Are "Clean" Cosmetic Products Really Clean? A Pilot Study on the Validity of Cosmetic Labeling

Jeannie H. Phan (✉ jphan2@stu.mcphs.edu)

MCPHS University <https://orcid.org/0000-0003-2352-0511>

Khine K. Kyaw

MCPHS University

Julia K. Tan

MCPHS University

Emily Q. Lam

MCPHS University

Cheng-Kuan Lin

MCPHS University

Research

Keywords: cosmetic products, EWG, Sephora, clean products

Posted Date: September 27th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-917383/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

As a result of increased demand for environmental and health-conscious cosmetics, retailers have increasingly marketed their products using terms such as “clean” or “non-toxic”. Sephora, a popular beauty retailer, implemented a clean seal system to label and promote their products. This study aims to compare the toxicity concerns of clean vs. non-clean products by using the Environmental Working Group (EWG)'s Skin Deep framework.

Methods

EWG's Skin Deep Framework was used to rank products based on their cancer, allergy and immunotoxicity, developmental and reproductive toxicity, and use-restriction concerns, with lower scores indicating a cleaner product (ranging 0–10). The distributions of EWG scores among clean and non-clean products were investigated and stratified by different types of products (i.e. fragrance, hair, makeup, or skincare). A multivariable linear regression model was further applied to evaluate the association between EWG scores and clean seals, adjusting by types of products, level of toxicity concerns.

Results

356 products were screened, including 180 (50.56%) clean products and 176 (49.44%) non-clean products. Clean products yielded a higher percentage of low hazard ingredients compared to non-clean products. EWG scores are positively correlated with the levels of toxicity concern for cancer, allergy and immunotoxicity, respectively. Clean products are also associated with a lower EWG score by 0.71 in the regression model. Fragrance products are most hazardous, with a 2.42 increase in EWG score.

Conclusions

Products under the binary “clean beauty” labeling system at Sephora may not necessarily capture the nuances of EWG's ten-point scoring system and various dimensions of health concerns. It may be insufficient for consumers to solely rely on the presence of the clean seal when purchasing beauty products. Consulting multiple frameworks and sources to inform decision-making is crucial in addressing knowledge gaps. Further research and increased data availability from EWG's database are necessary as well as public education on the application of Sephora's clean seal.

Background

In recent years, “clean” cosmetics have transformed the cultural and economic landscape of the cosmetic industry as mindful consumerism increases (1). When purchasing beauty products, consumers are often

influenced by perceptions in social media and do not account for contributing safety factors, such as ingredient quantity (2). Utilizing “clean” cosmetics has become a priority for many individuals as these products are applied onto the skin daily for long periods of time. According to a survey conducted by the Environmental Working Group (EWG), American women use an average of 12 skincare products and are exposed to almost 168 unique ingredients on a daily basis (2). As a result of increased demand for environmental and health-conscious cosmetics, retailers have increasingly marketed their products using terms such as “clean” or “non-toxic”. From 2017 to 2018, the clean beauty market increased by 23%, contributing to 25% of the overall annual skincare sales (3). Although the increase in demand may have led to an increase in supply of products that are more “natural”, “clean”, and “healthier” in the cosmetic market, clearer labels are still absent in the market and queries still exist regarding the science of products in this domain (4).

Sephora, a major cosmetics retailer, has a section in-store and online that is dedicated to promoting the “clean” beauty products of various brands. Sephora promotes its credibility and reliability of these products’ clean status through a green seal and associated signage. This clean beauty seal indicates that the product is “formulated without parabens, sulfates SLS and SLES, phthalates, mineral oils, formaldehyde, and more” (5). The reliability of the seal has been widely discussed by various online media outlets, such as EcoCult, who believes that although the clean seal could be helpful for consumers, it is “not ideal for corporations to be self-regulating and deciding what constitutes a safe product” (6). Further academic research and critical examination of the seal is required to support its reliability in distinguishing clean and non-clean products. For example, phthalates, which are commonly found in many cosmetic products, are often used to soften and increase flexibility. Exposure to phthalates, however, “increases the risk of damage to the liver, kidneys, lungs, and the reproductive system” (7).

Currently the Food and Drug Administration (FDA) and Cosmetic Act have no regulation for the approval of beauty products before it is distributed to the public market (8). Without a standardized definition by the regulatory authority, companies are able to determine and market ‘clean’ or ‘safe’ beauty products with their own standards and are not legally obligated to report adverse health events (9). The interchangeable use of terms such as “clean” and “natural” are also not closely regulated by the U.S. Federal Trade Commission or the FDA (10). The usage of these terms may persuade the consumer to purchase associated products and inadvertently contribute to unknown disease burden (10). For example, a company had developed a conditioner without sulfates in an attempt to promote “natural” and “chemical-free” products, which ultimately resulted in a product recall in response to complaints of hair loss and itching from approximately 21,000 consumers (2). Alternatively, utilizing words synonymous with “toxic” may instill fear in consumers, given that several well-known products have been recalled in the last decade due to adverse health effects, phenomena closely related to the increasing trend of clean cosmetics (4). Long-term health effects of cosmetic use have also not been thoroughly reported or researched (11). Thus, Sephora’s clean seal may play an integral role in mobilizing the cosmetic industry’s efforts to be more transparent about ingredient safety and labeling.

To reinforce the reliability of Sephora's clean seal and better discern the differences between clean and non-clean products, the reliable and widely adopted framework of Environmental Working Group (EWG)'s Skin Deep database can be implemented. EWG's framework fills the knowledge gap left by the lack of regulatory measures for ingredients that allows companies to use almost any ingredient they desire in their marketed products. Compared to other notable databases that focus on just ingredients, EWG's Skin Deep database consists of more than 73,000 products and their ingredients in which they rank them based on their toxicity and health concerns (12). The majority of the products within the database can easily be found in Sephora or drugstores and are analyzed by credible toxicologists, chemists, and public health specialists (10). Their ranking system is easy to navigate and comprehend, which may encourage consumers to use the database prior to purchasing products of interest. EWG has also established standard criteria for ingredients and products based on rigorous evidence-based approaches (13). Through this framework, consumers have the potential to be more informed about harmful ingredients and chemicals in the products they purchase and use every day.

By applying EWG's scoring system to Sephora's clean products, this study aims to 1) determine the validity of the cleanliness of their products; 2) examine the toxicity/concerns of cancer, developmental and reproductive toxicity, allergies and immunotoxicity, and use restrictions of Sephora's clean products; and 3) evaluate the levels of 'cleanliness' among different types of products.

Methods

EWG's Ranking System

EWG's Skin Deep database consists of approximately 64,983 products (14). These products are analyzed and evaluated by scientists to determine whether or not products are deemed safe by using a standardized review process. Each cosmetic product and ingredient evaluated by EWG is assigned a hazard score, ranging from 1 to 10 (least to most hazardous). EWG identifies products that meet its strictest criteria with the EWG Verified™ mark, which is indicated by a hazard score of 0 in this study. EWG determines the score by utilizing a weight-of-evidence approach that factors in all associated health factors and health impacts (14). The score is also determined by scientific literature for each ingredient and is reviewed and updated periodically (14). Data availability for products on the EWG's Skin Deep database are categorized in three levels of "none or limited, fair, and good or robust" data available and are aggregated based on the scope of the health hazard and the volume of evidence (14). EWG states that some ingredients may have insufficient scientific literature, resulting in lower data availability for some products. Consequently, products may have low hazard ratings attributed to a lack of research studies conducted by EWG for the ingredients in those products. EWG recommends consumers should look for products with low hazard ratings and at least "fair" or better data availability.

Each cosmetic product is also assessed for toxicity concerns, including the main concerns of cancer, developmental and reproductive toxicity, allergies and immunotoxicity, and use restrictions (14). First, cancer concerns are associated with ingredients linked to cancer in various studies or assessments. Developmental reproductive toxicity concerns are associated with ingredients linked to health effects

such as reproductive organ cancers, birth defects, and developmental delays. Allergies and immunotoxicity concerns are associated with ingredients linked to allergic reactions or immune system impairment. Lastly, use restriction concerns are linked to ingredients that are restricted for cosmetic use based on industry or government requirements from several countries, including the U.S. (14). In addition to toxicity concerns, EWG also evaluates whether a product is cruelty-free (i.e. yes or no). Additional details for product evaluation and methodology are located on EWG's Skin Deep website.

Sephora's Clean Seal

Since 2018, cosmetic products sold on Sephora's website have been reviewed by the Sephora Clean program group. Products that do not contain 'parabens, phthalates, formaldehydes and formaldehyde-releasers, oxybenzone, coal tar, hydroquinone, and triclosan' and other ingredients listed in Appendix A would be granted a clean seal and promoted by Sephora (5). Every product is reviewed by Sephora's Clean program members and a clean seal would be given once a product is free from the above-listed ingredients (5). Every cosmetic product on Sephora's purchase website is categorized into four main product types: fragrance, hair, makeup, and skincare.

Data Extraction and Analyses

Four independently trained researchers manually merged collected data from EWG's and Sephora's respective databases. A total of 356 products were reviewed, which included all products that were simultaneously available on both Sephora's website and EWG's database. Data that was extracted for each product included brand name, product category, EWG score, data availability (i.e. none, limited, fair, good, robust), acquisition of clean seal (i.e. yes or no), last updated year, cruelty-free status, toxicity concerns, and quantity of ingredients. We summarized the numbers and percentages of low, moderate, and high hazardous ingredients, by category (i.e. fragrance, hair, makeup, and skincare) and by clean status (i.e. clean products or non-clean products). Average EWG scores with exact 95% confidence intervals were calculated and plotted by product categories and clean seal, respectively. We further summarized and depicted the distribution of EWG scores across clean vs. non-clean products. The association between four types of toxicity concerns and average EWG scores among clean vs. non-clean products were also presented by bar figures, respectively. Furthermore, the number and percentage of cosmetic products at different levels of the four different toxicity concerns and cruelty status among clean vs. non-clean products were summarized in the study.

A generalized linear regression model was performed to examine the association between EWG score and clean seal. Factor of product category, levels of allergy & immunotoxicity concerns, developmental & reproductive toxicity concern, and use restriction concerns were further adjusted in a full multivariable regression model. We treated the EWG scores as a continuous variable and applied a linear regression model because the EWG scores are normally distributed and the interpretation of results would be comprehensively feasible for the readers. A significance level of 0.05 (p-value) was set up *a priori* and 95 % confidence intervals were estimated by using normal quantiles methodology of *confint()* function. R version 3.5.2 (2018-12-20) was used for all analyses and figure presentations in the study.

Results

Our sample includes 180 clean and 176 non-clean products (n=356). The mean EWG scores between clean and non-clean products show a difference of 0.88, with clean products having a mean score of 3.23 and non-clean products having a mean score of 4.11 (Table 1). Non-clean products consist of a higher number of ingredients compared to clean products (32 vs. 26 ingredients). Compared to non-clean products, clean products consist of a higher percentage of low-hazard ingredients and a lower percentage of moderate and high-hazard ingredients. Fragrance products have a higher mean for EWG score than the other product categories (Table 1).

Table 1. Basic characteristics of reviewed products

Products	n	EWG score*	Average ingredients per product	% of low hazard ingredients**	% of moderate hazard ingredients**	% of high hazard ingredients**
Clean	180	3.23	26.38	84.4	14.08	1.48
Makeup	54	2.69	21.94	87.85	10.97	1.18
Hair	24	3.92	27.96	78.54	18.48	2.98
Fragrance	2	5.5	7.75	25.81	54.84	19.35
Skincare	100	3.38	29.15	85.04	13.89	1.06
Non-clean	176	4.11	31.68	76.42	20.52	3.07
Makeup	54	3.59	24.5	79.14	17.23	3.63
Hair	20	4.85	32.15	69.98	25.51	4.51
Fragrance	4	6.5	10.75	27.9	62.79	9.3
Skincare	98	4.15	36.4	77.15	20.33	2.52

*Noted as mean (95%CI)

** Low: 0 – 2; moderate: 3 – 6; high: 7 – 10

Non-clean products have consistently higher EWG scores across all four product categories (Figure 1). Both clean and non-clean fragrances have the highest average EWG scores among all product categories, with a greater variation in scores.

Although Figure 1 indicates that non-clean products share higher scores across all product categories, the distribution of EWG scores shows a less clear trend with various overlap between clean and non-clean products (Figure 2). Generally, there is a greater distribution of clean products associated with lower EWG scores compared to non-clean products.

Table 2. Number and percentage of cosmetic products at different levels of different toxicity concerns

	Clean (n = 180)	Non-clean (n = 176)
Toxicity concern	n (%)	n (%)
Cancer concerns		
High	0 (0)	3 (1.70)
Moderate	3 (1.67)	6 (3.41)
Low moderate	30 (16.67)	31 (17.61)
Low	147 (81.67)	136 (77.27)
Developmental and reproductive toxicity		
High	2 (1.11)	8 (4.55)
Moderate	0 (0)	18 (10.23)
Low moderate	48 (26.67)	16 (9.09)
Low	130 (72.22)	134 (76.14)
Allergic and immunological concerns		
High	111 (61.67)	90 (51.14)
Moderate	22 (12.22)	52 (29.55)
Low moderate	29 (16.11)	15 (8.52)
Low	18 (10.00)	19 (10.80)
Use restrictions		
High	63 (35.00)	106 (60.23)
Moderate	77 (42.78)	62 (35.23)
Low moderate	1 (0.56)	0 (0)
Low	39 (21.67)	8 (4.55)
Cruelty free		
Yes	69 (38.33)	30 (17.05)
Unknown	111 (61.67)	146 (82.95)

The numbers and distributions of cosmetic products at different levels of toxicity concerns are different between clean and non-clean products. More than 95% of non-clean products have high or moderate concerns of using restricted ingredients, while the percentage among clean products is approximately 78% (Table 2). More than half of the products have high allergy and immunological concerns, regardless of the clean status. Notably, the percentage of high allergy and immunological concern is higher in clean products (61.67%) than non-clean products (51.14%). For cancer concerns and developmental and reproductive toxicity concerns, the majority of both clean (81.67%) and non-clean (77.27%) products share low concern. However, the number and percentages of high and moderate concerns of non-clean products are higher than that of clean products. There are more cruelty-free products among clean products, but cruelty-free status for most products for either group are unknown.

EWG scores capture the different levels of concerns of cancer and allergy and immunotoxicity, respectively, among non-clean products, but not in other concerns (Figure 3). Notably, the average scores of clean products are always lower than that in non-clean products, even within the same level of concern.

Clean products are associated with a 0.71 lower EWG score, adjusted for types of products, toxicity, and year (Table 3). Compared to the other product types, fragrances have a higher EWG score by 2.42. Products of high cancer concern are associated with an increase of EWG score by 1.61, while products of moderate and low moderate cancer concern are associated with a 0.50 increase of EWG score. Compared to products of low use restriction concerns, products of moderate and high use restriction concern are associated with an increase of EWG score by 1.71 and 2.15, respectively. The EWG score of products with high allergy and immunotoxicity concern is higher by 1.76 compared to low allergy and immunotoxicity concern. These trends do not necessarily apply to developmental and reproductive concerns.

Table 3. Regression Model

	Estimate	LL	UL	Std. error	t- value	Pr(> t)
Clean Status						
(Intercept)	-55.49	-187.17	76.19	67.19	-0.83	0.41 .
Clean	-0.71	-0.96	-0.45	0.13	-5.47	<0.001 ***
Non-clean	1 (reference)					
Product Category						
Makeup	0.2	-0.07	0.47	0.14	1.44	0.15 .
Hair	0.42	0.05	0.79	0.19	2.21	0.03 .
Fragrance	2.42	1.57	3.27	0.43	5.59	<0.001 ***
Skincare	1 (reference)					
Cancer Concern						
High	1.61	0.33	2.89	0.65	2.46	0.01 .
Moderate	2.26	1.51	3	0.38	5.93	<0.001 **
Low Moderate	0.5	0.18	0.82	0.16	3.03	<0.001 **
Low	1 (reference)					
Developmental & Reproductive Toxicity Concern						
High	0.93	0.21	1.65	0.37	2.52	0.01 .
Moderate	1.47	0.89	2.04	0.29	5	<0.001 ***
Low Moderate	0.4	0.08	0.71	0.16	2.47	0.01 .
Low	1 (reference)					
Allergies & Immunotoxicity Concern						
High	1.76	1.33	2.19	0.22	8.01	<0.001 ***
Moderate	0.3	-0.15	0.76	0.23	1.32	0.19 .
Low Moderate	0.74	0.25	1.23	0.25	2.98	0 **
Low	1 (reference)					
Use Restriction Concern						
High	2.15	1.76	2.54	0.2	10.84	< 0.001 ***
Moderate	1.71	1.34	2.09	0.19	8.96	<0.001 ***
Low Moderate	-0.09	-2.26	2.08	1.11	-0.09	0.93
Low	1 (reference)					
Updated.Year	0.03	-0.04	0.09	0.03	0.83	0.4

Signif. Codes: *** = 0; ** = 0.001; * = 0.01; . = 0.05; none = 1

Discussion

This study was conducted to understand the difference in toxicity concerns between 356 non-clean and clean cosmetic products by using EWG's Skin Deep framework. Compared to non-clean products, clean products are cleaner (0.71 lower EWG score). Fragrance is associated with a higher EWG score than the other products (2.42; 95%CI: 1.57–3.27), likely attributable to the presence of ingredients related to worse

health effects and toxicity concerns. Among both clean and non-clean products, the increasing level of concern of cancer, allergy, and immunotoxicity are significantly correlated to an increase in EWG score, respectively.

This study is the first pilot study to compare Sephora's clean seal system, a popular marker for clean beauty products, and EWG's Skin Deep review system. We provided additional insight into the different levels of toxicity concerns for different types of beauty and lifestyle products, as well as possible areas of discordance between clean and non-clean products. For example, the EWG score takes into account more "unacceptable" ingredients using their gathered scientific research than Sephora's clean standard list. This discrepancy is important as consumers are more likely to purchase clean cosmetics when products contain eco-friendly labels without any knowledge of toxicity concerns (15). This study also provides an example of a methodology for consumers to utilize and adopt a framework such as EWG's Skin Deep's ranking system to better understand the specific health effects of products and their ingredients.

Although products with a clean seal are cleaner than non-clean products based on their EWG scores, 3.23 and 4.11 respectively, clean seals, defined as products without containing parabens, phthalates, formaldehydes, and more, may not be a good marker for consumers to purchase healthy cosmetic products, and potentially misleading. There are several areas of discordance that are evident among clean and non-clean products based on their EWG scorings and levels of toxicity concerns. The distribution of EWG scores of clean and non-clean products are largely overlapped (Fig. 1). Also, labeling with a clean seal does not guarantee less toxicity or concerns of cancer, development, immunotoxicity, and use restriction. Our results showed that the distributions of the four types of toxicities are similar among both clean and non-clean products (Table 2). The majority of both clean and non-clean products are associated with lower cancer concern, while the lowest number of products are associated with high concern (Table 2). However, when examining the distribution of EWG scores for clean and non-clean products based on the four main categories of toxicity concern, only non-clean products are associated with a positive trend between EWG score and concern for cancer as well as allergy and immunotoxicity concern (Fig. 3). The distribution of EWG scores for developmental and reproductivity concerns and use-restriction concerns indicate no significant association between EWG score and toxicity concern level. Also, both products have a majority of their respective products having higher allergy and immunotoxicity concerns compared to lower concern levels, but clean products consist of a higher percentage (61.67%) compared to non-clean products (51.14%) (Table 2). Both majorities of clean and non-clean products also fall under high or moderate use-restriction concerns.

There are several possible reasons that fragrance has a higher EWG score. Firstly, ingredients commonly found in fragrance products (and not as commonly found in other product categories) may contribute to a higher score. These ingredients generally have a higher concern for the four main toxicity categories, as well as higher concern for other categories such as endocrine disruption, irritation, and non-reproductive organ system toxicity (14). For example, benzyl alcohol acts as an aromatic agent, preservative, and solvent, and is associated with allergies and occupational hazards as well as non-reproductive organ toxicity (16). Secondly, fragrance compounds can be present in many different forms and are represented

with different names including fragrance, alcohols, esters, ketones, aldehydes, and alkalis are usually classified differently (17). Fragrance products with labeled terms such as “floral”, “exotic”, and “musky” may not accurately disclose their exact formulations and omit mixtures of natural and synthetic chemicals linked to reproductive problems (18). Thirdly, fragrance products have 14 different unlisted chemicals and 80% of fragrance products are not tested, on average (18). Results from product testing show that unlisted ingredients may include galaxolide and tonalide, which have shown potential for endocrine disruption in vitro studies and environmental toxicity in fish and crustacean growth functions (19). Fourthly, according to analyses conducted by EWG, the fragrance industry has published safety assessments for only 34% of the unlabeled ingredients (19). Although the available data is limited for fragrances (as summarized in Appendix B), EWG’s product testing and assessments have resulted in fair data availability for ingredients such as fragrance and benzyl alcohol. These ingredients may have a larger influence on final product scores when compared to other ingredients and ultimately result in higher scores among fragrance products.

This study has a few limitations. Although we included only products from Sephora which have been simultaneously reviewed by EWG Skin Deep, the fraction of included products ($n = 356$) is relatively low, compared to 8,043 cosmetic products on Sephora’s platform, and 71,774 from EWG. Selection bias is less likely for the following reasons: 1) The numbers of products listed by both systems might be overcounted. Many products containing exactly the same main ingredients but minor modifications on the proportions of ingredients, flavors, aroma, and/or colors were considered as multiple products. 2) EWG reviews products independently, without cherry-picking specific products and regardless of the clean seals. Therefore, the reviewed products fairly represent products on Sephora.

Products included what was available in Sephora’s “clean beauty” section that were examined and scored by EWG’s team of scientists. An increase of data simultaneously available on both websites would be beneficial in creating a more representative sample, especially regarding fragrance and hair products. This may be achieved through a higher rate of frequency in product updates or more active participation from both ends to exchange data. Also, some of the products listed on EWG’s Skin Deep database included old formulations that were not updated, which may not be accurately reflected in some of the EWG scores. Some products were also limited in data availability, suggesting that much of the literature for various ingredients is lacking or needs further investigation by EWG. Thus, EWG Skin Deep’s weight-by-evidence approach to its scoring system may not factor unavailable or limited information on the hazardous effects of certain ingredients.

Products under the binary “clean beauty” labeling system at Sephora may not necessarily capture the nuances of EWG’s ten-point scoring system. It may be insufficient for consumers to solely rely on the presence of the clean seal when purchasing beauty products. Our study, by utilizing a reputable framework to evaluate products, fills various knowledge and regulatory gaps that exist for cosmetic products. Consulting supplementary frameworks and primary data sources may reinforce and expand on existing data and criteria provided by EWG as well as address remaining gaps. Additionally, it would be beneficial to examine the economic and regulatory implications of ‘clean’ beauty in other countries (i.e.

cosmetics of prohibited use, other prioritized measures of toxicity concern, and the role of regulatory bodies in the cosmetic industry).

Declarations

Ethics Approval and Consent to Participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

Data used and analyzed in this study are publicly available from EWG.org and Sephora.com.

Competing Interests

The authors declare that they have no competing interests.

Funding

This research received grant funding from MCPHS University for the purpose of supporting journal publication costs as well as conference reimbursements. Conference attendance fee waivers were provided by the 33rd Annual Conference of the International Society for Environmental Epidemiology.

Authors' contributions

JP was a major contributor in the data collection for makeup, analysis of results and the writing of the results section, and management of the project overall. KK was a strong contributor in the collection of hair product data and the writing of the introduction section of the manuscript. JT played a major role in the data collection of skincare products and the writing of the discussion section of the manuscript. EL contributed to the data collection for fragrance products and the writing of the methods section. CL made major contributions by performing the regression analyses and making final revisions to manuscript drafts. All authors played a role in the interpretation of results and approval of the final manuscript.

Acknowledgements

Not applicable.

References

1. Nielsen Wellness Track, (2017). "Natural cosmetics" are those with natural product claims. The Nielsen Company Accessed 14 April 2021

2. Chandon L, (2020). Do Claims about the Naturalness and Dose of Cosmetics Ingredients Affect the Public's Perception of Their Safety?. *MDPI*. Accessed 14 April 2021
3. Rubin CB, Brod B, (2019) Natural Does Not Mean Safe—The Dirt on Clean Beauty Products. *JAMA Dermatol.* 2019;155(12):1344–1345. doi:10.1001/jamadermatol.2019.2724. Accessed 26 February 2020
4. Wanner M, Nathan N, (2019). Clean cosmetics: The science behind the trend. *Harvard Health Publishing*. Accessed 6 July 2020
5. Sephora, (2021) Clean At Sephora. *Sephora*. Accessed 23 March 2021
6. Wicker A, (2019). Can You Trust the Sephora Clean Seal?. *Ecocult*. Accessed 21 June 2021
7. Huang, P. C., Liao, K. W., Chang, J. W., Chan, S. H., & Lee, C. C., (2018). Characterization of phthalates exposure and risk for cosmetics and perfume sales clerks. *Environmental pollution (Barking, Essex : 1987)*, 233, 577–587. Accessed 13 April 2021
8. FDA, (2021). FDA Authority Over Cosmetics: How Cosmetics Are Not FDA-Approved, but Are FDA-Regulated. *U.S. Food & Drug Administration*. Accessed 27 March 2021
9. Kwa, M., Welty, L. J., Xu, S., Adverse Events Reported to the US Food and Drug Administration for Cosmetics and Personal Care Products. *JAMA Internal Medicine* (2017)
10. Wischhover C, (2018) The "natural" beauty industry is on the rise because we're scared of chemicals. *Vox*. Accessed 30 June 2020
11. Panico, A., Serio. F., Bagordo, F., et. al, (2019). Skin safety and health prevention: an overview of chemicals in cosmetic products. *Journal of Preventive Medicine and Hygiene*.
12. EWG's Skin Deep, (2015). About EWG's Skin Deep. *Environmental Working Group* Accessed 12 April 2020.
13. Pearson, B, (2019) Clean Beauty Can Be A Dirty Business: Beautycounter, Sephora And P&G Are Changing That. *Forbes*. Accessed 1 July 2020
14. EWG's Skin Deep, (2021). Understanding Skin Deep Ratings. *Environmental Working Group* . Accessed 12 April 2020
15. Borges L, Paananen A: Consumer perception and purchasing behavior towards green cosmetics: A market research. *Theseus* (2020)
16. Nair B. (2001). Final report on the safety assessment of Benzyl Alcohol, Benzoic Acid, and Sodium Benzoate. *International journal of toxicology*, 20 Suppl 3, 23–50. Accessed 14 April 2021
17. Niederer, M., Bollhalder, R., & Hohl., C. (2006, September 1). *Determination of fragrance allergens in cosmetics by size-exclusion chromatography followed by gas chromatography–mass spectrometry*. *Journal of Chromatography* Accessed 14 April 2021
18. EWG, (2021). Frequently Asked Questions. *Environmental Working Group*. Accessed 12 April 2020
19. Sarantis H, Naidenk O, V., Gray, S., et. al, (2020). Not So Sexy The Health Risks of Secret Chemicals in Fragrances. *Campaign for Safe Cosmetics*. Accessed 29 May 2021

Figures

Figure 1

EWG scores for clean vs. non- clean products by product category

Figure 2

Distribution of EWG scores across clean vs. non-clean products

Figure 3

Average EWG scores across different levels of toxicity concerns

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AppendixAB.docx](#)