A quantitative weight of evidence assessment of Hill’s guidelines for causal inference for cosmetic talc as a cause of mesothelioma

A. Michael Ierardi a,*, Ania Urban b, Gary M. Marsh c

a Cardno ChemRisk, 231 Front Street, Suite 212, Brooklyn, NY, 11201, USA
b Cardno ChemRisk, 235 Pine Street, 23rd Floor, San Francisco, CA, 94104, USA
c Cardno ChemRisk, 20 Stanwix Street, Suite 505, Pittsburgh, PA, 15222, USA

A R T I C L E   I N F O
Keywords:
Sir Bradford Hill
Causal criteria
Causal viewpoints
Asbestos
Pleural mesothelioma
Risk assessment

A B S T R A C T
Cosmetic talc has been suggested to cause mesothelioma. To assess a potential causal relationship between cosmetic talc and mesothelioma, a quantitative weight of evidence analysis was performed in accordance with Hill’s nine original guidelines for causal inference using a published empirical model to weight each respective guideline. Various epidemiological, toxicological, and exposure studies related to cosmetic talc and risk of mesothelioma were included in an evaluation of each of Hill’s guidelines. Probabilities that the guidelines were true were assigned based on expert judgment. We applied a sensitivity analysis to evaluate the variability of our probability estimates. The overall probability of causality for cosmetic talc and mesothelioma was approximately 1.29% (range: 0.73%–3.96%). This low probability of causality supports the conclusion that cosmetic talc is not related to the development of mesothelioma.

1. Introduction

Causal inference is an important and nuanced art that is central to the charge of physicians, epidemiologists, risk assessors, and other public health professionals in identifying and defining a cause(s) of a specific disease (Weed, 2000). The application of causal inference allows health professionals to integrate evidence from various sources and fields of study in order to make an informed decision as to whether there is sufficient evidence to conclude that a causal relationship, as opposed to a non-causal association, exists between an agent and disease.

One of the most widely used tools for causal inference is the approach set forth by Sir Austin Bradford Hill in 1965 during the proceedings of a Royal Society of Medicine meeting (Hill, 1965). Based on his early investigative work on the relationship between cigarette smoking and the development of lung cancer, Hill outlined his guidelines for the evaluation of potential causal relationships between environmental and/or occupational exposures and disease outcomes. The specific intent for these guidelines was to aid health professionals in deciding whether causation is the “most likely interpretation” of an observed association between two variables (Hill, 1965, p. 295). The original guidelines as described by Hill include: 1) Strength, 2) Consistency, 3) Specificity, 4) Temporality, 5) Biological Gradient, 6) Plausibility, 7) Coherence, 8) Experiment, and 9) Analogy. Occasionally, Hill’s guidelines for causal inference are referred to as ‘Hill’s causal criteria.’ However, we specifically elected to refer to these as ‘guidelines,’ since none of these items is, in and of itself, a necessary characteristic of causality, except for ‘Temporality’ (Fedak et al., 2015; Gordis, 2013). As such, ‘Hill’s guidelines for causal inference’ is our preferred phrase herein. Moreover, Hill himself stressed:

“What I do not believe – and this has been suggested – is that we can usefully lay down some hard-and-fast rules of evidence that must be obeyed before we accept cause and effect. None of my nine viewpoints can bring indisputable evidence for or against the cause-and-effect hypothesis and none can be required as a sine qua non. What they can do, with greater or less strength, is to help us to make up our minds on the fundamental question – is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect?” (Hill, 1965, p. 295).

Regardless of how these guidelines are referred to, however, it has been generally agreed that they represent a reliable and reproducible framework to determine whether an observed association can, in fact, be considered causative. Since their initial publication, Hill’s guidelines for causal inference have been broadly applied to assess various potential

* Corresponding author.
E-mail addresses: Michael.Ierardi@cardno.com (A.M. Ierardi), Ania.Urban@cardno.com (A. Urban), Gary.Marsh@cardno.com (G.M. Marsh).

https://doi.org/10.1016/j.taap.2021.115461
Received 9 December 2020; Received in revised form 8 February 2021; Accepted 15 February 2021
Available online 20 February 2021
0041-008X/© 2021 Elsevier Inc. All rights reserved.

Article made available for educational purposes. Any further re-distribution or sharing is expressly prohibited.
exposure-disease relationships, including microbes and chronic gastrointestinal disease (Lowe et al., 2008), power-frequency electromagnetic fields and cancer (Moulder and Foster, 1995), and, more recently, talc and mesothelioma (Kanarek and Liegel, 2020) and ovarian cancer (Goodman et al., 2020).

Traditionally, investigators who undertake Hill’s analyses to assess whether observed associations between various agents and disease outcomes are causal do so in a qualitative manner by evaluating each of Hill’s guidelines on an equally weighted basis regarding the relative importance of each guideline to the overall evaluation of potential causality. To address some of the shortcomings sometimes associated with such a qualitative analysis, Swaan and van Amelsvoort (2009) provided an empirical basis by which each of Hill’s guidelines were weighted following a discriminant analysis. By following this systematic application of Hill’s guidelines, investigators can more transparently and systematically estimate the probability of a causal association (e.g., Hughes and colleagues (2014)). To our knowledge, this quantitative weight of evidence approach has not yet been applied to cosmetic talc and mesothelioma.

The potential causative relationship between exposure to cosmetic talc and mesothelioma is of particular interest considering recent, high visibility cases being litigated across the U.S. Briefly, the core argument in support of the assertion that cosmetic talc exposures cause mesothelioma is that cosmetic talc source mines contain detectable amphiboles, such as tremolite and anthophyllite, and possibly chrysotile (i.e., asbestiform serpentine); therefore, it has been alleged that consumer use of cosmetic talcum powder products that potentially contain asbestos can lead to an increased risk of developing this disease (Kanarek and Liegel, 2020; Emory et al., 2020; Gordon et al., 2014; Moline et al., 2020).

Yet multiple investigators, including epidemiologists and physicians who have conducted large cohort studies of cosmetic talc miners and millers over prolonged periods of time, have unanimously concluded that cosmetic talc source mines do not contain detectable asbestos (Lightfoot et al., 1972; Pooley, 1976; Rubino et al., 1976; Rubino et al., 1979; Boundy et al., 1979; Selevan et al., 1979; Parkes, 1982; Wegman et al., 1982; Wergeland et al., 1990; Wergeland et al., 2017; Wild et al., 2002; Coggiola et al., 2003; Pira et al., 2017; Fordyce et al., 2019; Pooley, n.d.). It would therefore follow that consumer use of products that contain cosmetic talc sourced from these mines would not lead to an increased risk of developing mesothelioma, as talc not containing asbestiform fibers has not been associated with an increased risk of cancer in general (International Agency for Research on Cancer (IARC), 2010).

Indeed, previous investigators have evaluated the pooled mortality experience of cosmetic talc miners and millers from Italy, Norway, France, Austria, and Vermont (Finley et al., 2017; Marsh et al., 2019; Ierardi and Marsh, 2020). These workers would have likely experienced much higher cumulative exposures to talc than a typical end-user of cosmetic talcum powder products (Aylott et al., 1979; Brown, 1985; Burns et al., 2019; Dement et al., 1972; Hildick-Smith, 1976; Moon et al., 2011; Rasmussen et al., 2019; Russell et al., 1979; Swanson, 1986; U.S. Environmental Protection Agency (USEPA), 1992; Zazenski et al., 1995), and would therefore be expected to have an increased risk of developing mesothelioma, if any such risk actually existed. Yet there is currently no epidemiological evidence to suggest that these workers are at an increased risk of mesothelioma (Ierardi and Marsh, 2020; Marsh and Ierardi, 2020).

The discrepancy between these two lines of argument is likely due, at least in part, to the analytical issues surrounding the proper identification of the amount, type, and habit of tremolite, anthophyllite, and serpentine minerals in talc (Gordon et al., 2014; International Agency for Research on Cancer (IARC), 2010; Swanson, 1986; Cralley et al., 1968; U.S. Food and Drug Administration (USFDA), 1971; Lewin, 1972; Snider et al., 1972; Cancer, 1973; Weissler, 1973; Rohl and Langer, 1974; Rohl et al., 1976; Krause, 1977; Rohl and Langer, 1979; Addison and Langer, 2000; Anderson et al., 2017; Pierce et al., 2017). It should be noted that the distinction between asbestiform and non-asbestiform is critical in the evaluation of the potential toxicity of amphibole and serpentine minerals, as it has been demonstrated that the non-asbestiform varieties of such minerals do not possess biological activity (Consumer Product Safety Commission (CPSC), 1988; American Thoracic Society (ATS), 1990; Occupational Safety and Health Administration (OSHA), 1992; Vu, 1993; Agency for Toxic Substances and Disease Registry (ATSDR), 2001; Addison and McConnell, 2008; Gamble and Gibbs, 2008; Mossman, 2008; Mossman, 2018; Williams et al., 2013). The asbestiform varieties of these amphibole and serpentine minerals (simply referred to herein as ‘asbestos’), however, may indeed lead to an increased risk of developing an asbestos-related disease at sufficient and prolonged exposures (Agency for Toxic Substances and Disease Registry (ATSDR), 2001; Finley et al., 2012; Pierce et al., 2008; Pierce et al., 2016; Gaffney et al., 2017).

In spite of these analytical fiber identification issues, it has been demonstrated by researchers, including scientists affiliated with the U.S. Food and Drug Administration (FDA), that even if one were to assume up to 0.1% asbestos content in cosmetic talc, as a “worst-case” scenario, the hypothetical cumulative asbestos exposures generated during routine use of cosmetic talcum powder products and subsequent risk of disease are “orders of magnitude below upper-bound estimates of cumulative asbestos exposure and risk at ambient levels, which have not been associated with increased incidence of asbestos-related disease” (Brown, 1985; Burns et al., 2019, p. 2272; Swanson, 1986).

It is evident that there currently exists ample information surrounding the potential human health effects associated with exposure to cosmetic talc. The aim of this study was therefore to apply each of Hill’s nine guidelines for causal inference to the available health effects data on cosmetic talc and to quantitatively assess the probability that any association between cosmetic talc and mesothelioma is causal.

2. Methods

2.1. Literature search

A review of the relevant literature related to cosmetic talc exposures and the disease endpoint of interest, mesothelioma, was undertaken. Specifically, we searched for information that described the epidemiological and toxicological data associated with cosmetic talc, and that quantified occupational and non-occupational cosmetic talc exposures (i.e., industrial hygiene or exposure studies). Articles published in the peer-reviewed literature, as well as unpublished reports, book chapters, and government documents were considered for inclusion. The available information was incorporated under one or more of the relevant Hill guidelines for causal inference, which is how our analysis is presented below.

2.2. Quantitative weight of evidence analysis

The overall probability of causality (P%) for mesothelioma was calculated according to Eq. 1 (Swaan and van Amelsvoort, 2009).

\[
P% = \frac{e^X}{e^X + e^Y} \times 100
\]

(1)

In which,

\[
X = -14.7799 + c_1 \times 0.06223 + c_2 \times 0.04061 - c_3 \times 0.02778 + c_4 \times 0.07657 - c_5 \times 0.03528 + c_6 \times 0.23025 - c_7 \times 0.0009621 + c_8 \times 0.00843 - c_9 \times 0.01294
\]

(2)

\[
Y = -10.80346 + c_1 \times 0.01923 + c_2 \times 0.01803 - c_3 \times 0.03877 + c_4 \times 0.08281 - c_5 \times 0.03534 + c_6 \times 0.21689 - c_7 \times 0.00334 - c_8 \times 0.00659 - c_9 \times 0.01011
\]

(3)
Where,
\[ \begin{align*}
    c_1 &= \text{Probability that the Strength guideline is true (\%)} \\
    c_2 &= \text{Probability that the Consistency guideline is true (\%)} \\
    c_3 &= \text{Probability that the Specificity guideline is true (\%)} \\
    c_4 &= \text{Probability that the Temporality guideline is true (\%)} \\
    c_5 &= \text{Probability that the Biological Gradient guideline is true (\%)} \\
    c_6 &= \text{Probability that the Plausibility guideline is true (\%)} \\
    c_7 &= \text{Probability that the Coherence guideline is true (\%)} \\
    c_8 &= \text{Probability that the Experiment guideline is true (\%)} \\
    c_9 &= \text{Probability that the Analogy guideline is true (\%)}
\end{align*} \]

The probabilities that each of Hill’s guidelines were true given the available data were based on expert judgment. These probabilities are presented as \( c_1 \) to \( c_9 \) at the end of the guideline evaluations and with justifications in Table 1. The coefficients (or weights) presented in Eqs. 2 and 3 were taken directly from the empirical discriminant function model published by Swaen and van Amelsvoort (2009). General recommendations for probability assignment were provided in part by Swaen and van Amelsvoort (2009).

### 2.3. Sensitivity analysis

While Swaen and van Amelsvoort (2009, p. 275) recognized that the assignment of the probabilities \( c_1 \) to \( c_9 \) is a “purely subjective exercise” based on expert judgment, they provided no guidelines on how to incorporate this source of variability into the estimation of \( P\% \). Thus, to evaluate the extent to which our estimate of \( P\% \) was robust with respect to variation in the assigned probabilities \( c_1 \) to \( c_9 \), we conducted a sensitivity analysis. In this analysis, we first determined which of the nine probabilities were reasonable to assume as fixed as 0% or 100%, and not subject to expert variability (e.g., \( c_4 \) Temporality = 100%). For the remaining probabilities, we bracketed our point or central estimate with reasonable low and high alternative values using two systematic approaches: Approach 1) \(-/+\) 50% of central estimate, and Approach 2) \(-/+\) 100% of central estimate. We then calculated \( P\% \) for all possible combinations of the fixed and variable (low, central, high) probabilities, and report descriptive statistics for the distribution of the resulting \( P\% \) values, including the minimum, 25%-tile, median, mean, 75%-tile, maximum, and range. This analysis provides an empirical estimation of the error associated with the expert judgment component of the process and the corresponding overall causal probability estimate (\( P\% \)).

### 3. Results

#### 3.1. Estimation of probabilities for Hill’s guidelines

##### 3.1.1. Strength

The ‘Strength’ of an association may be evaluated according to both the magnitude of an effect or risk estimate (e.g., Relative Risk [RR], Odds Ratio [OR], or Standardized Mortality Ratio [SMR]) and its statistical significance; notably, “[t]he stronger the association, the more likely it is that the relation is causal” (Gordis, 2013, p. 251). Contrarily, Hill (1965) suggested that small associations could more conceivably be attributed to other underlying contributors (i.e., bias or confounding) and, therefore, are less indicative of causation.

Based on the most recent data for each cosmetic talc miner/miller cohort (Wergeland et al., 2017; Wild et al., 2002; Pira et al., 2017; Fordyce et al., 2019), only one case of mesothelioma (in the Vermont cohort) has been reported and confirmed across all five cohorts. Ierardi and Marsh (2020) therefore calculated a deficit in mesothelioma risk (SMR = 0.299; 95% Confidence Interval (CI): 0.015, 1.42) for the pooled studies using a mid-value estimate of 3.34 total expected number of mesotheliomas and the one observed case of mesothelioma. Thus, according to the best available epidemiological data from robust cohort studies, this low (i.e., less than the null value of one) pooled SMR estimate suggests that cosmetic talc miners and millers are not at an increased risk of mesothelioma.

\( P\% \) = \( 100\% \) of central estimate, and Approach 2) \(-/+\) \( 100\% \) of central estimate. We then calculated \( P\% \) for all possible combinations of the fixed and variable (low, central, high) probabilities, and report descriptive statistics for the distribution of the resulting \( P\% \) values, including the minimum, 25%-tile, median, mean, 75%-tile, maximum, and range. This analysis provides an empirical estimation of the error associated with the expert judgment component of the process and the corresponding overall causal probability estimate (\( P\% \)).

### Table 1

Quantitative weight of evidence analysis using Hill’s guidelines of causal inference for the association between cosmetic talc and mesothelioma.

<table>
<thead>
<tr>
<th>Guideline</th>
<th>WeightX</th>
<th>WeightY</th>
<th>Probability of guideline being true (%)</th>
<th>Justification</th>
<th>WeightX × probability</th>
<th>WeightY × probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-14.7799</td>
<td>-10.80346</td>
<td>0</td>
<td>-</td>
<td>-14.7799</td>
<td>-10.80346</td>
</tr>
<tr>
<td>Strength (c1)</td>
<td>0.06223</td>
<td>0.01923</td>
<td>20</td>
<td>Deficit in mesothelioma risk (pooled SMR = 0.299; 95% CI: 0.015, 1.42) among cosmetic talc miner/miller cohorts</td>
<td>1.245</td>
<td>0.3846</td>
</tr>
<tr>
<td>Consistency (c2)</td>
<td>0.04061</td>
<td>0.01803</td>
<td>0</td>
<td>No increased risk of mesothelioma consistently reported across various cosmetic talc miner/miller cohort studies and pleurodesis studies</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Specificity (c3)</td>
<td>-0.02787</td>
<td>-0.03877</td>
<td>0</td>
<td>Talc not containing asbestos or asbestiform fibers is not a known human carcinogen</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Temporality (c4)</td>
<td>0.07657</td>
<td>0.08281</td>
<td>100</td>
<td>Exposure preceded disease outcomes in the cosmetic talc miner/miller cohort studies</td>
<td>7.657</td>
<td>8.281</td>
</tr>
<tr>
<td>Biological Gradient (c5)</td>
<td>-0.03528</td>
<td>-0.03534</td>
<td>30</td>
<td>No clear dose-response was observed for cosmetic talc exposures (consumer v. occupational exposures), but was investigated</td>
<td>-1.058</td>
<td>-1.060</td>
</tr>
<tr>
<td>Plausibility (c6)</td>
<td>0.23025</td>
<td>0.21689</td>
<td>0</td>
<td>Talc is non-genotoxic under in vitro and in vivo conditions, and experimental animal studies are negative for mesothelioma</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Coherence (c7)</td>
<td>0.0009621</td>
<td>-0.00034</td>
<td>30</td>
<td>Increased risk of mortality from pneumoconiosis and NMRD among some cosmetic talc miner/miller cohorts</td>
<td>0.02886</td>
<td>-0.1002</td>
</tr>
<tr>
<td>Experiment (c8)</td>
<td>0.00843</td>
<td>-0.00659</td>
<td>5</td>
<td>No increase in mesothelioma incidence among U.S women following peak talc usage</td>
<td>0.0422</td>
<td>-0.0330</td>
</tr>
<tr>
<td>Analogy (c9)</td>
<td>-0.01294</td>
<td>-0.01011</td>
<td>30</td>
<td>No evidence of mesothelioma risk in analogous non-asbestiform-exposed cohorts</td>
<td>-0.38820</td>
<td>-0.30300</td>
</tr>
</tbody>
</table>

**Sum**

**Overall Probability of Causality (P%)**

NMRD: Non-Malignant Respiratory Disease.

\(^a\) According to the empirical discriminant function model previously published by Swaen and van Amelsvoort (2009).

\(^b\) As reported by Ierardi and Marsh (2020).

\(^c\) Calculated according to Eq. 1.
Toxicology and Applied Pharmacology 417 (2021) 115461

4

The available cosmetic talc miner/miller cohort studies represent distinct populations of men, spanning five countries, and over a wide range of years (1940 to 2013), yet have a common occupational exposure agent and consistently report no increased risk of mesothelioma. In fact, up until 2019, not a single case of mesothelioma had been confirmed in any of these cohorts in the published literature. With the recent mesothelioma case described by Fordyce and colleagues (2019), there is now one confirmed case of mesothelioma among 4151 cosmetic talc miners and millers, contributing a total of 130,514 person-years of observation (Ierardi and Marsh, 2020). As noted by Fordyce and colleagues (2019), their finding of one case of mesothelioma in the Vermont cohort (a case they noted, which was unlikely related to cosmetic talc exposure) was not statistically significant; thus, when pooled with the other cohort studies, this single case of mesothelioma results in a deficit in mesothelioma risk, as detailed above (Ierardi and Marsh, 2020). In addition, after a follow-up period ranging from 14 to 40 years, no mesotheliomas were observed in any of over 300 patients who received talc pleurodesis treatments; this procedure is used to treat pleural effusion or collapsed lung and entails the injection of massive doses (2–10 g) of pharmaceutical-grade talc directly into the pleural cavity to prevent further accumulation of fluid (Finley et al., 2017; Biau et al., 2020; Xia et al., 2014; Grippi et al., 2015; Chappell et al., 1979; Lange et al., 1988; Viskum et al., 1989).

c₂ (Central Estimate = 0%; Fixed)

3.1.3. Specificity

For an association to be ‘Specific,’ a certain exposure must only be associated with one disease outcome. The application of this guideline can be difficult, as a given exposure may result in a variety of diseases. Indeed, Hill (1965, p. 297) noted that “[w]e must also keep in mind that diseases may have more than one cause ... [t]o-one-to-one relationships are not frequent.”

Sufficient and prolonged exposures of various types of asbestos fibers increase the risk of developing mesothelioma, and fiber type-specific exposure-response relationships have indeed been derived (Agency for Toxic Substances and Disease Registry (ATSDR), 2001; Finley et al., 2012; Pierce et al., 2008; Pierce et al., 2016; Gaffney et al., 2017). In fact, it has been reported that approximately 70% to 90% of pleural mesotheliomas in men in Europe and North America are attributable to asbestos, but that “some 60% to 90% of mesotheliomas in US women (pleural and peritoneal sites, respectively), and a substantial proportion of peritoneal mesotheliomas in men are likely unrelated to asbestos” (Attanoos et al., 2018, p. 758). Other non-asbestos risk factors, such as erionite and other mineral fibers, and radiation, have been implicated in the development of a portion of mesotheliomas (Attanoos et al., 2018).

Importantly, however, no government agency or public health organization has concluded that cosmetic talc exposure is associated with an increased risk of mesothelioma. For example, IARC in its review of talc exposures concluded that “[i]nhaled talc not containing asbestos or asbestosiform fibers is not classifiable as to its carcinogenicity” or Group 3 (International Agency for Research on Cancer (IARC), 2010, p. 412), and the American Conference of Governmental Industrial Hygienists (ACGIH) similarly classifies talc containing no asbestos fibers as “Not Classifiable as a Human Carcinogen” (American Conference of Governmental Industrial Hygienists (ACGIH), 2001, p. 1). Moreover, in a recent analysis of seven cohorts of talc miners and millers, Garabrant and Pastula (2018, p. 134) concluded that the current evidence suggests that occupational talc exposures are “non-potent” for mesothelioma. Thus, cosmetic talc exposures are neither specific for, nor associated with, mesothelioma.

c₁ (Central Estimate = 100%; Fixed)

3.1.4. Temporality

As noted, ‘Temporality’ is arguably the only one of Hill’s guidelines for causal inference that is required to be met to conclude a causal relationship between exposure and disease exists. In other words, an exposure must precede the development of disease for the exposure to be considered causative.

Due to the extended latency period of mesothelioma of approximately 20 to 40 years (Mazurek et al., 2017), it is also necessary to ensure that the cosmetic talc occupational cohorts were followed for a sufficient length of time to overcome the latency period associated with mesothelioma. In the most recently published update for each cohort, follow-up durations ranged from approximately 20 to 80 years (Wergeland et al., 2017; Wild et al., 2002; Pira et al., 2017; Fordyce et al., 2019), indicating sufficient follow-up durations in these studies were achieved. Furthermore, Marsh and colleagues (2019) recently performed a time since first employment, or latency, analysis for all of the cohorts, apart from the Vermont cohort, in order to approximate the total number of mesotheliomas contributed by those members of the Italian, Norwegian, French, and Austrian cohorts who had a latency period of at least 30 years. The authors estimated that 97.9% of the total expected mesotheliomas in the pooled cohort came from the subcohort of workers with a latency period of at least 30 years, indicating that there is sufficient latency across the cosmetic talc occupational cohorts to be able to detect mesothelioma development.

3.1.5. Biological gradient

The concept of ‘Biological Gradient,’ or dose-response, is central to the fields of epidemiology, toxicology, and human health risk assessment, among others. As dose increases, so should the response. The presence of a dose-response relationship between an agent and disease can be strong evidence of a causal relationship (Gordis, 2013). Regarding dose-response in the context of dust exposures experienced in industry, Hill (1965, p. 298) stated that “[t]he dustier the environment the greater the incidence of disease we would expect to see.”

There is a lack of evidence of the existence of a dose-response relationship between cosmetic talc exposure and mesothelioma. For example, assuming that an individual powdered herself once per day after showering throughout her 70-year lifetime, a reasonable estimate of 0.12 mg/m²-years can be calculated for her lifetime cumulative consumer exposure to talc (Russell et al., 1979). However, for a worker exposed to talc dust at the Permissible Exposure Limit (PEL) of 20 mmpcf set forth by the Occupational Safety and Health Administration (OSHA), which is equivalent to 3 mg/m³ (Occupational Safety and Health Administration (OSHA), 1992; National Institute of Occupational Safety and Health (NIOSH), 1988), a cumulative talc exposure of 135 mg/m³-years can be estimated for a 45-year working lifetime.

Even if the worker’s actual occupational talc exposure were 1/100th of the PEL (or 0.03 mg/m³), their cumulative occupational talc exposure (0.03 mg/m³ × 45 years = 1.35 mg/m²-years) would still be at least an order of magnitude greater than the hypothetical consumer’s lifetime cumulative talc exposure, as calculated above. It is also apparent that most cosmetic talc miners and millers experienced very high occupational talc dust exposures indicated by the significantly increased mortality from pneumoconiosis (i.e., “dusty lung” disease), as many investigators have previously noted (Fordyce et al., 2019; Marsh et al., 2019; Buffetta et al., 2018). Thus, it is illogical, from a dose-response perspective, to infer that consumer talcum powder exposures are a causal factor of mesothelioma, while occupational talc exposures, which have been demonstrated to be much more intense (i.e., higher
concentrations) and prolonged in nature, are not associated with an increased risk of mesothelioma.

c5 (Central Estimate = 30%; Range: Approach 1 − 15%–45%; Approach 2 = 0%–60%) 3.1.6. Plausibility

For an agent to be considered causal of some disease endpoint, it should be biologically plausible that the agent has caused the disease of interest. In the case of talc and cancer, no such biological plausibility exists. For instance, the formation of cancer is a multi-factorial process that generally involves cancer initiation followed by tumor promotion. Substances that can initiate cancer are typically referred to as genotoxic agents or mutagens due to their potential for damaging genetic material, which may lead to mutations (Clapp et al., 2008). However, numerous studies have documented that cosmetic talc is non-genotoxic when tested under both in vitro and in vivo conditions (International Agency for Research on Cancer (IARC), 2010; Endo-Capron et al., 1990; Endo-Capron et al., 1993; International Agency for Research on Cancer (IARC), 1987; Litton Bionetics Inc, 1974). Further, experimental animal studies, in which very high doses of talc were either inhaled by or directly injected into the pleura, peritoneum, or trabecula of rats, mice, and hamsters have not reported an increase in mesotheliomas following exposure (Endo-Capron et al., 1996; Pott et al., 1974; Jagatic et al., 1967; Ozesmi et al., 1985; Stenback and Rowlands, 1976; Wehner et al., 1977; Wehner et al., 1979; Wagner et al., 1977; Wagner et al., 1979; National Toxicology Program (NTP), 1993; Pickrell et al., 1989). No other proposed mechanisms of carcinogenicity have been demonstrated. Therefore, there is insufficient evidence to support a plausible biological mechanism by which talc could cause mesothelioma.

c6 (Central Estimate = 0%; Fixed) 3.1.7. Coherence

In describing ‘Coherence,’ Hill (1965, p. 298) stated that “the cause-and-effect interpretation of our data should not seriously conflict with the generally known facts of the natural history and biology of the disease … it should have coherence.” Swaen and van Amelsvoort (2009, p. 271) explain that this guideline “refers to other observed biological effects possibly relevant to the etiologic pathway that make a causal association more likely, for example, histological changes in the target organ.” Such histological lung data is unavailable for the cosmetic talc miner/miller cohorts; however, some of these workers have demonstrated a significantly increased risk of mortality due to pneumoconiosis and non-malignant respiratory disease (Wild et al., 2002; Pira et al., 2017; Fordyce et al., 2019).

c7 (Central Estimate = 30%; Range: Approach 1 − 15%–45%; Approach 2 = 0%–60%) 3.1.8. Experiment

As described in Hill (1965)’s original guidelines, ‘Experiment’ entails how an observed association might change following the introduction of a preventive measure, such as removal of the causative agent. Because experimental cancer studies in humans are unethical, an ecological analysis of cosmetic talc usage data and mesothelioma rates over time can be used as a reasonable proxy.

Cosmetic talc use is known to have steadily increased during the 1960s and peaked in the mid-to-late 1970s with usage tapering until present. Burns and colleagues (2018) hypothesized that if cosmetic talc use was associated with an increased risk of mesothelioma, peaks in cosmetic talc usage would be followed by an increase in mesothelioma incidence beginning in the late 1990s/early 2000s. However, an evaluation of annual usage of cosmetic talc by consumers in the U.S. (data sourced from the U.S. Geological Survey (USGS)) compared to the incidence of mesothelioma in females (data sourced from National Cancer Institute’s [NCI] Surveillance, Epidemiology, and End Results (SEER) 9th registry) indicated that while cosmetic talc usage peaked in approximately 1977 at nearly 70,000 metric tons, no subsequent increase in female mesothelioma rates was observed in the following decades (Burns et al., 2018). This finding is consistent with the other epidemiological evidence discussed herein.

c8 (Central Estimate = 5%; Range: Approach 1 = 2.5%–7.5%; Approach 2 = 0%–10%) 3.1.9. Analogy

Hill’s final guideline, ‘Analogy,’ can be fulfilled by identifying potential causal associations from similar enough exposure-disease relationships. For instance, it has been repeatedly demonstrated that talc deposits mined for cosmetic purposes do not contain detectable asbestos and may only contain non-asbestiform amphibole and serpentine minerals, if any at all (Lightfoot et al., 1972; Pooley, 1976; Rubino et al., 1976; Rubino et al., 1979; Boundy et al., 1979; Selevan et al., 1979; Parkes, 1982; Wegman et al., 1982; Wergeland et al., 1990; Wergeland et al., 2017; Wild et al., 2002; Coggiola et al., 2003; Pira et al., 2017; Fordyce et al., 2019; Pooley, n.d.). Therefore, an analogous example could entail other mineral deposits that potentially contain non-asbestiform amphibole and serpentine minerals, and whether these other mineral exposures have been associated with an increased risk of mesothelioma.

The series of Homestake gold miner epidemiological studies (McDonald et al., 1978; Gillam et al., 1976; Steenland and Brown, 1995a; Steenland and Brown, 1995b) comprise a useful analogous example since gold mined from this deposit has been associated primarily with non-asbestiform cummingtonite-grunerite (referred to as amosite when asbestiform), among other non-asbestiform varieties, including tremolite-actinolite (National Institute for Occupational Safety and Health (NIOSH), 2011). Although McDonald and colleagues (1978, p. 276) identified one case of pleural mesothelioma in their cohort, the authors noted that the individual’s diagnosis was “in doubt” and that he was “conceivably exposed to insulation materials.” To the best of our knowledge, no other cases of mesothelioma have been reported among these workers. It has therefore been concluded that “[t]here is very little evidence of an excess of mesothelioma in the studies of Homestake gold miners” (National Institute for Occupational Safety and Health (NIOSH), 2011, p. 26).

Additionally, taconite iron ore miners in Minnesota, who were also potentially exposed to non-asbestiform cummingtonite-grunerite, provide another informative analogous example, albeit slightly more nuanced with respect to potential mesothelioma risk. Mandel and Odo (2018) most recently summarized various epidemiological studies describing these workers and concluded that taconite miners are at an increased risk of mesothelioma, though Lambert and colleagues (2016, p. 108) noted that “[t]o date, the finding of excess mesothelioma in taconite workers is unique among studies of non-asbestiform amphiboles.”

Indeed, in their case-control study that stratified disease risk by zones of the mine, Lambert and colleagues (2016) did not find an elevated risk of mesothelioma (RR = 0.88; 95% CI: 0.71, 1.09) among taconite miners who worked in the eastern zone where non-asbestiform amphiboles were located, suggesting that non-asbestiform exposures did not increase the workers’ overall risk of mesothelioma. Rather, it has been estimated that <1% of the amphibole minerals in the taconite deposit are asbestiform and could possibly contribute to the observed mesothelioma risk (Mandel and Odo, 2018). Mandel and Odo (2018, p. 110) also acknowledged that “commercial asbestos was typically widely used in all taconite processing facilities,” which further confounds a definitive conclusion of increased mesothelioma risk due to prior taconite exposures (National Institute for Occupational Safety and Health (NIOSH), 2011).
A lack of mesothelioma risk related to non-asbestiform mineral exposures among these various occupational cohorts is not surprising, as it has been demonstrated that the non-asbestiform varieties of such minerals do not possess biological activity (Consumer Product Safety Commission (CPSC), 1988; American Thoracic Society (ATS), 1990; Occupational Safety and Health Administration (OSHA), 1992; Vu, 1993; Agency for Toxic Substances and Disease Registry (ATSDR), 2001; Addison and McConnell, 2008; Gamble and Gibbs, 2008; Mossman, 2008; Mossman; Williams et al., 2013). Thus, pure cosmetic talc exposures are benign for mesothelioma (International Agency for Research on Cancer (IARC), 2010), as are other analogous non-asbestiform mineral exposures. Additionally, no increased risk of asbestos-related disease has been associated with cosmetic talcum powder products, even when up to 0.1% asbestos content in cosmetic talc was assumed (Brown, 1985; Burns et al., 2019; Swanson, 1986).

\( c_9 \) (Central Estimate = 30%; Range: Approach 1 = 15%–45%; Approach 2 = 0%–60%)

3.2. Calculation of overall probability of causality (P%)

Table 1 summarizes the results of calculations leading to the estimate of P% based on our central estimates for \( c_1-c_9 \). The overall probability of causality (P%) for the association between cosmetic talc and mesothelioma was 1.29%.

3.3. Sensitivity analysis

As shown above, \( c_2, c_3, \) and \( c_6 \) were fixed as 0%, and \( c_4 \) as 100%. The resulting five variable probabilities \( (c_1, c_5, c_7-c_9) \) under each bracketing approach (low, central, high) led to \( 3^5 = 243 \) possible combinations of \( c_1-c_9 \) values and corresponding estimates of P%. For each central estimate bracketing approach, Fig. 1 shows the distribution of P% estimates (via box and whisker plots showing from bottom to top: minimum, 25%-tile, median, mean, 75%-tile, and maximum). Table 2 shows corresponding descriptive statistics. Table 2 also shows the median P% values under both Approach 1 and 2 (1.29%) were identical to the estimate shown in Table 1, which was based on the central estimates of \( c_1-c_9 \). Both approaches led to an overall small range of P% values of 0.73% to 3.96%. Thus, the most conservative assignment of values for \( c_1-c_9 \) led to only a 3.96% probability that cosmetic talc is causally associated with mesothelioma.

4. Discussion

The current study represents the first of its kind in which a quantitative weight of evidence analysis of cosmetic talc as a potential causative agent of mesothelioma was performed. Based on our analysis, we found that the data does not support a causal relationship between cosmetic talc and mesothelioma. Following our search of the relevant literature, we identified one study (Kanarek and Liegel, 2020) in which a Hill’s analysis was undertaken to evaluate the potential relationship between talc and mesothelioma.

The conclusions of Kanarek and Liegel (2020), however, directly contradict our own. These authors reported to have summarized and

$$\text{Table 2: Results of sensitivity analysis of overall causal probability (P%) estimation by central estimate bracketing approach.}$$

<table>
<thead>
<tr>
<th></th>
<th>Approach 1a</th>
<th>Approach 2a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>0.73</td>
<td>0.41</td>
</tr>
<tr>
<td>25%-tile</td>
<td>0.89</td>
<td>0.62</td>
</tr>
<tr>
<td>Median</td>
<td>1.29</td>
<td>1.29</td>
</tr>
<tr>
<td>Mean</td>
<td>1.37</td>
<td>1.62</td>
</tr>
<tr>
<td>75%-tile</td>
<td>1.86</td>
<td>3.13</td>
</tr>
<tr>
<td>Maximum</td>
<td>2.27</td>
<td>3.96</td>
</tr>
<tr>
<td>Range</td>
<td>1.54</td>
<td>3.55</td>
</tr>
</tbody>
</table>

\( ^a \) Based on all 243 possible combinations of fixed and variable probability estimates for \( c_1-c_9 \).

\( ^b \) –/+ 50% of central estimate.

\( ^c \) –/+ 100% of central estimate.

Fig. 1. Sensitivity analysis: Distribution of overall causal probability estimate (P%).

\( ^a \) The Figure shows box and whisker plots for the distributions of the causal probabilities computed under each bracketing approach used in the sensitivity analysis. Each box shows the interquartile range of the corresponding values (25%-tile is bottom of box; 75%-tile is top of box), the whiskers or lines below and above the box show values outside the interquartile range (minimum and maximum values shown as short horizontal lines), the mean value is shown as an ‘X’ within the box, and the median value as a horizontal line within the box. The specific summary statistics shown in the plots are provided in Table 2.

\( ^b \) Based on all 243 possible combinations of fixed and variable probability estimates for \( c_1-c_9 \).
analyzed the available talc epidemiological literature using the “Hill criteria of causality,” ultimately concluding that although the available epidemiological studies suffer from deficiencies, their analysis of “all the toxicological, biological and human studies” resulted in “compelling evidence for the existence of a causal relationship between talc and mesothelioma” (Kanarek and Liegel, 2020, p. 1, 8). However, their conclusion is largely based on misinterpretations of the available data.

For example, Kanarek and Liegel (2020) reported that Coggiola and colleagues (2003) found two peritoneal cancers in the Italian cohort, and so Kanarek and Liegel (2020) classified these as two cases of mesothelioma in Table 2 of their study. On the contrary, however, Coggiola and colleagues (2003) found no cases of peritoneal cancer; rather, it was Pira and colleagues (2017) who identified two cases of peritoneal cancer among the Italian cohort, yet Pira and colleagues (2017) specifically noted that the “[t]wo deaths from peritoneal cancer were from neoplasms other than mesothelioma.” Additionally, Kanarek and Liegel (2020) noted in their Table 2 that Mirabelli (2018) identified a case of pleural mesothelioma in an Italian talc mill worker; however, Pira and colleagues (2018), the original authors of the Italian cohort study mentioned by Mirabelli (2018), reported in a response to Mirabelli (2018) that they were unable to identify this case in their cohort roster and concluded that “[t]he number of observed deaths from pleural mesothelioma in [the Italian] cohort remains therefore zero.”

Kanarek and Liegel (2020) also highlighted many of the perceived limitations of the epidemiological cohort studies of cosmetic talc miners/millers, yet do not attempt to similarly describe limitations of the recently published cosmetic talc case series they cited (Emory et al., 2020; Moline et al., 2020), as others have previously noted (Ierardi and Marsh, 2020; Geyer, 2020a; Geyer, 2020b). Overall, Kanarek and Liegel (2020) focus mainly on critiques of the cosmetic talc epidemiological literature, and do not attempt to contextualize these epidemiological findings in their written analysis with any other available evidence, such as animal toxicity studies, risk assessments, and/or industrial hygiene/exposure studies (contrary to their statement that they included “all the toxicological, biological and human studies” in their analysis) (Kanarek & Liegel, 2020, p. 8), which all indicate no increased risk of mesothelioma following exposure to cosmetic talc, as demonstrated herein.

4.1. Limitations

Some researchers have previously argued that a Hill’s causal inference analysis should not be undertaken in the absence of a positive association between exposure and disease (Yarborough, 2006). With specific regard to cosmetic talc and mesothelioma, Ierardi and Marsh (2020) calculated a non-significant deficit in mesothelioma risk (SMR = 0.299; 95% CI: 0.015, 1.42) among five cohorts of cosmetic talc miners and millers, and determined that with 3.34 expected mesotheliomas among the pooled cohorts, eight or more mesotheliomas (or an SMR of 8/3.34 = 2.40 or greater) would need to be observed across the five pooled cohorts to reject at the 0.05 significance level the null hypothesis of no association (i.e., SMR = 1.0) between exposure to cosmetic talc and mesothelioma.

Accordingly, there is currently no overwhelmingly positive association that has been reliably demonstrated between cosmetic talc and mesothelioma in the available epidemiological literature. Yet an analysis of Hill’s guidelines for causal inference in this case (i.e., the absence of an initial positive association between cosmetic talc and mesothelioma) remains a useful and informative analysis.

Swan and van Amelsvoort (2009) derived their empirical model based solely on Group 1 and 2A carcinogens, as classified by IARC. Group 2B, 3, and 4 agents were excluded because of a general lack of epidemiological data for these agents in the IARC database. It was therefore recommended by the authors to use this quantitative Hill’s analysis approach “only in those instances that resemble the data sets available for category 1 and 2A agents”; otherwise, the authors noted that “[h]ad the category 2B and category 3 agents been addressed in the model, the more epidemiological weights might have received lower weights” (Swan and van Amelsvoort, 2009, p. 276). Following our review of the available literature, it is clear that the dataset for cosmetic talc and mesothelioma does not resemble those of Group 1 and 2A agents. Indeed, IARC has classified talc not containing asbestos or asbestiform fibers as Group 3 (International Agency for Research on Cancer (IARC), 2010). Even so, the low overall probability of causality for mesothelioma that was calculated in the current study (P% = 1.29%) suggests that any reasonable deviations in the assigned weights for each of Hill’s guidelines would likely have minimal impact on the overall conclusions of the evaluation. In fact, when such deviations (−/+ 50% and −/+ 100% of central estimate) were explored in a sensitivity analysis, a range of 0.73% to 3.96% was calculated, indicating that P% is sufficiently robust to even large deviations from the central estimates.

Furthermore, it should be noted that two guidelines (c4 [Temporal] and c5 [Analogy]) are not logically linked to the final value of P%. For instance, as probability estimates for each of these two guidelines increase, the overall P% should intuitively increase as well since greater probabilities that each guideline is true would support a stronger causal association. However, when probability estimates for c4 and c5 are increased, the overall P% actually slightly decreases, likely representing some statistical aspect of Swan and van Amelsvoort (2009)’s model that warrants further exploration. Regardless of this counterintuitive observation, the overall P% remains low (2.58%) if both c4 and c5 are set to 0% in the current analysis.

As described above, Hill (1965) indicated that the totality of evidence should be evaluated when assessing a potential causal relationship between exposure and disease. Notably, Gordis (2013, p. 253) also confirms that “[a]ny conclusion that an observed association is causal is greatly strengthened when different types of evidence from multiple sources support such reasoning. Thus, it is not so much a count of the number of guidelines present that is relevant to causal inference but rather an assessment of the total pattern of evidence observed that may be consistent with one or more of the guidelines.” Taken together, the pattern of evidence observed following an overall assessment according to Hill’s guidelines for causal inference indicates that cosmetic talc is not a causative agent of mesothelioma.

5. Conclusion

The current quantitative weight of evidence analysis using Hill’s guidelines for causal inference allows for the simultaneous integration and assessment of multiple lines of scientific evidence, including epidemiological data (e.g., cohort studies, case-control studies, ecological analyses), animal toxicity studies, industrial hygiene/exposure data, as well as carcinogenicity conclusions of government agencies and international scientific organizations. Based on this reproducible analysis of health effects information for cosmetic talc spanning multiple scientific disciplines, the available evidence supports the conclusion that cosmetic talc is not related to the development of mesothelioma.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

All of the authors are employed by Cardno ChemRisk, a consulting firm that provides scientific advice to the government, corporations, law firms, and various scientific/professional organizations. G. M. M. is Professor Emeritus of Biostatistics and Epidemiology, and Founding Director of the Center for Occupational Biostatistics and Epidemiology at the University of Pittsburgh, Graduate School of Public Health. This manuscript was prepared and written exclusively by the authors, and no external funding was received for this study. No organizations other
than to those which the authors belonged were aware of the preparation of this manuscript, and no other organizations reviewed any part of this manuscript prior to its submission for publication. Two of the authors (A. M. L., G. M. M.) have been retained as defense experts in cosmetic talc- and asbestos-related litigation.

References


Ierardi et al., 119–146.


Koch, Hill, and Crohn

Koch, Hill, and Crohn


March 2011. Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, Cincinnati, OH.