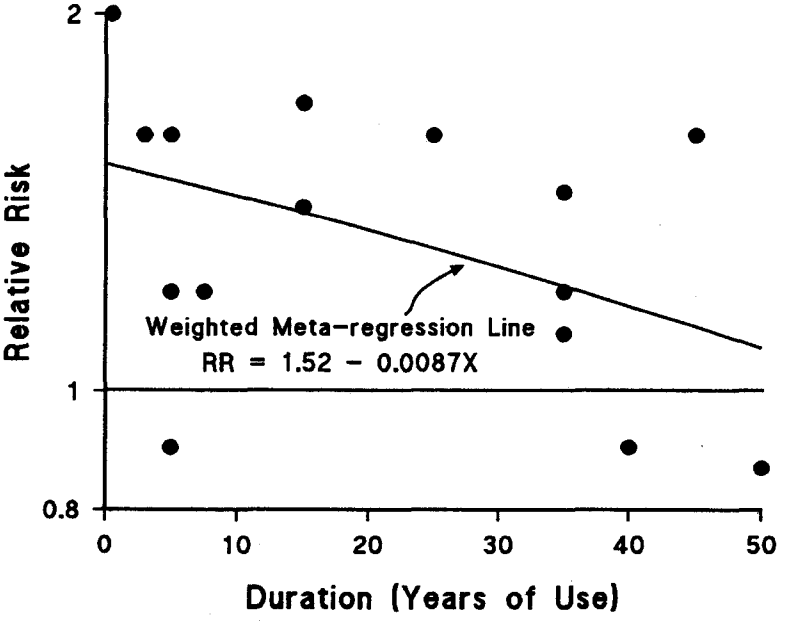


Table 1
Relative Risk Estimates of Ovarian Cancer by Frequency and Duration of Talc Use*

Citation	Frequency (Applications/wk)	Relative Risk	95% Confidence Interval
Booth et al.1989	7.00	1.30	0.80-1.90
	1.00	2.00	1.30-3.40
	0.25	0.70	0.30-1.80
Chang and Risch 1997	1.25	2.00	1.24-2.73
	4.40	1.13	0.74-1.72
	7.00	0.95	0.61-1.49
Cramer et al. 1999	4.00	2.21	1.37-3.56
	7.00	1.17	0.78-1.76
	7.00	1.57	0.80-3.10
Gertig et al. 2000	0.50	1.14	0.81-1.59
	4.00	0.99	0.67-1.46
	7.00	1.12	0.82-1.55
Harlow et al. 1992	0.55	1.50	0.80-2.70
	4.10	1.20	0.60-2.20
	7.00	1.80	1.10-3.00
Whittemore et al. 1988	2.14	1.27	0.82-1.96
	5.65	1.45	0.94-2.22

Citation	Duration (years)	Relative Risk	95% Confidence Interval
Chang and Risch 1997	15	1.70	1.09-2.64
	35	1.44	0.96-2.15
	50	0.86	0.54-1.38
Harlow et al. 1992	5	1.20	0.50-2.60
	25	1.60	1.00-2.70
	45	1.60	1.00-2.70
Ness et al. 2000	1	2.00	1.00-4.00
	3	1.60	1.10-2.30
	7.5	1.20	0.80-1.90
	35	1.20	1.00-1.50
Whittemore et al. 1988	5	1.60	1.00-2.57
	35	1.11	0.74-1.65
Wong et al. 1999	5	0.90	0.60-1.50
	15	1.40	0.90-2.20
	40	0.90	0.60-1.20

* For Open-ended Categories, the Values Assigned Assume that the Upper Category Boundary Corresponds to a Maximum Frequency Equal to Daily Use and a Maximum Duration of Use of 60 Years

Biologic Mechanism

The most plausible biological mechanism relating to the development of ovarian cancer concerns ovulation and the hormonal factors affecting it. Specifically, factors that suppress ovulation, such as gravidity, breast feeding, oral contraceptive use, tubal ligation and hysterectomy appear to reduce strongly the risk of ovarian cancer. Body mass index may also affect ovarian cancer risk. Medical conditions that may affect ovulation and also appear to increase the risk of ovarian cancer include endometriosis, ovarian cysts, and hyperthyroidism.

It does not appear plausible, however, that talc exposure has a direct effect on ovulation. If talc exposure is correlated with factors that affect ovulation, that correlation would produce confounding, as discussed above. If talc were a cause of ovarian cancer, it is presumably through a different mechanism than the many risk factors already known to affect ovarian cancer risk. There is no other evidence regarding such a mechanism, nor any clear evidence that talc applied perineally or on diaphragms makes its way physically to the ovaries. Ness et al suggest that inflammation may mediate ovarian cancer risk and that talc may play a role by causing inflammation.[17] This theory merits further investigation, although the tenability of the theory rests on the issue of whether talc particles physically reach the ovaries. Without a clear biologic mechanism for talc to cause ovarian cancer, an inference that talc does cause ovarian cancer would be an example of a "black-box" inference, meaning that the inference lacks a biologic foundation. "Black-box" inferences, such as the inference some draw that electromagnetic fields increase the risk for various cancers, are not necessarily invalid, but they are inherently more tenuous than inferences that are rooted in biologic explanations.

Conclusion

The only evidence to support a causal interpretation is the overall modest positive association seen in most of the epidemiologic studies that we have cited. The association is weak enough to be plausibly explained by unidentified bias. Recall bias is one possibility, but unidentified confounding could also readily give rise to the weak level of association that confronts us from these studies. Bias and causation are competing explanations for the weak positive association observed, an association that could be an underestimate of a stronger real association if nondifferential misclassification has diluted it. In considering these competing explanations, the lack of a plausible biologic mechanism based on the evidence to date weighs against a causal interpretation. More important, there is also positive evidence against a causal association: the inverse dose-response trend for both duration of use and frequency of use, a pattern that could not be explained by a causal relation. Based on these considerations, we suggest that the evidence to date does not indicate that talc can be "reasonably anticipated to be a human carcinogen."

References

1. Booth M, Beral V, Smith P: Risk factors for ovarian cancer: A case-control study. Br J Cancer 1989;60:592-598.

2. Chang S, Risch HA: Perineal talc exposure and risk of ovarian carcinoma. Cancer 1997;79:2396-2401.

3. Chen Y, Wu PC, Lang JH, Ge WJ, Hartge P, Brinton LA: Risk factors for epithelial ovarian cancer in Beijing, China. Int J Epidemiol 1992;21:23-29.

4. Cook LS, Kamb ML, Weiss NS: Perineal powder exposure and the risk of ovarian cancer. Am J Epidemiol 1997:145:459-465.

5. Cramer DW, Liberman RF, Titus-Ernstoff L, Welch WR, Greenberg ER, Baron JA, Harlow BL: Genital talc exposure and risk of ovarian cancer. Int J Cancer 1999; 81:351-356.

6. Cramer DW, Xu H: Epidemiologic evidence for uterine growth factors in the pathogenesis of ovarian cancer. Ann Epidemiol 1995;5:310-314.

7. Cramer DW, Welch WR, Scully RE, Wojciechowski CA: Ovarian cancer and talc. Cancer 1982;50:372-376.

8. Eltabbakh GH, Piver MS, Natarajan N, Mettlin CJ: Epidemiologic differences between women with extraovarian primary peritoneal carcinoma and women with epithelial ovarian cancer. Obstet Gynecol 1998;91:254-259.

9. Gertig DM, Hunter DJ, Cramer DW, Colditz GA, Speizer FE, Willett WC, Hankinson SE: Prospective study of talc use and ovarian cancer. J Natl Cancer Inst 2000;92:249-252.

10. Godard B, Foulkes WD, Provencher D, Brunet JS, Tonin PN, Mes-Masson AM, Narod SA, Ghadirian P: Risk factors for familial and sporadic ovarian cancer among French Canadians: A case-control study. Am J Obstet Gynecol 1998;179:403-410.

11. Green A, Purdie D, Bain C, Siskind V, Russell P, Quinn M, Ward B, and the Survey of Women's Health Study Group: Tubal sterilisation, hysterectomy and decreased risk of ovarian cancer. Int J Cancer 1997;71:948-951.

12. Hankinson SE, Hunter DJ, Colditz GA, Willett WC, Stampfer MJ, Rosner B, Hennekens CH, Speizer FE: Tubal ligation, hysterectomy, and risk of ovarian cancer. A prospective study. JAMA 1993;270:2813-2818.

13. Harlow BL, Cramer DW, Bell DA, Welch WR: Perineal exposure to talc and ovarian cancer risk. Obstet Gynecol 1992; 80:19-26.

14. Harlow BL, Weiss NS: A case-control study of borderline ovarian tumors: The influence of perineal exposure to talc. Am J Epidemiol 1989;130:390-394.

15. Hartge P, Hoover R, Lesher LP, McGowan L: Talc and ovarian cancer. JAMA 1983;250:1844.

16. Hartge P, Stewart P: Occupational and ovarian cancer: A case-control study in the Washington, DC, metropolitan area, 1978-1981. J Occup Med 1994;36:924-927.

17. Ness RB, Grisso JA, Cottreau C, Klapper J, Vergona R, Wheeler JE, Morgan M., Schlesselman JJ: Factors related to inflammation of the ovarian epithelium and risk of ovarian cancer. Epidemiology 2000;11:111-117.

18. Purdie D, Green A, Bain C, Siskind V, Ward B, Hacker N, Quinn M, Wright G, Russell P, Susil B: Reproductive and other factors and risk of epithelial cancer: An Australian case-control study. Survey of Women's Health Study Group. Int J Cancer 1995;62:678-684.

19. Rosenblatt KA, Szklo M, Rosenshein NB: Mineral fiber exposure and the development of ovarian cancer. Gynecol Oncol 1992;45:20-25.

20. Shushan A, Paltiel O, Iscovich J, Elchalal U, Peretz T, Schenker J: Human menopausal gonadotropin and the risk of epithelial ovarian cancer. Fertil Steril 1996;65:13-18.

21. Tzonou A, Polychronopoulou A., Hsieh CC, Rebelakos A, Karakatsani A, Trichopoulos D: Hair dyes, analgesics, tranquilizers and perineal talc application as risk factors for ovarian cancer. Int J Cancer 1993;55:408-410.

22. Whittemore AS, Wu ML, Paffenbarger RS, Sarles DL, Kampert JB, Grosser S, Jung DL, Ballon S, Hendrickson M: Personal and environmental characteristics related to epithelial ovarian cancer. II. Exposures to talcum powder, tobacco, alcohol, and coffee. Am J Epidemiol 1988;128:1228-1240.

23. Wong C, Hempling RE, Piver MS, Natarajan N, Mettlin CJ: Perineal talc exposure and subsequent epithelial ovarian cancer: A case-control study. Obstet Gynecol 1999;93:372-376.

24. Rothman KJ, Fyler DC, Goldblatt A, Kreidberg MB: Exogenous hormones and other drug exposures of children with congenital heart disease. Am J Epidemiol 1979;109:433-439.

25. Zierler S, Rothman KJ: Congenital heart disease in relation to maternal use of Bendectin and other drugs in early pregnancy. N Engl J Med 1985;313:347-352.

26. Horwitz RI, Feinstein AR: Alternative analytic methods for case-control studies of estrogens and endometrial cancer. N Engl J Med. 1978;299:1089-94.

27. Maclure M: Demonstration of deductive meta-analysis: ethanol intake and risk of myocardial infarction. Epidemiol Rev. 1993;15:328-51.

Appendix

Literature Search Methodology

The literature search was designed to find published epidemiologic studies specifically relating to the perineal use of non-asbestiform talc. The 2000 NTP Draft Report was used as the initial resource to locate applicable studies. To identify other relevant publications, an on-line search was performed in Dialog and using the internet. In addition, medical and scientific resources such as Medline, Toxline, and SciSearch were queried using various keyword terms including "talc," "non-asbestiform," "ovarian cancer," and "perineal." The search was limited to papers published after 1980, because asbestiform products were removed from the market in 1976. Once relevant articles were obtained, bibliographies were "tree-searched" to identify other applicable studies that may have been omitted during the on-line search. "Tree-searching" involves reading an article's bibliography, and then identifying citations that may contain appropriate information based on the title or author. "Tree-searching" identified early studies or those not recorded in on-line databases.

Non-asbestiform Talc and the Risk of Epithelial Ovarian Cancer.

A statement submitted to the Cosmetics, Toiletries and Fragrance Association (CTFA) for presentation to the NTP Board of Scientific Counselors.

December 1, 2000

Samuel Shapiro, MB, FRCP(E)
Emeritus Director, Slone Epidemiology Unit, Boston University School of Public Health
Visiting Professor of Epidemiology, Columbia University School of Public Health

EXECUTIVE SUMMARY

In response to NTP's consideration of non-asbestiform talc as reasonably anticipated to be a carcinogen to the human ovary, a critical review of the epidemiologic evidence was conducted in the context of its potential for carcinogenicity. Associations identified in 16 case-control studies and one cohort study are overall of very low magnitude. In addition, there is strong evidence to suggest bias and confounding in the published studies. These alternative explanations have not been ruled out. Further, the application of standard criteria for causation to the data set as a whole indicates that none of these criteria are satisfied. Particularly troublesome are the inconsistencies seen across studies in subgroup analyses including different modes of exposure and different cancer outcome subtypes. Further, dose-response effects in terms of frequency or duration of exposure in those studies where the information was available have generally not been evident; in fact, much of the data contradict any dose response effects. Lastly, pending the opinion of expert toxicologists there is no plausible biological evidence to support the hypothesis that non-asbestiform talc increases the risk of ovarian cancer.

It is concluded that the epidemiologic data on non-asbestiform talc and ovarian cancer do not satisfy standard criteria for causality, and that classification of non-asbestiform talc as "reasonably anticipated to be a human carcinogen" is not supported by the epidemiological evidence.

INTRODUCTION

The National Toxicology Program (NTP) has submitted a draft report for consideration by its Board of Scientific Counselors entitled "Report on Carcinogens. Background Document for Talc: Asbestiform and Non-Asbestiform." In the Summary Statement (pages iii and iv), the following claim is made:

> The use of talc for perineal dusting on sanitary napkins and diaphragms has been associated with ovarian cancer. Fourteen of 16 case control studies of human ovarian cancer provided evidence for an association with the use of talc (presumably cosmetic grade, but information on fibrous content is lacking). A recent large prospective cohort study did not demonstrate an overall increase in risk for ovarian cancer with talc use (Gertig et al. 2000). However, in this study talc use was significantly associated with one subtype of ovarian cancer, invasive serous ovarian cancer. Risk of this tumor type was also elevated in several case-control studies (Harlow et al. 1992, Chang and Risch 1997, Cook et al. 1997, Wong et al. 1999, and Cramer et al. 1999). There is conflicting evidence concerning transport of talc through the genital tract to the ovary (Hamilton et al. 1984). Several studies provided evidence that factors preventing translocation of talc to the ovary, such as tubal ligation or hysterectomy, reduce the risk associated with talc use (Harlow et al. 1992, Whittemore et al. 1988, Cramer et al. 1999). Risk of ovarian cancer associated with talc use is unlikely to be a consequence of confounding or other biases.

Taken together, the findings are adduced as constituting sufficient evidence to classify non-asbestiform talc as a substance reasonably anticipated to be a carcinogen to the human ovary.

I have been asked by the CTFA to evaluate the epidemiologic evidence concerning the use of non-asbestiform talc in relation to the risk of epithelial ovarian cancer. To that end I have reviewed the studies referred to in the Draft Report, together with other relevant material. Based on my evaluation, I conclude that there are no valid epidemiologic data to implicate non-asbestiform powders in the etiology of epithelial ovarian cancer. In addition, the published epidemiologic evidence concerning elevated risks for all talc exposures, in whatever form or composition, is of questionable validity. Even the study that initially generated the hypothesis that exposure to talc may increase the risk of ovarian cancer (Cramer et al. 1982) is of questionable validity.

Contrary to what is stated in the NTP Report, bias and confounding are not only possibilities, but the published studies contain data to suggest that they are likely. The studies also suggest that several biases are generally present, and that they tend to be in the same direction in all the studies that have not yielded null findings. Moreover, the magnitude of the reported statistically significant overall relative risk (RR) estimates has been low—in all instances 1.9 or less, and mostly 1.5 or less. For such low estimates, epidemiologic methods are seldom, if ever, capable of making the distinction between bias, confounding, and causality (Shapiro 2000).

As it is beyond the scope of this evaluation to undertake a study by study review of the evidence, I have prepared a global critique. Before commencing that critique, however, it is necessary to point out two major, and critical, errors in the Draft Report.

First, contrary to what is stated in the Report, there are virtually no data that associate the use of diaphragms (including diaphragms known to have been stored in talc powder) with an increased risk of ovarian cancer. The findings have been so uniformly negative that Cramer et al. (1999), in their most recent case control study elected not to collect data on diaphragm use on the grounds that earlier studies had established that such exposure has been shown not to increase the risk of ovarian cancer (see also below: selection bias).

Second, contrary to the claim that the epidemiologic studies have evaluated exposure to talc that is "presumably cosmetic grade," (*i.e.,* non-asbestiform) there are at most only indirect and imprecise data on the risk associated with non-asbestiform talc. That claim is only inferred, based on exposure that took place after 1976. However, supplies of talc-containing powders may commonly have been stored before they were sold, after which they may then have been used for appreciable periods of time before fresh supplies were purchased. Thus the inference that after 1976, exposure was to non-asbestiform talc, may not be justified. All that can be assumed is that at some unknown time after 1976, the ratio of the use of asbestiform to non-asbestiform talc presumably declined.

CRITIQUE

Below, the assembled epidemiologic evidence to implicate talc in the etiology of ovarian cancer is evaluated according to standard criteria used to assess causality in epidemiologic research (Hill 1965, Susser 1991). The criteria are not mutually exclusive, but to justify any causal inference a reasonable combination of them must be present.

Temporality

An absolute requirement for causal inference is that the exposure must antedate the onset of the illness. For ovarian cancer, it is impossible to determine the time of onset with any precision. In that circumstance, lag-time analysis must be used in order to ensure that all exposures that are assessed have a strong likelihood of having taken place before the disease commenced. Exposures in the distant past have been assessed in the different studies, but they have mostly been ignored in the evaluation of dose response effects, in terms either of duration or frequency of use. Lower abdominal symptoms, such as vaginal bleeding, could lead to the use of powders. That is, ovarian cancer could sometimes have "caused" talc exposure, rather than the reverse. For the estimation of overall risk, only one case-control study (Cook et al. 1997), and two follow-up studies (Hankinson et al. 1993, Gertig et al. 2000) have clearly specified the exposures in an effort to establish an unambiguous temporal sequence of events.

Strength (Magnitude) of the Association

In observational studies, bias and confounding can never be entirely excluded as possible explanations for an observed association. However, if in any given study, the risk among the exposed is increased many-fold relative to the non-exposed (high relative risk estimate), and if the methods used in the research are reasonably adequate, it is likely that even if plausible sources of bias could be eliminated, the RR would remain elevated. For example, the relative risk of lung cancer among heavy smokers has been shown to be increased some 30-fold in countless studies. Some of the studies (including the original pioneering work by Doll and Hill (1950) that was among the first to document the association) are known not to have been free of bias (Doll and Hill 1964, Doll and Peto 1976). However, there are no plausible biases that could conceivably have accounted for a 30-fold increase in the risk. Had it been possible to avoid the various biases, doing so would have had (at most) only a minor effect on the magnitude of the RR—a strong if slightly attenuated association would still have been present.

For weak associations (*i.e.*, RRs of 2.0 or less), the situation is entirely different: minor sources of bias and confounding may readily account for them. If all sources of bias could be avoided, there may be no association at all. For this reason, making a confident distinction between bias and causality is almost invariably beyond the resolving power of observational research (Shapiro 2000). For RRs that are considerably below 2.0 (say, 1.5 or less), it is virtually impossible to do so (Shapiro 2000).

In the case-control studies of talc exposure and ovarian cancer, the finding that first gave rise to the hypothesized association with ovarian cancer was a statistically significant overall RR estimate of 1.92 (Cramer et al. 1982). In one subsequent study, it was 1.60 (Cramer et al. 1999). In the remaining studies, all statistically significant overall RR estimates were 1.6 or less (Chang and Risch 1997, Cook et al. 1997, Cramer and Xu 1995, Green et al. 1997, Ness et al. 2000, Purdie et al. 1995). This general trend of declining RR is strongly suggestive of the well-known phenomenon of regression to the mean and it favors chance as an explanation for the initial hypothesis-generating association (Cramer et al. 1982).

A meta-analysis combining the various studies has been published in which the summary overall RR estimate was 1.3 (Gross and Berg 1995). The validity of the application of meta-analysis to observational data is highly questionable (Shapiro 1997). In addition, as is the case here, when there is marked heterogeneity among the studies in terms of their methods (see below: consistency) and the definition of the outcomes and the exposures studied, as well as in the confounding factors that were taken into account, it is generally accepted that findings derived from meta-analysis are not only questionable, but uninterpretable (Shapiro 1997, Greenland 1994). In the present case, that lack of interpretability is compounded by the low magnitude of the summary RR estimates.

Some have argued that the fact that the reported risk estimates, although small, have been positive in 14 out of 16 case control studies, as well as in a meta-analysis (which, incidentally, is not independent of the 16 studies), points to causality. That argument assumes that bias and confounding did not tend to be in the same direction across the studies. Yet in observational research the same biases are commonly present, and in the same direction, in more than one study. In addition, as explained below, there is evidence in the published studies to suggest that this was the case (see below: selection bias; information bias).

In addition to overall RRs derived from analyses of the total data in the various studies, analyses of subgroups have also been performed. Some of the findings in the subgroups have been invoked as evidence to support causality. For example, as mentioned above, in their initial hypothesis-generating study, Cramer et al. (1982) reported an overall RR of 1.92 (95% confidence interval, 1.27-2.89) for women exposed to perineal dusting, or to talc applied to sanitary napkins, or to both. That estimate was based on 92 exposed cases and 61 exposed controls. They also examined multiple subgroups, in one of which (exposure to perineal dusting *in combination* with exposure to sanitary napkins dusted with talc, and *in combination* with exposure to diaphragms stored in talc) the RR was 3.28 (1.68-6.42). In that subgroup analysis there were only 32 exposed cases and 13 exposed controls. Moreover, as indicated by the 95% confidence intervals, the RR of 3.28 was statistically compatible with the overall and best RR estimate of 1.92. The "blip" of 3.28 was identified in the course of multiple comparisons in which "significant" associations would have been expected to occur by chance.

The procedure of exploring multiple subsets of data in the search for "statistically significant" associations is known as "data dredging," and it is not valid unless there are well formulated hypotheses, specified *a priori*.

Other examples of subgroup analyses abound in the published studies. Thus, in some instances there have been overall associations with all types of ovarian cancer (*e.g.*, Green et al. 1997). In others overall associations have not been identified (*e.g.*, Hartge et al. 1983), or they have been weak and nonsignificant (*e.g.*, Harlow and Weiss 1989). Instead, causality has been suggested based on subgroup analyses carried out in the search for elevated RRs. Some "significant" associations, virtually all of them with RRs below 1.5, have been identified only with specific subtypes of epithelial ovarian cancer. Again, the use of multiple comparisons in order to "dredge" for positive associations is not valid; inevitably, some will turn up by chance. In addition, as a separate issue discussed below (see below: consistency), the associations with the specific subtypes have been inconsistent across the studies.

In summary, the overall associations identified in the published studies have almost without exception been of low magnitude. Even the associations identified in subgroups have generally been of low magnitude. None of the studies satisfy the causal criterion of high magnitude associations.

Statistical Stability

The criterion of statistical stability requires that data derived from relatively unbiased and unconfounded studies should be sufficiently robust so that chance, as an alternative explanation of any given association, is only a remote possibility. That requirement, however, is not independent of the requirement that the association should also be of sufficient magnitude to ensure that bias and confounding are also remote possibilities. Weak associations simply impose the need for greater numbers in order to achieve "statistical significance." Thus, biased data, if plentiful enough, can become "robust", and "statistically significant." In the face of the low magnitude of the associations that have generally been observed, the "statistical significance" of the reported associations cannot be invoked to support causality.

Dose-Response Effects

If talc increases the risk of ovarian cancer there should be evidence to suggest a dose-response effect either in terms of frequency or duration of exposure, or both. Yet in the preponderance of the studies, when possible dose-response effects have been looked for, they have not been evident. Moreover, much of the data have contradicted any dose-response effects. For example, Booth et al. (1989) reported RRs of 0.7, 2.0, and 1.3 for monthly, weekly, and daily exposure, respectively. Under causal assumptions, it is not plausible that daily use would carry a lower risk than weekly use. Or to give another example, Cramer et al. (1999) reported RRs of 1.84, 1.43, and 1.43, respectively, for

lifetime total talc applications of <3,000, 3,000-10,000, and >10,000; again it is not plausible that the smallest number of applications would carry the highest risk.

Virtually the only instances in which suggestions of dose-response effects have been found were again in subgroup analyses; for example, among women classified by frequency of exposure after the exclusion of exposures during pregnancy, oral contraceptive use, and following sterilization (Cramer et al. 1999). Such *post hoc* analysis of the data is not valid, and it is again based on multiple comparisons.

Consistency

An inference of causality is supported if the findings among several studies, carried out by different investigators who use different strategies, nevertheless converge on the same relatively invariant associations. For talc exposure, the studies have been markedly inconsistent. Some have reported statistically significant overall associations between talc use and invasive cancer (Chang and Risch 1997, Cramer at al. 1999, Gertig et al. 2000), while another study found an association between talc use and tumors of low malignant potential (Harlow et al. 1992). Other studies have not reported significant overall associations, but only associations within subgroups classified according to cancer subtypes: sometimes the associations have been with serous tumors (Cook et al. 1997, Cramer et al. 1999, Gertig et al. 2000), sometimes with endometrioid tumors (Harlow et al. 1992). Even Cramer et al, the originators of the talc hypothesis, have published contradictory findings. Their original study, conducted in 1982, found that histologic characteristics of tumors developing in women with perineal exposure to talc did not differ significantly from those in women without perineal exposure to talc. In contrast, their subsequent study of 1999 found a statistically significant association between talc use and serous invasive cancer.

The studies have also been inconsistent in terms of other subgroup associations. Thus the great majority of the findings have suggested that the use of diaphragms stored in talc does not increase risk. With regard to other routes of exposure, sometimes the strongest associations have been with perineal dusting (Cook et al. 1997, Eltabbakh et al. 1998, Green et al. 1997, Harlow et al. 1992), sometimes with powder application to sanitary napkins (Ness et al. 2000, Rosenblatt et al. 1992), and sometimes with applications that have not been further described (Cramer et al. 1999, Cramer and Xu 1995).

There has also been no consistent pattern across the relevant studies that would make clinical or biological sense. For example, it has been proposed that sexual intercourse might facilitate the migration of talc up the female genital tract (Cramer et al. 1999). In that case, the highest risk, surely, might be for diaphragm use; yet none has been found. Remarkably, in the face of this striking inconsistency, Cramer et al. (1999) elected not to study diaphragm use in their last study. When testing a hypothesis, it is not valid to exclude factors that constitute genuine exposures even if their association with ovarian cancer has been negative in other studies (see also below: information bias). The next highest risk might be for regular perineal dusting, an association that has been repeatedly sought, but only inconsistently found. By the same reasoning a somewhat weaker

association might be expected for application of talc to sanitary napkins, as they would usually only be used for four or five days per month. However, the RRs for those exposures have commonly been higher than for regular applications of talc to the perineum. Finally, it might be expected that the RRs would be lowest for generalized occasional body exposure to talc, or for application to unspecified sites; yet in some studies the risk associated with such applications has been higher than for more intense applications (see above: dose-response effects).

In summary, the studies have not only failed to satisfy the criterion of consistency, but if anything, they have commonly revealed trends that were opposite to what might reasonably have been expected, under causal assumptions, as dose-response effects.

Systematic Bias

As mentioned above, bias can never be entirely ruled out in observational research, and in the face of RR estimates below 2.0 its possible existence, even when studies are well performed, limits interpretability. That interpretability is even further limited when, as in the present instance, there is strong evidence to suggest the presence of bias. The role of selection bias and information bias is discussed below.

Selection Bias

Selection bias exists when, on the null, the cases and noncases selected for study are not independent of the exposure of interest. In the majority of the case-control studies, the response rates have been low, generally below 70% among the cases, and seldom much higher among the controls (and probably lower still in those studies that recruited controls through random digit dialing).

It is generally accepted that possible selection bias becomes a major concern when enrollment rates in the targeted population are low. For example, among the cases, talc use could have been markedly different among women with invasive cancer, compared to those with tumors of low malignant potential. Specifically, it is reasonable to assume that women who are hygiene conscious would preferentially use powders, and undergo more frequent gynecological examinations, so that borderline tumors might be preferentially diagnosed among powder users. For the same reason, even among women with invasive cancer, the diagnosis might be made earlier among users than among nonusers. Thus cases who are users could fall inside the time frame of the study, while nonusers may only come to diagnosis after the study is completed.

The same problems are relevant to low recruitment rates among targeted controls. For example, hygiene-conscious women would tend to use talc more frequently than non-hygiene-conscious women. If there is a difference in the recruitment rates according to hygiene consciousness, this could bias the results.

Perhaps the most egregious example of selection bias was the original study of Cramer et al. (1982) that generated the hypothesis. In that study, only 45% of the eligible controls, and 72% of the eligible cases participated. For the cases it might be argued that an appreciable proportion had died, and that their pattern of powder use was unlikely to be different from that of the survivors. That argument is speculative, and open to question. Moreover, with only a 45% recruitment rate among the controls, no hypothesis was justified in the first place, and the findings must be categorically rejected.

The only way to confidently avoid selection bias due to under-enrollment would be to recruit 100%, or close to 100%, of all targeted cases or controls in a study base, specified *a priori*. With ovarian cancer a high recruitment rate is difficult to accomplish, but difficulty is not a criterion of causality. In the face of the poor recruitment rates in the majority of the case control studies, it is simply not possible to claim that the associations are unlikely to be accounted for by selection bias. It should also be noted that exactly the same selection biases would operate in the identification of cases and noncases in a cohort study.

Finally, possible selection bias was not adequately assessed in the published studies. Firstly, as explained above, borderline tumors should have been excluded, since their inclusion was much too likely to be dependent on life style factors, including the use of talc-containing powders. Secondly, as also explained, even invasive tumors may selectively have been included in the various studies because of earlier diagnosis among powder users. This possibility could have been assessed, at least in part, by the evaluation of risk stratified according to the staging of the cancer. Under causal assumptions the association should then have been evident even among the most advanced cases that would inevitably have come to diagnosis without any further possibility of delay. An assessment according to stage has not been done.

Further strong evidence to support the likelihood of selection bias is that in two successive analyses of follow-up data from the Nurses Health Study, there was no overall association with the use of talc (Hankinson et al. 1993, Gertig et al. 2000). Follow-up was for more than 20 years, and at most, only a few cases could have been missed during the last year or two of follow-up.

In summary, given the multiple potential sources of selection bias present among virtually all the studies, selection bias could readily have accounted, partially or wholly, for the statistically significant overall relative risks of 1.9 or less that have been observed in the various studies. Given the low recruitment rate in the original hypothesis-generating study, even that hypothesis itself was based on unsatisfactory data. Contrary to what has been stated by some authors, selection bias is not only possible in all the studies that reported positive associations, but likely.

Information Bias

Information bias exists when, under the null, the recording of exposure status is not independent of status as a case or a noncase. Following the initial hypothesis-generating study in 1982, the possibility that talc-containing powders may increase the risk of ovarian cancer was given extensive and repeated publicity. Thus, the strong likelihood is that the cases would repeatedly have probed their memories (as well as have had their memories probed for them by medical attendants) in order to remember every possible occasion when they had used talc-containing powders. They may also have tended to overestimate the duration or frequency of use, especially use that took place years previously. Healthy control women, by contrast, would not have been motivated to remember with the same intensity.

There is strong evidence in the published data to suggest that there was major information bias. Some of that evidence has already been mentioned. Thus, the absence of an association with diaphragms stored in talc could be explained by a lack of awareness among the cases that the use of diaphragms could have represented exposure to talc. In addition, the paradoxical data, in multiple studies, in which RRs were higher for short duration than for long duration exposures, and higher for a small number of talc applications than for a large number, also strongly support the likelihood of information bias. Indeed, it is difficult to conceive of another explanation, and information bias is a far more plausible explanation of the patterns that have been observed than possible causality.

Further strong evidence to support the likelihood of information bias is that no overall association was observed in two successive analyses of follow-up data in the Nurses Health Study (Hankinson et al. 1993, Gertig et al. 2000). The exposures were recorded before the cancers were diagnosed, and these data had the considerable advantage of being free of information bias. Only one positive association was observed in this study, a statistically significant RR of 1.4 for serous tumors. Again, however, that association was discovered in the course of multiple comparisons.

Confounding

Confounding exists when a factor is associated both with the exposure, and independently, with the outcome. The known risk factors for ovarian cancer include infertility or low parity (or both), a family history of breast or ovarian cancer (BRCA1), history of tubal ligation or hysterectomy, and probably, high socioeconomic status (Purdie et al. 1995). Several of the published studies have not adequately allowed for these factors. Factors such as socioeconomic status, for example, could both have been determinants of talc use, and independently, of the risk of ovarian cancer.

Coherence

The criterion of coherence requires that the findings should be broadly consistent with other epidemiologic data on talc exposure, including occupational exposure. Hartge and Stewart 1994 have evaluated the risk of ovarian cancer among persons occupationally exposed to talc; no associations were found. Heller et al. 1996 found no difference in the number of talc particles present in the ovaries of ovariectomized women who did and did not regularly use talc.

Biological Plausibility

A causal inference would receive some support if there were experimental data in animals to suggest that talc applied to the perineum travels up the female genital tract, and then increases the incidence of ovarian cancer. Again I am not qualified to judge, and I must defer to my colleagues in toxicology. However, my understanding is that the only evidence to hint at any carcinogenicity is one study referred to in the Draft Report that documented the occurrence of tumors in the rat lung and adrenal gland. There is no evidence that talc is carcinogenic to the ovary. There is also no experimental evidence that talc applied to the perineum even reaches the ovaries, and there are animal (monkey) data to refute this hypothesis (Wehner et al. 1986).

CONCLUSIONS

I have reviewed the epidemiologic studies of the risk of epithelial ovarian cancer in relation to exposure to talc. When assessed against the standard criteria of causality, none of the positive associations satisfy any of them. All of the associations have been of low magnitude; plausible and even likely sources of bias have not been ruled out. In addition, the more valid studies, the follow-up studies in particular, do not suggest that talc increases the overall risk of ovarian cancer. Subgroup associations have not been consistent, and the analyses that have given rise to those associations were not valid.

Contrary to what is stated in the Draft Report, the published evidence does not meet the standard required to classify talc as being reasonably anticipated to be a carcinogen.

REFERENCES

Booth, M., Beral, V., and Smith, P. 1989. Risk factors for ovarian cancer: A case-control study. Br J Cancer 60:592-598.

Chang, S. and Risch, H.A. 1997. Perineal talc exposure and risk of ovarian carcinoma. Cancer 79(12):2396-2401.

Cook, L.S., Kamb, M.L., and Weiss, N.S. 1997. Perineal powder exposure and the risk of ovarian cancer. Am J Epidemiol. 145:459-465.

Cramer, D.W., Liberman, R.F., Titus-Ernstoff, L., Welch, W.R., Greenberg, E.R., Baron, J.A., and Harlow, B.L. 1999. Genital talc exposure and risk of ovarian cancer. Int J Cancer 81(3):351-356.

Cramer, D.W. and Xu, H. 1995. Epidemiologic evidence for uterine growth factors in the pathogenesis of ovarian cancer. Ann Epidemiol. 5:310-314.

Cramer, D.W., Welch, W.R., Scully, R.E., and Wojciechowski, C.A. 1982. Ovarian cancer and talc. Cancer 50:372-376.

Doll, R. and Hill, A.B. 1964. Mortality in relation to smoking: 10 years' observations of British doctors. Br Med J. 5395:1399-410.

Doll, R. and Hill, A.B. 1950. Smoking and carcinoma of the lung. Br Med J. 2:739.

Doll R and Peto R. 1976. Mortality in relation to smoking: 20 years' observations on male British doctors. Br Med J. 2(6051):1525-36.

Gertig, D.M., Hunter, D.J., Cramer, D.W., Colditz, G.A., Speizer, F.E., Willett, W.C., and Hankinson, S.E. 2000. Prospective study of talc use and ovarian cancer. J Natl Cancer Inst. 92:249-252.

Green, A., Purdie, D., Bain, C., Siskind, V., Russell, P., Quinn, M., Ward, B., and the Survey of Women's Health Study Group. 1997. Tubal sterilisation, hysterectomy and decreased risk of ovarian cancer. Int J Cancer 71:948-951.

Greenland, S. 1994. A Critical Look at Some Popular Meta-Analytic Methods. Am J Epidemiol. 140(3): 290-296.

Gross, A.J. and Berg, P.H. 1995. A meta-analytical approach examining the potential relationship between talc exposure and ovarian cancer. J Expo Anal Environ Epidemiol. 5(2):181-95.

Hamilton, T.C., Fox, H., Buckley, C.H., Henderson, W.J., and Griffiths, K. 1984. Effects of talc on the rat ovary. Br J Exp Pathol. 65:101-106.

Hankinson, S.E., Hunter, D.J., Colditz, G.A., Willett, W.C., Stampfer, M.J., Rosner, B., Hennekens, C.H., and Speizer, F.E. 1993. Tubal ligation, hysterectomy, and risk of ovarian cancer. A prospective study. JAMA. 270(23):2813-2818.

Harlow, B.L., Cramer, D.W., Bell, D.A., and Welch, W.R. 1992. Perineal exposure to talc and ovarian cancer risk. Obstet Gynecol. 80(1):19-26.

Harlow, B.L. and Weiss, N.S. 1989. A case-control study of borderline ovarian tumors: The influence of perineal exposure to talc. Am J Epidemiol. 130(2):390-394.

Hartge, P., Hoover, R., Lesher, L.P., and McGowan, L. 1983. Talc and ovarian cancer. JAMA. 250(14):1844.

Heller DS, Westhoff C, Gordon RE, and Katz N. 1996. The relationship between perineal cosmetic talc usage and ovarian talc particle burden. Am J Obstet Gynecol. 174(5):1507-10.

Hill, A.B. 1965. The environment and disease: association or causation? Proc R Soc Med. 58:295-300.

Ness, R.B., Grisso, J.A., Cottreau, C., Klapper, J., Vergona, R., Wheeler, J.E., Morgan, M., and Schlesselman, J.J. 2000. Factors related to inflammation of the ovarian epithelium and risk of ovarian cancer. Epidemiology. 11(2):111-117.

Purdie, D., Green, A., Bain, C., Siskind, V., Ward, B., Hacker, N., Quinn, M., Wright, G., Russell, P., and Susil, B. 1995. Reproductive and other factors and risk of epithelial cancer: An Australian case-control study. Survey of Women's Health Study Group. Int J Cancer 62(6):678-684.

Shapiro, S. 2000. Bias in the evaluation of low magnitude associations: An empirical perspective. Am J Epidemiol. 151(10):939-45.

Shapiro, S. 1997. Is meta-analysis a valid approach to the evaluation of small effects in observational studies? J Clin Epidemiol. 50(3):223-229.

Susser, M. 1991. What is a cause and how do we know one? A grammar for pragmatic epidemiology. Am J Epidemiol. 133(7): 635-48.

Wehner A.P., Weller R.E., and Lepel E.A. 1986. On talc translocation from the from the vagina to the oviducts and beyond. Food Chem Toxicol. 24(4):319-38.

Whittemore, A.S., Wu, M.L., Paffenbarger, R.S., Sarles, D.L., Kampert, J.B., Grosser, S., Jung, D.L., Ballon, S., and Hendrickson, M. 1988. Personal and

environmental characteristics related to epithelial ovarian cancer. II. Exposures to talcum powder, tobacco, alcohol, and coffee. Am J Epidemiol. 128(6):1228-1240.

Wong, C., Hempling, R.E., Piver, M.S., Natarajan, N., and Mettlin, C.J. 1999. Perineal talc exposure and subsequent epithelial ovarian cancer: A case-control study. Obstet Gynecol. 93(3):372-376.

1 Dana Road
Valhalla, New York 10595
Telephone: (914) 592-2600
FAX: (914) 592-6317

American Health Foundation

EXECUTIVE SUMMARY

Re: Solicitation of Public Comment to NTP Board of Scientific Counselors; Nomination of NonAsbestiform Talc as "reasonably anticipated to be a human carcinogen in the 10th Report on Carcinogens."

The NTP reviewed 16 case-control studies and one cohort study that reported relative risk estimates of ovarian cancer associated with cosmetic talc use. These studies are critically reviewed in the Appendix. The majority of these studies were not specifically designed to test the talc hypothesis, but included at least one question on ever having applied cosmetic talc to the perineum. The first report on this association came from a community-based case-control study (Cramer et al. 1982). All but three of the subsequent studies were based on a similar case-control design, except for two hospital-based case-control studies (Rosenblatt et al. 1992, Wong et al. 1999) and one cohort study (Gertig et al. 2000). A number of methodologic and interpretational concerns have been raised in these studies that would mitigate against either a causal interpretation (Muscat et al. 1998) or a "reasonably anticipated" conclusion.

These include:

1. The actual biological 'exposure' of ovarian tissue to talc has been measured by poorly defined proxy measures. The assumption that ovarian tissue is contaminated with talc from perineal dusting is based on putative physiological actions including penetration of the female reproductive tract from external talc dust, and translocation of talc particles to the ovaries. This is supported by studies showing no relationship between cosmetic talc burden in healthy ovarian tissue and lifelong perineal talc dusting. In contrast, the use of talc-dusted diaphragms provides a more valid exposure measurement for testing the hypothesis that talc burden in the ovaries of women with ovarian cancer is greater than the talc burden in the ovaries of woman with healthy ovaries. The use of talc-dusted diaphragms, by definition, results in female tract exposure to talc. Considering the anatomy of the female reproductive tract, it is difficult to describe mechanisms by which an inert particle such as talc is able to translocate to the ovaries under non-experimental conditions. A meta-analysis of the association between talc-dusted diaphragm usage and ovarian cancer risk resulted in a summary odds ratio of 0.79 (95% CI 0.46, 1.38).

2. Recall bias in questions on "ever versus never" using perineal talc exposure might have resulted in spurious associations. Post-operative side effects of radiation treatment of malignant ovaries might include perineal skin rash and irritation, leading to increased talc use. The four case-control studies based on case interviews that were conducted in the hospital immediately after diagnosis found little or no association with perineal talc use. In studies with positive associations, recall bias might also have occurred due to heightened awareness of the hypothesis in case subjects. It has not been demonstrated that recall bias can be ruled out as an alternate explanation for positive findings.

3. The validity and the reliability of questions on perineal talc dusting exposure have never been measured. The validation of exposure measures is a necessary component of proper epidemiologic research.

4. In most studies, there was no trend in the odds ratios associated with increasing duration, frequency or cumulative perineal talc exposure. In many of the studies with dose information, the risk was lower with the highest lifetime exposures relative to the lowest lifetime exposures. These epidemiologic findings are inconsistent with known carcinogenic mechanisms.

5. There are conflicting findings on ovarian cancer risk associated with perineal talc exposure in sterilized women compared to women with intact reproductive organs.

6. Publication bias, in the form of incomplete presentation of numerical study results, needs to be addressed as a limitation of the epidemiologic studies being evaluated here.

The following describes in greater detail the outline presented in the Executive Summary.

1. Testing the talc hypothesis using different epidemiologic measures

The presence of talc dust in ovarian cancer tissue was first reported in 1971 (Henderson et al. 1971). The observation of a high rate of ovarian cancer in World War II female gas mask assemblers exposed to asbestos (Berry et al. 2000), a fiber with a similar chemical but not morphologic composition to talc, helped raise the hypothesis that cosmetic talc may cause ovarian cancer. It currently seems questionable if asbestos can induce ovarian malignancy. Malignant mesothelioma in the peritoneal cavity, which is believed to be caused by asbestos (Price 1997), can present as ovarian cancer. Even using modern diagnostic methods, peritoneal mesothelioma is difficult to distinguish from ovarian masses (Clement et al. 1996, Sato et al. 2000). Asbestos fiber burden in healthy ovaries was found to correlate with reported asbestos exposure, although comparisons with malignant tissue are unavailable (Heller et al. 1999).

Heller et al. (1996) determined talc particle counts in ovarian specimens from 24 women undergoing oophorectomy and compared these counts to reported history of talc dusting. No relationship was found between cosmetic talc burden in healthy ovarian tissue and lifelong perineal talc dusting. Although the lack of an association in the Heller et al. findings could be due to contamination or reflect nonuniform distribution of talc particles in ovarian tissue, this recent study raises questions over whether reported associations between perineal talc exposure and ovarian tumors in case-control studies reflects a carcinogenic action of talc. The validity of these epidemiologic associations has also been questioned because it is unknown whether talc dust in the perineal area can actually penetrate the female reproductive tract, and then translocate to the ovaries against physiological forces working in the opposite direction.

Although the epidemiologic literature has focused primarily on external perineal exposure, the talc hypothesis would appear to be tested with greater precision and validity by questions on the use of talc-dusted diaphragms. The use of talc-dusted diaphragms, by definition, results in female reproductive tract exposure to talc. In fact, experimental translocation of talc particles to the ovary in women was achieved with deliberate or inadvertent manipulation of patients in the supine position (as cited in Wehner 1998). Although data on the use of talc-dusted diaphragms have been reported in some epidemiologic studies, this literature is sparsely referred to, and no formal evaluation of the results has been conducted. The reasons for this probably reflect simply that perineal talc-dusting is more common than the use of talc-dusted diaphragms and can be examined with greater statistical precision in epidemiologic studies. If the use of talc-dusted diaphragms occurred more commonly than perineal dusting, it is likely that the epidemiologic literature would have focused primarily on talc-dusted diaphragms as the relevant risk factor. In any case, it can be concluded that use of talc-dusted diaphragm results in exposure of the reproductive tract to talc. Intuitively, the association with talc-dusted diaphragms appears to provide a better test of the talc hypothesis.

Consequently, a meta-analysis of the association between talc-dusted diaphragm and ovarian cancer risk was conducted (Greenland 1987). The results are shown in Tables 1 and 2 of this document. Crude odds ratios and 95% confidence intervals were calculated based on the

exposure rates in cases and controls. In some studies, the OR was calculated but was based on an inappropriate control group; e.g. subjects who reported no exposure to any talc. For these studies, the OR was recalculated based on women who never used talc-dusted diaphragm as the reference group. The summary crude odds ratio associated with use of talc-dusted diaphragm was 0.79 (95% CI 0.46, 1.38).

Using this exposure measure, it can be concluded that there is no relationship between ovarian cancer and talc exposure. Limitations in this conclusion include lack of any data with cumulative exposure and possible confounding. In a meta-analysis of ten studies that examined talc dusting and ovarian cancer risk, Gross and Berg (1995) found little difference between the summary crude odds ratio and the summary adjusted odds ratio. Similarly, it appears unlikely that confounding could have obscured a positive association with talc-dusted diaphragms. There are some uncertainties in the interpretation of these findings. Diaphragms are used with contraceptive jelly, which could affect migration of talc particles. However, the jelly is normally applied to the rim and the surface of the diaphragm that faces the vaginal entrance. Jelly is not normally applied to the surface of the diaphragm that faces the cervix. Cramer et al. (1982) point out that talc-dusted diaphragms might be washed prior to their use. If this does occur, the extent to which this impacts on the study findings depend on how these questions were asked, or whether women volunteered to provide this information if not asked.

2. Recall bias

A pattern of small but elevated odds ratios for a particular exposure is not uncommon in epidemiologic research. Although it might seem intuitive that consistent findings across case-control studies could not be attributed to bias, some examples will serve to refute this notion. For example, in a meta-analysis of 12 case-control studies of dietary fat and breast cancer, the OR for the upper quartile of fat consumption relative to the lowest quartile was 1.46 ($p<0.0001$) (Howe et al. 90). In contrast, a pooled analysis of 7 cohort studies of dietary fat intake and breast cancer found an OR of 1.05 (95% CI 0.94-1.16) in the highest quintile relative to the lowest quintile of fat consumption (Hunter et al. 1996). One possible explanation for the discrepancy in the findings is recall bias in the case-control studies. Alternatively, changes in the levels of the exposure (e.g., dietary fat consumption) might have obscured an association in the cohort studies. In order to determine which explanation is more likely where case-control and cohort data conflict, it is useful to find examples in which there would not be any changes in the exposure categorization in the cohort studies. Examples would include studies in which the exposure is classified as ever versus never in middle age or older age adults.

One example is based on the literature of environmental tobacco smoke exposure and the risk of breast cancer. In a meta-analysis of 5 case-control studies of ever versus no exposure, the OR was 1.8 (95% CI 1.4-2.5). However in a pooled analysis of raw data from three cohort studies, the relative risk was 1.1 (95% CI 0.9-1.4). (Wartenberg et al. 2000). Similarly, while the majority of case-control studies that examined the association between having ever dusted the perineum with talc and ovarian cancer found increased risks, no association was observed in the Nurses Health Study (Gertig et al. 2000). The lack of an association in the Nurses Health Study was not due to exposure misclassification since middle aged nurses in their 40's or 50's who reported never using dusting powder would start using talc powder after the baseline period. By

definition, in a never versus ever classification, exposure misclassification could not occur in cohort members who reported ever using perineal talc. An increased risk was observed in one histologic subgroup of ovarian cases in the Nurses Health Study, a finding that requires further study.

The above examples of conflicting findings in case-control versus cohort studies were selected from the breast cancer literature since the methodologic issues in female breast cancer studies might be expected to be similar to that for ovarian cancer. However, the above examples do illustrate that recall or other forms of bias might result in spurious findings in case-control studies even when an increased risk has been observed in multiple case-control studies.

Recall bias in case-control studies can either be nondifferential or differential. Nondifferential recall bias occurs when the extent of bias is equal between cases and controls and results in an attenuated association. Nondifferential bias results in an underestimation of risk, whereas differential bias can result in a spurious association. It seems plausible that differential recall bias occurred in the case-control studies of talc perineal dusting and ovarian cancer. It is possible to speculate that for a cancer with few known causes and a poor prognosis, there might be a greater interest or a greater attention paid to questions on talc powder use in cases than controls. In these studies, the primary focus was on hormonal and reproductive risk factors. A question or series of questions on genital talc dusting might be considered quite unusual and striking to case subjects who are most likely unfamiliar with the hypothesis of talc dusting. Healthy controls would not have any incentive to ponder this question. Although it is possible to speculate on these scenarios, it is necessary to examine the epidemiologic literature to determine the presence and extent of recall bias. For example, differential recall bias could be inferred by comparing studies that used women with gynecologic or obstetric conditions with studies that used population controls or hospital controls with non-gynecologic conditions. An ideal control group would be women with ovarian cysts who might also be expected to respond to questions on talc dusting in a similar way as cases with ovarian cancer. If studies using women with ovarian cysts as controls also showed an increased risk for ovarian cancer associated with talc dusting, recall bias could be eliminated as a possible explanation. However, none of the case-control studies of ovarian cancer used a design based on controls with reproductive diseases.

In most of the case-control studies on talc and ovarian cancer, cases were interviewed after discharge from the hospital or were identified from cancer registries. Post-operative treatment of ovarian cancer includes radiation to the pelvic area, which might result in skin rashes or irritation in the perineum/genital area. It is feasible that this leads to increased talc use following diagnosis. The increased risks observed in these studies might reflect post-diagnosis use of talc. This is supported by the fact that in the four studies in which cases were interviewed directly in the hospital at diagnosis, there was little or no association with overall talc use or perineal talc use (Hartge 1983, Harlow and Weiss 1989, Rosenblatt et al. 1992, Whittemore et al. 1988). The summary of individual studies attached to this response provides more information on case-control methodologies. In some studies that did find positive associations, subjects were asked about exposure prior to diagnosis. However, it is uncertain whether case subjects with a high degree of morbidity that used talc after diagnosis would make the distinction.

In summary, there have been no attempts to scientifically determine whether alternative explanations such as recall bias resulted in spurious increased risks in case-control studies of talc and ovarian cancer. Since recall bias is a plausible alternative explanation, as has been shown in other examples of case-control studies, the associations with talc dusting cannot be considered a causal relationship. The burden of proof in carcinogen identification depends on the demonstration of scientific fact and ruling out alternate explanations. There have been no attempts to identify and rule out possible alternate explanations despite research efforts that have been ongoing for almost two decades.

3. Validity and reliability

In many fields of epidemiologic research, especially studies with exposures that give rise to low odds ratios, extensive validation work is done prior to using the exposure measure in case-control or cohort studies. For example, in nutritional epidemiology, the validity of a food frequency questionnaire (FFQ) is usually determined by comparison of responses to food diaries. FFQs with low-validity scores are often considered too imprecise for its proper use in epidemiology and few studies are published without demonstrating at least moderate single-order or partial correlations with food diary data. Studies of environmental tobacco smoke and cancer or coronary disease outcomes have been supported by studies demonstrating a high degree of concordance between reported smoke exposure and biological markers of tobacco smoke metabolites. Occupational studies of cancer and environmental air pollutants, water pollutants and electromagnetic fields also commonly employ complex exposure measures using environmental hygiene measurements.

The validity of reported exposures to genital talc powder has never been determined. There have been no attempts to compare reported talc usage and frequency with a log maintained over a defined amount of time (e.g., two weeks, month). It is possible to conduct validity tests on current but not past talc usage patterns.

In the absence of validity testing, it is common to infer the validity of the exposure measurement by repeated questionnaires. If the same response is provided on two or more separate occasions, it suggests that the exposure measure is possibly valid, although inaccurate information can be obtained from both readings. There have been no attempts to determine the reliability of self-reported history of perineal talc powder use. In one study that evaluated the reliability of questions on menstrual history, percent agreement for age at menopause and at menarche and other variables ranged from 70-90% (Bean et al. 1980). Recollection of menstrual cycle length was considered unreliable, and it might be suspected that questions on perineal talc dusting are to some extent unreliable also. Indeed lifelong usage patterns are unknown, but respondents in epidemiologic studies are forced to provide information in terms of a regular lifetime pattern. Talc powder use might be sporadic, seasonal or change with circumstances (e.g., sexual activity, parity).

Concerns over measurement error are done prior to the design, collection and analysis of epidemiologic studies. Epidemiologic studies of perineal talc dusting and ovarian cancer have been conducted over the past 18 years, yet no attempts have been made to determine the extent to

which talc usage questions measure what they purport to measure. Measurement error can have a major impact on the results of epidemiologic studies (Kelsey et al. 1986).

4. Dose-response relationship

The fundamental aspect of carcinogenic mechanisms is that the likelihood of tumor initiation and promotion is directly related to cumulative genetic damage and cellular insults (Wynder et al. 1992). In the studies that provide information on dose response, the ORs associated with the intensity, duration or cumulative exposure do not generally show an increasing trend, as would be expected in a causal association. In fact, by and large the smallest associations are seen with the lowest levels of exposure. These findings are inconsistent with known mechanisms of carcinogenic action. The NTP document notes that the "evidence for causality is weakened by the absence of the exposure response trends in most studies." A dose-response trend is a necessary condition for inferring causality and the absence, much less an inverse trend, is incompatible with accepted paradigms of causality. The NTP notes that this lack of a trend might be due to "the difficulty of measuring exposures by retrospective recall." If this is the case, then the arguments presented in this response regarding adequate exposure definition, validity of measurement, and recall bias would argue against a causal interpretation.

5. Tubal ligation

The NTP document noted that the association with talc dusting was apparent only in women who never underwent tubal ligation in the study by Cramer et al. (1999). Similar findings were noted in the Harlow et al. (1992) and Whittemore et al. (1988) studies, but not in the Gertig et al. (2000) study.

The NTP document did not make mention that the association between after bath talc use and ovarian cancer in the Chang and Risch (1997) study was the same for women who underwent tubal ligation and women who did not have this procedure. In the study by Wong et al. (1999), there was no difference in the risk of ovarian cancer associated with talc use between women who had tubal ligation or hysterectomy (OR=0.9, 95% CI 0.4, 2.2) and those with no history (OR=1.2, 95% CI 0.8, 1.6). Although not stated in the NTP document, in the Whittemore et al. (1988) study women who were exposed to talc prior to tubal ligation or hysterectomy had an increased risk only for 1-9 years of exposure. The OR for 10+ years was 1.11 (0.74-1.65). In addition, inverse trends with duration of talc exposure were found after adjustment for tubal ligation (Ness et al. 2000).

6. Publication bias

Publication bias is the failure to report numerical negative findings. The investigation by Chang and Risch (1997) was specifically designed to assess talc exposure and ovarian cancer risk. However, the authors omitted from the publication findings on the association with talc-dusted diaphragm. Similarly, Cramer et al. (1999) readily acknowledge "we did not assess potential talc exposure from diaphragms or condoms, exposures not found to be associated with ovarian cancer in our previous studies." Thus, the authors specifically omitted data that would potentially refute their hypothesis. Chen et al. (1992) found a positive association with talc dusting. Other sources

of talc exposure were collected but odds ratios were not calculated because it was stated that there was no association.

Summary

A number of epidemiologic studies have found small but consistent associations with ever having dusted the perineum with talc powder or use of a talc-dusted sanitary napkin and the risk of ovarian cancer. Some studies show a greater risk in women with tubal ligation, whereas an equal number or more find no difference. These findings raise the hypothesis that talc use is associated with ovarian cancer, but do not test the hypothesis. Proper epidemiologic methods require determining the extent to which alternative explanations account for study findings. These include bias, confounding and the accuracy of exposure measurement—methodologic issues inherent in all epidemiologic designs. The one cohort study of talc dusting and ovarian cancer risk found no overall association, raising the possibility of recall bias as an alternative explanation in case-control studies. Misclassification in the cohort study is unlikely using an "ever versus never" exposure measure. An increased risk in one histologic subgroup in the cohort study adds to the uncertainty in these data. Many studies had no information on dose of exposure, and the lack of an overall dose-response relationship in those studies with this information argues against a causal interpretation. Indeed the inverse trend found in several studies is incompatible with the known mechanisms of carcinogenesis.

There was no summary association between ovarian cancer and using a talc-dusted diaphragm, an exposure measurement that perhaps has greater validity than perineal talc dusting in reflecting ovarian exposure to talc. The overall risk fell below unity although the association was not statistically significant. The precisions of these results were affected by publication bias in which some studies that were designed specifically to test the talc hypothesis failed to report or test the association with talc-dusted diaphragms.

The epidemiologic data on talc exposure are conflicting, but do not support the hypothesis that cosmetic talc is "reasonably anticipated" to be a human carcinogen for the ovary.

References

1. Acheson ED, Gardner MJ, Pippard H et al. Mortality of two groups of women who manufactured gas masks from chrysotile and crocidolite asbestos: a 40-year follow-up. Brit J Med 1982;39:344-348.

2. Bean JA, Leeper JD, Wallace RB et al. Variation in the reporting of menstrual histories. . Am J Epidemiol 1979;109:181-185.

3. Berry G, Newhouse ML, Wagner JC. Mortality from cancers of asbestos factory workers in east London 1933-80. Occup Environ Med 2000;57:782-785.

4. Booth M, Beral V, Smith P: Risk factors for ovarian cancer: A case-control study. Br J Cancer 1989;60:592-598.

5. Chang S, Risch HA: Perineal talc exposure and risk of ovarian carcinoma. Cancer 1997;79:2396-2401.

6. Chen Y, Wu PC, Lang JH, Ge WJ, Hartge P, Brinton LA: Risk factors for epithelial ovarian cancer in Beijing, China. Int J Epidemiol 1992;21:23-29.

7. Clement PB, Young RH, Scully RE: Malignant mesotheliomas presenting as ovarian masses. A report of nine cases, including two primary ovarian mesotheliomas. Am J Surg Pathol 1996;20:1067-1080.

8. Cook LS, Kamb ML, Weiss NS: Perineal powder exposure and the risk of ovarian cancer. Am J Epidemiol 1997;145:459-465.

9. Cramer DW, Liberman RF, Titus-Ernstoff L, Welch WR, Greenberg ER, Baron JA, Harlow BL: Genital talc exposure and risk of ovarian cancer. Int J Cancer 1999; 81:351-356.

10. Cramer DW, Xu H: Epidemiologic evidence for uterine growth factors in the pathogenesis of ovarian cancer. Ann Epidemiol 1995;5:310-314.

11. Cramer DW, Welch WR, Scully RE, Wojciechowski CA: Ovarian cancer and talc. Cancer 1982;50:372-376.

12. Eltabbakh GH, Piver MS, Natarajan N, Mettlin CJ: Epidemiologic differences between women with extraovarian primary peritoneal carcinoma and women with epithelial ovarian cancer. Obstet Gynecol 1998;91:254-259.

13. Gertig DM, Hunter DJ, Cramer DW, Colditz GA, Speizer FE, Willett WC, Hankinson SE: Prospective study of talc use and ovarian cancer. J Natl Cancer Inst 2000;92:249-252.

14. Godard B, Foulkes WD, Provencher D, Brunet JS, Tonin PN, Mes-Masson AM, Narod SA, Ghadirian P: Risk factors for familial and sporadic ovarian cancer among French Canadians: A case-control study. Am J Obstet Gynecol 1998;179:403-410.

15. Green A, Purdie D, Bain C, Siskind V, Russell P, Quinn M, Ward B, and the Survey of Women's Health Study Group: Tubal sterilisation, hysterectomy and decreased risk of ovarian cancer. Int J Cancer 1997;71:948-951.

16. Greenland S. Quantitative methods in the review of epidemiologic literature. Epidem Rev 1987;9:1-30.

17. Gross AJ, Berg PH: A meta-analytical approach examining the potential relationship between talc exposure and ovarian cancer. J Exp Anal Environ Epidemiol 1995;5:181-195.

18. Hankinson SE, Hunter DJ, Colditz GA, Willett WC, Stampfer MJ, Rosner B, Hennekens CH, Speizer FE: Tubal ligation, hysterectomy, and risk of ovarian cancer. A prospective study. JAMA 1993; 270:2813-2818.

19. Harlow BL, Cramer DW, Bell DA, Welch WR: Perineal exposure to talc and ovarian cancer risk. Obstet Gynecol 1992; 80:19-26.

20. Harlow BL, Weiss NS: A case-control study of borderline ovarian tumors: The influence of perineal exposure to talc. Am J Epidemiol 1989;130:390-394.

21. Hartge P, Hoover R, Lesher LP, McGowan L: Talc and ovarian cancer [letter]. JAMA 1983;250:1844.

22. Hartge P, Stewart P: Occupation and ovarian cancer: A case-control study in the Washington, DC, metropolitan area, 1978-1981. J Occup Med 1994;36:924-927.

23. Heller DS, Westhoff C, Gordon RE, Katz N: The relationship between perineal talc usage and ovarian talc particle burden. Am J Obstet Gynecol 1996; 174:1507-1510.

24. Heller DS, Gordon RE, Katz N: Correlation of asbestos fiber burden in fallopian tubes and ovarian tissue. Am J Obstet Gynecol 1999; 181:346-347.

25. Henderson WJ, Joslin CAF, Turnbull AC et al. Talc and carcinoma of the ovary and cervix. J Obstet Gynecol Br Commonw 1971;78:266-272.

26. Howe GR, Hirohata T, Hislop TG et al. Dietary factors and risk of breast cancer: combined analysis of 12 case-control studies. J Natl Cancer Inst 1990;82:561-569.

27. Hunter DJ, Spiegelman D, Adami HO et al. Cohort studies of fat intake and the risk of breast cancer-a pooled analysis. N Engl J Med 1996;334:356-361.

28. Kelsey J, Thompson WD, Evans AS: Methods in Observational Epidemiology. Oxford University Press, New York, 1986, pg 306.

29. Muscat JE, Wynder EL. Cigarette smoking, asbestos exposure, and malignant mesothelioma. Cancer Research 1991;51:2263-2267.

30. Muscat JE, Barish M: Epidemiology of talc exposure and ovarian cancer. A critical appraisal. Comments on Toxicology 1998;6:327-335.

31. Ness RB, Grisso JA, Cottreau C, Klapper J, Vergona R, Wheeler JE, Morgan M., Schlesselman JJ: Factors related to inflammation of the ovarian epithelium and risk of ovarian cancer. Epidemiol 2000;11:111-117.

32. Price B. Analysis of current trends in United States mesothelioma incidence. Am J Epidemiol 1997;145:211-218.

33. Purdie D, Green A, Bain C, Siskind V, Ward B, Hacker N, Quinn M, Wright G, Russell P, Susil B: Reproductive and other factors and risk of epithelial ovarian cancer: An Australian case-control study. Survey of Women's Health Study Group. Int J Cancer 1995;62:678-684.

34. Rosenblatt KA, Szklo M, Rosenshein NB: Mineral fiber exposure and the development of ovarian cancer. Gynecol Oncol 1992;45:20-25.

35. Rosenblatt KA, Mathews WA, Daling JR, Voigt LF, Malone K. Characteristics of women who use perineal powder. Obstet & Gynecol 1998;92:753-756.

36. Sato S, Ito K, Konno R, Yajima A: Differential diagnosis of mesothelial and ovarian cancer cells in ascites by immunohistochemistry using Ber-EP4 and calretinin. Acta Cytol 2000;44:85-88.

37. Shushan A, Paltiel O, Iscovich J, Elchalal U, Peretz T, Schenker J: Human menopausal gonadotropin and the risk of epithelial ovarian cancer. Fertil Steril 1996;65:13-18.

38. Tzonou A, Polychronopoulou A., Hsieh CC, Rebelakos A, Karakatsani A, Trichopoulos D: Hair dyes, analgesics, tranquilizers and perineal talc application as risk factors for ovarian cancer. Int J Cancer 1993;55:408-410.

39. Wartenburg D, Calle E, Thun M et al. Passive smoking exposure and female breast cancer mortality. J Natl Cancer Inst 2000;92:1666-1673.

40. Wehner A: Is cosmetic talc safe? Comments in Toxicology 1998;6:337-366.

41. Whittemore AS, Wu ML, Paffenbarger RS, Sarles DL, Kampert JB, Grosser S, Jung DL, Ballon S, Hendrickson M: Personal and environmental characteristics related to epithelial ovarian cancer. II. Exposures to talcum powder, tobacco, alcohol, and coffee. Am J Epidemiol 1988;128:1228-1240.

42. Wong C, Hempling RE, Piver MS, Natarajan N, Mettlin CJ: Perineal talc exposure and subsequent epithelial ovarian cancer: A case-control study. Obstet Gynecol 1999;93:372-376.

43. Wynder EL, Williams GM. Metabolic overload and carcinogeneisis from the viewpoint of epidemiology. In: Chemical Carcinogenesis. (A. Somogyi et al., eds.). Munich 1992, 17-22.

Table 1. Crude odds ratios for talc-dusted diaphragm and ovarian cancer.

Study, year	Exposure (y/n)	Exposure note	Cases	Controls	Crude odds ratio	Cornfield 95% CI
1. Booth,1989	Ever		30	74		
	Never		178	329	0.75	0.33, 2.02
2. Cook, 1997	Ever		154	256	0.96	0.52, 1.76
	Never		22	35		
3. Cramer, 1982	Ever		22	19	1.18	0.59, 2.35
	Never		193	196		
4. Harlow, 1992	Ever	Diaphragm or partner	20	21	0.97	0.49, 1.92
	Never		215	218		
5. Harlow, 1989	Ever		11	27	0.51	0.22, 1.13
	Never		105	131		
6. Hartge, 1983	Ever		25	41	0.72	0.40, 1.3
	Never		110	130		
7. Ness, 2000	Ever		10	33	0.53	0.25, 1.13
	Never		757	1334		
8. Rosenblatt, 1992	Ever	Education-adjusted OR=3.0	14	5	1.82	0.55, 6.34
	Never		60	39		
9. Whittemore, 1988	Ever		9	19	1.38	0.57, 3.28
	Never		179	520		

Table 2. Meta-analysis of studies of talc-dusted diaphragm and ovarian cancer

Study, year	Crude odds ratio	95% CI	Variance	Weight (W)
1. Booth, 1989	0.75	0.85, 2.02	0.1754	5.70
2. Cook, 1997	0.96	0.52, 1.76	0.0978	10.22
3. Cramer, 1982	1.18	0.59, 2.35	0.1251	7.99
4. Harlow, 1992	0.97	0.49, 1.92	0.1214	8.24
5. Harlow, 1989	0.51	0.22, 1.13	0.1840	5.43
6. Hartge, 1983	0.72	0.40, 1.3	0.0899	11.12
7. Ness, 2000	0.53	0.25, 1.13	0.1470	6.80
8. Rosenblatt, 1992	1.82	0.55, 6.34	0.3728	2.68
9. Whittemore, 1988	1.38	0.57, 3.28	0.2035	4.91
Σ				63.09
OR	0.86	0.59, 1.40		

APPENDIX: ANALYTIC EPIDEMIOLOGY STUDIES
USE OF COSMETIC TALC AND RISK OF OVARIAN CANCER

CITATION, LOCATION	STUDY TYPE, POPULATION	EXPOSURE RATE OF GENITAL TALC IN CONTROLS AND DEFINITION	RISK ESTIMATES	COMMENTS
Booth et al. 1989 England	Hospital-based case-control study 235 cases with histologically confirmed epithelial ovarian cancer and 451 age-matched hospitalized controls.	59% The exposure is simply called 'talc use in the genital area.'	**Reported Frequency of Talc Use** / **Relative Risk (95% CI)** Never 1.0 (reference) Rarely 0.9 (0.3-2.4) Monthly 0.7 (0.3-1.8) Weekly 2.0 (1.3-3.4)* Daily 1.3 (0.8-1.9)	Cases do not appear to be newly diagnosed and study might contain prevalent cases. Case response rate was 84 percent. An evaluation of response bias not done. No information on duration of exposure. No association found with talc-dusted diaphragm.

* Denotes statistically significant *increase* in risk.
▼ Denotes statistically significant *decrease* in risk.

APPENDIX: ANALYTIC EPIDEMIOLOGY STUDIES
COSMETIC TALC AND RISK OF OVARIAN CANCER (Continued)

CITATION, LOCATION	STUDY TYPE, POPULATION	EXPOSURE RATE OF GENITAL TALC IN CONTROLS AND DEFINITION	RISK ESTIMATES		COMMENTS
Chang and Risch 1997 Canada	Case-control (population-based) 450 cases with borderline and invasive ovarian carcinoma. 564 community-based age-matched controls.	After bathing: 32%	**Talc Use** None Any **After Bath Talc Use** <10 10-25 >25 **Years of After-Bath Talc Use** <30 30-40 >40 **Histologic Type** Invasive Borderline Serous Mucinous Endometroid	**Odds Ratio (95% CI)** 1.000 (reference) 1.420 (1.08-1.86)* 1.836 (1.24-2.73)* 1.128 (0.74-1.72) 0.951 (0.61-1.49) 1.697 (1.09-2.64)* 1.435 (0.96-2.15) 0.865 (0.54-1.38) 1.513 (1.13-2.02)* 1.237 (0.76-2.02) 1.336 (0.96-1.85) 1.585 (0.97-2.58) 1.671 (1.00-2.79)	Interviews in cases occurred 3-4 months after diagnosis. Information pertaining to one year prior to diagnosis was excluded. The authors state in the abstract that "a borderline statistical association was detected between duration of talc exposure and risk (OR 1.09, 95% CI 0.98-1.21, per 10 years of exposure." This conclusion is a misrepresentation of the data. It is unstated but the 'statistical association' that was found appeared to be based on a linear model. (The model parameters were not stated.) The data clearly do not show a linear increase with duration after 30 years of use. The exposure-response relationship is U-shaped, with a decreased risk associated with >40 years of use. There was no difference in risk between subjects ever having had sterilization and other subjects. The authors conclusion that 'this investigation supports previous contentions that exposure to talc may increase risk of ovarian cancer' is unsubstantiated by the data on duration of use, frequency of use and risk estimates stratified by history of sterilization. Risk estimates were adjusted for several possible confounders.

* Denotes statistically significant *increase* in risk.
▼ Denotes statistically significant *decrease* in risk.

APPENDIX: ANALYTIC EPIDEMIOLOGY STUDIES
COSMETIC TALC AND RISK OF OVARIAN CANCER (Continued)

CITATION, LOCATION	STUDY TYPE, POPULATION	EXPOSURE RATE OF GENITAL TALC IN CONTROLS AND DEFINITION	RISK ESTIMATES		COMMENTS
Chen et al. 1992 Bejing, China	Case-control (population-based) Cases were 112 women with newly diagnosed epithelial ovarian cancer from 1984-1986. Controls were 224 women matched to cases by age.	2% Subjects were asked if they used dusting powder on their abdomen or perineal region for 3 or more months. All exposure information sought was with reference to events occurring 3 or more years before the date of diagnosis (equivalent date in controls).	**Dusting Powder ≥3 Months** No Yes	**Relative Risk (95% CI)** 1.0 (reference) 3.9 (0.9-10.6)	Despite nonsignificant finding with very wide confidence interval, the authors report this finding in the abstract. No details are provided on what types of powder are used in China. Prevalent cases were interviewed after discharge. Response rate in cases was 51% although findings for hormone/reproductive factors similar to other studies, the same conclusion cannot be reached with regard to talc exposure. Information obtained for 3 or more years prior to diagnosis. The authors reported investigating several sources of talc exposure. They state that the only exposure associated with an increased risk was dusting powder. The authors provided no data and failed to present odds ratios associated with talc exposures that were unrelated to ovarian cancer risk. No information on risk by frequency and duration of exposure. Risk estimate was adjusted for education and parity.

* Denotes statistically significant *increase* in risk.
▼ Denotes statistically significant *decrease* in risk.

APPENDIX: ANALYTIC EPIDEMIOLOGY STUDIES
COSMETIC TALC AND RISK OF OVARIAN CANCER (Continued)

CITATION, LOCATION	STUDY TYPE, POPULATION	EXPOSURE RATE OF GENITAL TALC IN CONTROLS AND DEFINITION	RISK ESTIMATES		COMMENTS
Cook et al. 1997 Washington	Case-control (population-based) Cases were 313 white women aged 20-79 years in three counties of western Washington who were diagnosed with borderline (n=79) or invasive (n=234) ovarian cancer from 1986-1988. Controls were 422 white women living in the same area identified via random digit dialing. All were matched to cases by age.	39.3% Users were asked about the circumstances in which they used genital powder, as well as the duration, frequency and types of powder used. Cases were asked to refer only to the period before their diagnoses; controls were asked to consider a comparable period.	**Lifetime Genital Powder Application** None Any **Types of Exclusive Users** Never users Perineal dusting after bathing only Diaphragm storage only Powder on sanitary napkins only Genital deodorant spray only **Histologic type** Controls Serous tumors Mucinous tumors Endometroid tumors Other tumors	**Relative Risk (95% CI)** 1.0 (reference) 1.5 (1.1-2.0)* 1.0 (reference) 1.8 (1.2-2.9)* 0.8 (0.4-1.4) 1.5 (0.6-3.6) 1.5 (0.8-3.0) 1.0 (reference) 1.7 (1.1-2.5)* 0.7 (0.4-1.4) 1.2 (0.6-2.3) 1.8 (1.1-2.8)*	Prevalent cases interviewed after discharge. Cases were identified by a registry and interviewed months after diagnosis, although the exact number is unstated. Risk factor information was recorded only for exposures that occurred before diagnosis. The authors concluded that a history of perineal dusting or use of genital deodorant sprays had a modest influence on the development of epithelial ovarian tumors, whereas storing a diaphragm in powder or powdering sanitary napkins had no effect. Selection bias may have affected the results, especially since prevalent cases are used. Interviews were obtained for 64.3 percent of eligible cases, and 72.3 percent of eligible controls. Furthermore, the authors noted that the completeness of reporting may have differed between cases and controls. The Odds ratio in the highest exposure category (>10,000 applications) was the same (1.8) as in the lowest exposure category <<2,000) applications. This is a five-fold difference in reported exposure, a difference unlikely to be due poor recall. Risk estimates were adjusted for age. There were no other statistical confounders. When specific histologic categories of ovarian tumor were examined, any genital powder use was associated with an elevated risk for serous tumors and the nonspecific category of "other tumors."

* Denotes statistically significant *increase* in risk.
▼ Denotes statistically significant *decrease* in risk.

APPENDIX: ANALYTIC EPIDEMIOLOGY STUDIES
COSMETIC TALC AND RISK OF OVARIAN CANCER (Continued)

CITATION, LOCATION	STUDY TYPE, POPULATION	EXPOSURE RATE OF GENITAL TALC IN CONTROLS AND DEFINITION	RISK ESTIMATES		COMMENTS
Cramer et al. 1999 Massachusetts and New Hampshire	Case-control (population-based) Cases were 563 women diagnosed with invasive or borderline epithelial ovarian cancer from 5/1992-3/1997. Controls were 523 women from the general population matched by age and residence.	Any perineal: 18.2%. Perineal dusting:15.9% Subjects were asked about their exposure to talc one year before diagnosis for cases and one year before the interview for controls. Women were asked whether they had "regularly used talc, baby, or deodorizing powders dusted or sprayed. Data collected included method and site of application, husband's use, age at first use, types used, applications per month, and total years of genital use.	**Genital Use** No genital exposure Any genital exposure **Type of Personal Use** No personal use Dusting perineum Dusting sanitary napkin Dusting underwear Multiple uses genital area **Frequency of Use/Month** <30 30-39 40+ **Age at First Use** <20 20-25 >25 **Years of Use** <20 20-30 >30 **Total Applications** <3,000 3,000-10,000 >10,000 **Applications Censored** <3,000 3,000-10,000 >10,000	**Odds Ratio (95% CI)** 1.00 (reference) 1.60 (1.18-2.15)* 1.00 (reference) 1.45 (0.97-2.18) 1.45 (0.68-3.09) 1.21 (0.40-3.63) 2.15 (1.30-3.57)* 2.21 (1.37-3.56)* 1.17 (0.78-1.76) 1.57 (0.80-3.10) 1.46 (1.03-2.07)* 1.87 (1.03-3.39)* 1.54 (0.64-3.72) 1.86 (1.16-3.00)* 1.33 (0.76-2.30) 1.44 (0.91-2.26) 1.84 (1.12-3.03)* 1.43 (0.84-2.41) 1.43 (0.92-2.22) 1.54 (1.01-2.35)* 1.72 (1.08-2.76)* 1.80 (1.02-3.18)*	Prevalent cases were identified from registries and interviews occurred months after diagnosis. Information was collected on exposures that occurred at least one year prior to diagnosis. The authors concluded that the data demonstrate a significant association between the use of talc in genital hygiene and risk for ovarian cancer. Only 52 percent of identified cases were included. The authors state that recall bias is not likely to be a factor in this study since the exposure occurred over many years. Consistent dose-response trends were not observed across age at first use, years of use or total applications. While risk increased with increasing censored application (*i.e.*, when uses following hysterectomy or tubal ligation and uses during pregnancy or OC use were excluded), this model included women who used talc in non-genital areas. Risk estimates were adjusted for age, study center, tubal ligation, BMI, parity, OC use, primary relative with breast or ovarian cancer, and other categories of genital talc use (except where noted). Data not presented in "Risk Estimate" column: • A significant risk was observed among ever users who were parous before their first live birth. • A significant risk was observed only among women with the serous invasive histologic sub-type of ovarian cancer. • An elevated risk was not observed among married women with husbands who used talc.

* Denotes statistically significant *increase* in risk.
▼ Denotes statistically significant *decrease* in risk.

APPENDIX: ANALYTIC EPIDEMIOLOGY STUDIES
COSMETIC TALC AND RISK OF OVARIAN CANCER (Continued)

CITATION, LOCATION	STUDY TYPE, POPULATION	EXPOSURE RATE OF GENITAL TALC IN CONTROLS AND DEFINITION	RISK ESTIMATES		COMMENTS
Cramer et al. 1982 Boston, Massachusetts	Case-control (population-based) Cases were 215 white females with epithelial and borderline ovarian malignancies. Controls were 215 women from the general population matched by age, race and residence.	Any perineal: 28.4% Perineal dusting: 30.5%.	**Talc Use on Perineum** None Any **Type of Exposure** None Dusting powder *or* on napkins Dusting powder *and* on napkins	**Odds Ratio (95% CI)** 1.0 (reference) 1.92 (1.27-2.89)* 1.0 (reference) 1.55 (0.98-2.47) 3.28(1.68-6.42)*	Prevalent cases were interviewed at least several months after diagnosis. No mention is made if risk factor information pertained to prior to the diagnosis date. Only 45 percent of eligible controls and 72 percent of eligible cases participated in the study. There is an increased risk with perineal dusting, but no dose-response information is presented. No association was found with talc-dusted diaphragms.

* Denotes statistically significant *increase* in risk.
▼ Denotes statistically significant *decrease* in risk.

APPENDIX: ANALYTIC EPIDEMIOLOGY STUDIES
COSMETIC TALC AND RISK OF OVARIAN CANCER (Continued)

CITATION, LOCATION	STUDY TYPE, POPULATION	EXPOSURE RATE OF GENITAL TALC IN CONTROLS AND DEFINITION	RISK ESTIMATES		COMMENTS
Gertig et al. 2000 USA	**Prospective cohort** (population-based) Subjects were participants of the Nurses' Health Study, a prospective cohort of 121,700 registered nurses living in 11 of the larger states in the United States. All subjects were married, female nurses aged 30-35 years. After exclusions, 78,630 women formed the cohort for analysis (984,212 person-years). Ovarian cancers were diagnosed in 307 women, 121 of whom were talc users.	Subjects were asked about the frequency with which they applied "talcum, baby powder, or deodorizing powder" to their "perineal (private) areas." (Response categories included: no, < 1x/week, 1-6x/week and daily.) Subjects were also asked if they applied any of these agents on their sanitary napkins (yes, no). "Ever talc use" was classified as ever talc use on either the perineal area or sanitary napkins.	**Ever Perineal Talc Use** No Yes **Talc Use on Perineum** Never <1x/week 1-6x/week Daily **Serous Invasive Cancers, Ever Perineal Use** No Yes <1x/week 1-6x/week Daily	**Relative Risk (95% CI)** 1.00 (reference) 1.09 (0.84-1.37) 1.00 (reference) 1.14 (0.81-1.59) 0.99 (0.67-1.46) 1.12 (0.82-1.55) 1.00 (reference) 1.40 (1.02-1.91)* 1.29 (0.81-2.04) 1.49 (0.77-2.11) 1.49 (0.98-2.26)	The authors concluded that the results provide little support for any substantial association between perineal talc use and ovarian cancer risk overall; however, perineal talc use may modestly increase the risk of invasive serous ovarian cancer. As this study used a prospective cohort design, recall bias was avoided and selection bias was reduced. A dose-response trend was not observed with increasing frequency of use. Risk estimates were adjusted for age, parity, duration of oral contraceptive use, body mass index, tubal ligation history, smoking status, and postmenopausal hormone use. While the association between ever perineal use and invasive serous cancers was statistically significant, women over age 45 seemed to account for this association. These women may have been exposed to asbestiform talc. Furthermore, a stratified analysis by frequency of use (< 1x/wk, 1-6x/wk, daily) did not reveal significant associations for any sub-group, nor did it reveal a clear dose-response pattern. While the talc hypothesis depends on the ability of fibers to migrate up a patent genital tract to the ovaries, no differences in risk were observed between women who had reported tubal ligation and those who had not. NOTE: Similar to Hankinson et al. 1993, subjects were participants of the Nurses' Health Study.

* Denotes statistically significant *increase* in risk.
▼ Denotes statistically significant *decrease* in risk.

APPENDIX: ANALYTIC EPIDEMIOLOGY STUDIES
COSMETIC TALC AND RISK OF OVARIAN CANCER (Continued)

CITATION, LOCATION	STUDY TYPE, POPULATION	EXPOSURE RATE OF GENITAL TALC IN CONTROLS AND DEFINITION	RISK ESTIMATES		COMMENTS
Godard et al. 1998 Montreal, Canada	Case-control (population-based) Cases were 170 women aged 20-84 years with histologically confirmed primary ovarian carcinomas or borderline tumors diagnosed from 1995-1996. Controls were randomly selected 170 women frequency-matched to cases by age and ethnic group.	Any: 4.7%.	**Use of Talc on Perineum** Never Ever **Sporadic Cancer Patients** Never Ever **Familial Cancer Patients** Never Ever	**Relative Risk (95% CI)** 1.00 (reference) 2.49 (0.94-6.58) 1.0 (reference) 2.45 (0.85-7.07) 1.0 (reference) 3.25 (0.85-12.4)	Prevalent cases were ascertained in oncology clinics. These are not newly diagnosed cases. No mention is made as to whether risk factor information was collected prior to diagnosis. The very low exposure rate in controls raises questions on what exactly is being measured. Another Canadian study (Chang and Risch 97) found that 35.6% of controls reported talc use. Of the eligible cases and controls, the response rates were 87 percent and 89 percent, respectively. Data on frequency and duration of talc use were not collected.

* Denotes statistically significant *increase* in risk.
▼ Denotes statistically significant *decrease* in risk.

APPENDIX: ANALYTIC EPIDEMIOLOGY STUDIES
COSMETIC TALC AND RISK OF OVARIAN CANCER (Continued)

CITATION, LOCATION	STUDY TYPE, POPULATION	EXPOSURE RATE OF GENITAL TALC IN CONTROLS AND DEFINITION	RISK ESTIMATES		COMMENTS
Harlow et al. 1992 Boston, Massachusetts	Case-control (population-based) Cases were 235 white women diagnosed with histologically confirmed epithelial ovarian cancer from 1984-1987. Controls were 239 women matched on race, age, and residence.	Any genital: 39.3% Perineal dusting: 25.5% Data regarding perineal talc use included the method of application, as well as brands used, age at first use, total years of use, and frequency of use per month.	**Genital Talc Application**	**Odds Ratio (95% CI)**	Cases appear to be prevalent cases and interviewed after discharge. Hospital-based case interview, neighborhood controls used. The authors concluded that these data support the concept that a lifetime pattern of perineal talc use may increase the risk for epithelial ovarian cancer, but is unlikely to be the etiology for the majority of epithelial ovarian cancers. The authors discouraged the use of talc for daily genital hygiene. As this study used a case-control design, recall and/or selection bias may have affected the results. Only 60 percent of identified cases and 45 percent of identified controls were included in the analysis. A consistent dose-response trend was not observed across the number of talc applications per month, years of talc use, age at first talc use, or years since last talc use. A statistically significant linear trend was observed across censored applications. Risk estimates were adjusted for parity, education, marital status, religion, use of sanitary napkins, douching, age, and weight. Sub-analyses of years of talc use did not reveal any significant associations. Additional data not presented in "Risk Estimate" column: • A significantly elevated risk was not observed among any category of years of talc use, total applications, applications excluding use after sterilization, and women with mid-cycle pain. • A significant risk was observed among women with a regular period, women with no history of PID or ectopic pregnancy and women with ovarian tumors of endometrioid type or borderline grade.
			None	1.0 (reference)	
			Any	1.5 (1.0-2.1)	
			Talc Applications/Month		
			<5	1.5 (0.8-2.7)	
			5-29	1.2 (0.6-2.2)	
			≥30	1.8 (1.1-3.0)*	
			Age at First Talc Use		
			<20	1.7 (1.1-2.7)*	
			20-25	1.2 (0.6-2.2)	
			>25	1.6 (0.8-3.2)	
			Years Since Last Talc Use		
			Within last 6 mos.	2.3 (1.3-4.0)*	
			Between 6 mos.-10 yrs.	1.1 (0.7-1.9)	
			≥ 10 yrs.	1.4 (0.8-2.6)	
			Era of Use		
			Exclusive use after 1960	1.1 (0.6-2.1)	
			Any use before 1960	1.7 (1.1-2.7)*	
			Applications Excluding Use After Sterilization or During Nonovulatory Months		
			None	1.0 (reference)	
			<1000	1.5 (0.8-2.9)	
			1000-10,000	1.3 (0.8-2.0)	
			>10,000	2.8 (1.4-5.4)*	

* Denotes statistically significant *increase* in risk.
▼ Denotes statistically significant *decrease* in risk.

APPENDIX: ANALYTIC EPIDEMIOLOGY STUDIES
COSMETIC TALC AND RISK OF OVARIAN CANCER (Continued)

CITATION, LOCATION	STUDY TYPE, POPULATION	EXPOSURE RATE OF GENITAL TALC IN CONTROLS AND DEFINITION	RISK ESTIMATES		COMMENTS
Harlow and Weiss 1989 Washington State	Case-control (population-based) Cases were 116 white women aged 20-79 diagnosed with serous and mucinous borderline ovarian tumors during the years 1980-1985. Controls were 158 white women recruited via random digit dialing and matched to controls on age and county of residence.	Any perineal talc: 40.5% Perineal dusting: 23.4%. Subjects were asked about their perineal exposure to talc, as well as the types of powder used and various methods of application. Ever-users included women who reported using either one or more of three types of talc containing powders or cornstarch. Exposure information pertained to use prior to diagnosis (or a similar date for controls).	**Perineal Exposure to Talc** None Any **Type of Powder Used** None Cornstarch only Baby powder only Baby powder only or combined use Talc, unspecified Deodorizing powder only Deodorizing powder only or combined use	**Relative Risk (95% CI)** 1.0 (reference) 1.1 (0.7-2.1) 1.0 (reference) 0.8 (02-3.8) 0.8 (0.4-1.9) 0.9 (0.5-2.0) 1.0 (0.4-2.4) 3.5 (1.2-28.7)* 2.8 (1.1-11.7)*	Based on prevalent cases although only 5% were deceased. Subjects were asked about exposure before diagnosis. The authors concluded that the application of talc to diaphragms is not associated with increased risk of borderline ovarian tumors, but that there was a modest increase in risk among women who applied talc-containing powders to the perineum or sanitary napkin. Interviews were obtained for 68 percent of eligible cases and 74 percent of eligible controls. Data on frequency and duration of talc use were not collected. Risk estimates were adjusted for age, parity, and use of oral contraceptives. No association was found for talc-dusted diaphragm.

* Denotes statistically significant *increase* in risk.
▼ Denotes statistically significant *decrease* in risk.

APPENDIX: ANALYTIC EPIDEMIOLOGY STUDIES
COSMETIC TALC AND RISK OF OVARIAN CANCER (Continued)

CITATION, LOCATION	STUDY TYPE, POPULATION	EXPOSURE RATE OF GENITAL TALC IN CONTROLS AND DEFINITION	RISK ESTIMATES		COMMENTS
Hartge et al. 1983 Washington, DC	Case-control (hospital-based) Cases were 135 women with pathologically confirmed primary epithelial ovarian cancer identified from 1974-1977 in Washington, DC metropolitan hospitals. Controls were 171 women treated at the same hospitals for conditions other than gynecologic, psychiatric, or malignant diseases or pregnancy. Controls were frequency-matched to cases on age, race and hospital.	Any perineal talc: 35.7% Subjects were asked questions regarding the use of talc (no use vs. any) and the method of application.	**Talc Use**	**Relative Risk (95% CI)**	Hospital-based interviews of newly diagnosed cases and hospital controls. No overall association was found with perineal talc dusting. Data on frequency and duration of talc use were not collected; consequently, dose-response information is not available. No information is provided on response rates. Little methodologic information is supplied.
			No talc mentioned	1.0 (reference)	
			Any talc mentioned	0.7 (0.4-1.1)	
			Diaphragm-related Talc Use		
			No diaphragm used	1.0 (reference)	
			Diaphragm used, no talc	1.6 (0.7-3.7)	
			Diaphragm, with talc	0.8 (0.4-1.4)	
			Body Talc Use		
			No body talc	1.0 (reference)	
			Some body talc	0.8 (0.5-1.2)	
			"All over"	0.7 (0.4-1.2)	
			Genital	2.5 (0.7-10.0)	
			Legs only	NA	
			Not genital	0.8 (0.3-2.5)	
			Unknown where	0.3 (0.1-1.2)	

* Denotes statistically significant *increase* in risk.
▼ Denotes statistically significant *decrease* in risk.

APPENDIX: ANALYTIC EPIDEMIOLOGY STUDIES
COSMETIC TALC AND RISK OF OVARIAN CANCER (Continued)

CITATION, LOCATION	STUDY TYPE, POPULATION	EXPOSURE RATE OF GENITAL TALC IN CONTROLS AND DEFINITION	RISK ESTIMATES		COMMENTS
Ness et al. 2000 Delaware Valley	Case-control (population-based) Cases were 767 women aged 20-69 diagnosed with epithelial ovarian cancer within 6 months prior to the interview. Controls were 1367 women from the community. Ages were similarly distributed for cases and controls.	Genital/rectal: 16.0% Any perineal: not able to calculate. Women were asked if they used talc at least once per month for 6 months or more prior to 6 months before the interview. The types of use and duration of use for each type of use were recorded.	**Type of Talc Use** Never Feet, etc. Genital/rectal Sanitary napkin Underwear Diaphragm/cerv cap Male partner **Talc Use (Genital/Rectal and Feet)** Never <1 year 1-4 years 5-9 years 10+ years	**Odds Ratio (95% CI)** 1.0 (reference) 1.4 (1.1-1.6)* 1.5 (1.1-2.0)* 1.6 (1.1-2.3)* 1.7 (1.2-2.4)* 0.6 (0.3-1.2) 1.0 (0.7-1.4) 1.0 (reference) 2.0 (1.0-4.0) 1.6 (1.1-2.3)* 1.2 (0.8-1.9) 1.2 (1.0-1.5)	Cases diagnosed 6 months before the interview. Random population-based controls. Talc exposure information obtained for prior to diagnosis and for at least 6 months. The participation rate was 88% for cases, and 72% for controls. The dose-response trend with increasing duration of use was inverse. Risk estimates were adjusted for age, number of pregnancies, family history of ovarian cancer, race, oral contraceptive use, tubal ligation, hysterectomy, and breast-feeding. A nonsignificant protective effect was observed with talc-dusted diaphragm

* Denotes statistically significant *increase* in risk.
▼ Denotes statistically significant *decrease* in risk.

APPENDIX: ANALYTIC EPIDEMIOLOGY STUDIES
COSMETIC TALC AND RISK OF OVARIAN CANCER (Continued)

CITATION, LOCATION	STUDY TYPE, POPULATION	EXPOSURE RATE OF GENITAL TALC IN CONTROLS AND DEFINITION	RISK ESTIMATES		COMMENTS
Purdie et al. 1995 Australia	Case-control (population-based) Cases were 824 women aged 18-79 years with histologically confirmed cases of primary epithelial ovarian cancer identified from 1990-1993. Controls were 860 women drawn at random from electoral rolls, stratified by age and geographic location.	Subjects were administered a detailed questionnaire about reproductive and contraceptive history, as well as other factors of interest.	**Talc Use** No use On abdomen/perineum	**Relative Risk (95% CI)** 1.0 (reference) 1.27 (1.04-1.54)*	NOTE: All of the data in this study appear to have been re-analyzed subsequently in Green et al. 1997 As this study used a case-control design, recall and/or selection bias may have affected the results. There is some evidence of recall bias: interviews were conducted for 90 percent of eligible cases, but only 73 percent of eligible controls. Furthermore, while most cases were interviewed in a clinical setting, all controls (and only some cases) were interviewed in their homes. Data on frequency and duration of talc use were not collected; consequently, dose-response information is not available. The reported risk estimate was adjusted for parity.

* Denotes statistically significant *increase* in risk.
▼ Denotes statistically significant *decrease* in risk.

APPENDIX: ANALYTIC EPIDEMIOLOGY STUDIES
COSMETIC TALC AND RISK OF OVARIAN CANCER (Continued)

CITATION, LOCATION	STUDY TYPE, POPULATION	EXPOSURE RATE OF GENITAL TALC IN CONTROLS AND DEFINITION	RISK ESTIMATES		COMMENTS
Rosenblatt et al. 1992 Baltimore, Maryland	Case-control (hospital-based) Cases were 77 women diagnosed with ovarian cancer from 1981-1985. Controls were 46 women matched by age and race.	Any genital fiber: 88%. Subjects were asked about the manner and frequency of their talc use.	**Genital Talc Bath** No Yes **Use of Talc on Sanitary Napkin** No Yes	**Odds Ratio (95% CI)** 1.0 (reference) 1.7 (0.7-3.9) 1.0 (reference) 4.8 (1.3-17.8)* *Sanitary napkin or other sanitary product.	Cases newly diagnosed. Hospital-based controls. The dose-response odds ratios are calculated incorrectly. Recalculating length of fiber use relative to never users, the crude OR is: <37 years: 0.62 >37 years: 1.2 The very high exposure rate in the controls raises questions on the validity of this data. It is difficult to understand how there was a lack of available matching controls at Johns Hopkins hospital. Since control matching was partially unsuccessful, the authors performed a secondary matching. Further, questionnaire data was administered by telephone and directly. No information was given regarding how the percentages of the two methods in the control group, and this methodology was not taken into account statistically. Combined with the low response rate in cases (55%), the methodology is highly flawed. Data on frequency and duration of talc use were not collected. The risk estimate for exposure to a genital talc bath does not appear to have been adjusted for any potential confounding factors.

* Denotes statistically significant *increase* in risk.
▼ Denotes statistically significant *decrease* in risk.

APPENDIX: ANALYTIC EPIDEMIOLOGY STUDIES
COSMETIC TALC AND RISK OF OVARIAN CANCER (Continued)

CITATION, LOCATION	STUDY TYPE, POPULATION	EXPOSURE RATE OF GENITAL TALC IN CONTROLS AND DEFINITION	RISK ESTIMATES		COMMENTS
Tzonou et al. 1993 Athens, Greece	Case-control (hospital-based) Cases were 189 women less than 75 years of age who underwent surgery for a histologically confirmed common malignant epithelial ovarian tumor from 6/1989-3/1991. Cases were residents of greater Athena. Controls were 200 hospital visitors less than 75 years of age who were visiting patients in the same wards as the cancer patients at the same time.	Any perineal exposure: 3.5%. Subjects were asked to report the frequency of use (over an extended period before the onset of the present disease for cases, or a comparable period before the interview for controls) of talc in the perineal region (no, yes).	**Talc Application in the Perineum** No Yes	**Relative Risk (95% CI)** 1.0 (reference) 1.05 (0.28-3.98)	Cases and controls interviewed in the hospital. The authors concluded that although the number of talc users is in general small and the respective confidence intervals fairly large, there is clearly no evidence of an increased risk associated with the perineal application of talc. Interviews were conducted with 90 percent of eligible cases and 94 percent of eligible controls. Data on frequency and duration of talc use were not collected. The reported risk estimate was adjusted for a number of possible confounders although adjustment with such a low exposure rate might not be meaningful.

* Denotes statistically significant *increase* in risk.
▼ Denotes statistically significant *decrease* in risk.

APPENDIX: ANALYTIC EPIDEMIOLOGY STUDIES
COSMETIC TALC AND RISK OF OVARIAN CANCER (Continued)

CITATION, LOCATION	STUDY TYPE, POPULATION	EXPOSURE RATE OF GENITAL TALC IN CONTROLS AND DEFINITION	RISK ESTIMATES		COMMENTS
Whittemore et al. 1988 Northern California	Case-control (population and hospital-based) Cases were 188 women aged 18-74 diagnosed with primary epithelial ovarian cancer from 1/83 to 12/85 at one of seven hospitals. Of the 539 controls, 280 were hospitalized women and 259 were women selected from the general population via random digit dialing. Controls were matched to cases on age, race and additional criteria.	Perineal dusting; 41%.	**Type of Talc Use** None Perineum only Sanitary pads only Diaphragm only Any two of above All three of above **Years of Talc Use** None 1-9 10+ **Applications of Talc Per Month** None 1-20 20+	**Relative Risk (95% CI)** 1.00 (reference) 1.45 (0.81-2.60) 0.62 (0.21-1.80) 1.50 (0.63-3.58) 1.36 (0.91-2.04) 0.35 (0.04-2.94) 1.00 (reference) 1.60 (1.00-2.57) 1.11 (0.74-1.65) 1.00 (reference) 1.27 (0.82-1.96) 1.45 (0.94-2.22)	As this study used a case-control design, recall and/or selection bias may have affected the results. The most salient biases of this study include the failure to interview all eligible ovarian cancer patients and a completely random sample of controls, as well as the potential pitfalls of combining the two control groups. A significant trend was not observed with increasing duration or frequency of use. All risk estimates were adjusted for parity. In addition, the risk estimates stratified by type of talc use were also adjusted for oral contraceptive use. No association was found with talc-dusted diaphragm.

* Denotes statistically significant *increase* in risk.
▼ Denotes statistically significant *decrease* in risk.

APPENDIX: ANALYTIC EPIDEMIOLOGY STUDIES
COSMETIC TALC AND RISK OF OVARIAN CANCER (Continued)

CITATION, LOCATION	STUDY TYPE, POPULATION	EXPOSURE RATE OF GENITAL TALC IN CONTROLS AND DEFINITION	RISK ESTIMATES		COMMENTS
Wong et al. 1999 New York	Case-control (hospital-based) Cases were 499 patients with epithelial ovarian cancer treated at Roswell Park Cancer Institute from 10/1982-10/1995. Controls were 755 patients who were treated for non-gynecologic malignancies during the same period. Controls were frequency matched to cases on age at diagnosis.	Ever use: 48.7%. Subjects were asked about their method of talc application, and the duration of use.	**Talc Use by Site** Never used Sanitary napkin Genital or thigh area Both **Duration of Talc Use** None 1-9 yr. 10-19 yrs. ≥20 yrs. **No History of Surgical Interruption of Genital Tract** Nontalc user Talc user **History of Tubal Ligation or Hysterectomy** Nontalc user Talc user **No History of Hysterectomy within 5 Years of Diagnosis** Nontalc user Talc user	**Odds Ratio (95% CI)** 1.0 (reference) 0.9 (0.4-2.0) 1.0 (0.8-1.3) 1.1 (0.7-1.7) 1.0 (reference) 0.9 (0.6-1.5) 1.4 (0.9-2.2) 0.9 (0.6-1.2) 1.0 (reference) 1.2 (0.8-1.6) 1.0 (reference) 0.8 (0.5-1.2) 1.0 (reference) 0.9 (0.4-2.2)	Incident cases were interviewed in the hospital. Response rate was 93 percent of cases and 92 percent of controls. A dose-response trend was not observed with increasing duration of use. The reported risk estimates were adjusted for parity, oral contraceptive use, smoking history, family history of epithelial ovarian cancer, age at menarche, menopausal status, income, education, geographic location, and history of tubal ligation or hysterectomy. The authors note that the current study was limited to the use of talc on the perineum or sanitary napkin and did not address potential talc exposure from condom and diaphragm use. There were no significant associations between talc use and specific histologic subtypes of ovarian cancer.

* Denotes statistically significant *increase* in risk.
▼ Denotes statistically significant *decrease* in risk.