Volume 24
Number 6
21 March 2022
Pages 2253-2640

# Green Chemistry

Cutting-edge research for a greener sustainable future

rsc.li/greenchem

ISSN 1463-9262

ROYAL SOCIETY OF CHEMISTRY

**PAPER**
Matthew J. Eckelman *et al.*
Applying green chemistry to raw material selection and product formulation at The Estée Lauder Companies

Green Chemistry

PAPER

View Article Online
View Journal | View Issue



Cite this: *Green Chem.*, 2022, **24**, 2397

# Applying green chemistry to raw material selection and product formulation at The Estée Lauder Companies†

Matthew J. Eckelman, *[a] Matthew S. Moroney,[a] Julie B. Zimmerman,[a] Paul T. Anastas,[a] Eva Thompson,[b] Paul Scott,[b] Maryann McKeever-Alfieri,[b] Paul F. Cavanaugh[b] and George Daher[b]

Advances in green chemistry over the past 25 years have improved sustainability in the development of new cosmetic and personal care products. Product formulators benefit from an expanding palette of "greener" natural and synthetic ingredients but need clear guidance on how to choose among options to optimize formula sustainability while also evaluating for performance. As greener can have a variety of meanings, for the purpose of this article, we define greener as being aligned with green chemistry principles. Here, we report the development of a quantitative green chemistry scoring methodology incorporating human health (HH), ecosystem health (ECO), and environmental (ENV) endpoints to specifically characterize cosmetic and personal care products. Using a hazard-based approach, a "Green Score" for cosmetic ingredients was calculated incorporating the HH, ECO, and ENV categories. Ingredient and chemical component data were obtained from manufacturers, open-source databases, or computer model estimates. There are 8 individual metrics: 3 each for HH (acute, ocular, and dermal toxicity) and ECO (bioaccumulation, persistence, and aquatic toxicity) and 2 for ENV (feedstock sourcing and greenhouse gas emissions). All metrics and data quality measures can be examined by formulators for individual endpoints, averaged by category, or further averaged to an overall Green Score. This scoring framework has been applied across ingredients and product formulations at The Estée Lauder Companies to establish Green Score baseline values, identify priority raw materials for replacement, and guide future innovation. Actual scores and statistical results are presented here at the ingredient, formula, and product subcategory levels to demonstrate the functionality of the tool as a measure of green chemistry performance.

Received 24th August 2021,
Accepted 7th December 2021

DOI: 10.1039/d1gc03081g

rsc.li/greenchem



## Introduction

Informed selection of ingredients and raw materials is a key process in the development of personal care and cosmetic products. Product formulators must integrate available environmental, ecosystem, and human health data to improve sustainability across their product portfolios. Application of advances in green chemistry and engineering, along with judicious choice of ingredients, promises to yield significant improvements in product sustainability across product portfolios in many industries.[1,2]

A challenge for any organization seeking to integrate sustainability into their product development-related decision-making is the need for a standard definition for product sustainability and a comprehensive framework of metrics with which to measure progress. The field of green chemistry fortunately offers a wealth of knowledge that can be incorporated into enabling and quantifying sustainability.[3] The significant progress in global green chemistry initiatives provides an expanding framework for balancing the large-scale considerations of sustainability with tangible actions and metrics that integrate those considerations into product design and business decisions across the product life cycle.[2,4,5] These incremental advances are important steps for reaching specific aspects of the United Nations Sustainable Development Goals.[6,7] This framework is particularly valuable for formulators within the beauty care (*i.e.*, cosmetic) industry because of the direct application of products onto the skin and hair. Tools based in green chemistry principles can assist in harmonizing sustainability goals with the formulation design process, providing a quantitative approach to make decisions and measure progress.

A variety of bespoke tools have emerged for green chemistry and life cycle assessments for pharmaceutical and personal

[a]*Sustainability A to Z, LLC, Guilford, CT, USA. E-mail: matt@sustainabilityatoz.com*
[b]*The Estée Lauder Companies, New York, NY, USA*

†Electronic supplementary information (ESI) available. See DOI: 10.1039/d1gc03081g

care products (PPCPs). Tools for green chemistry assessments of PPCPs can be categorized along 2 continuums: hazard (inherent nature of ingredients) and risk (concentration in final product) as well as cradle to cradle (life cycle) and gate to gate (internal operations). Tools within this space include those that consider a number of green chemistry principles,[8–11] and help inform the design of "greener" raw materials and/or products to yield more sustainable PPCPs. Each incorporates unique environmental and toxicologic metrics, weighting, and scoring algorithms.

As the data underlying these tools evolve to be more robust and transparent, there are opportunities to improve on green chemistry measurement frameworks to account for inherent hazards, consider life cycle implications, enable the design of greener raw materials and products, and most importantly, drive innovation to more sustainable means of providing the necessary function in PPCP formulations. To be trusted, effective tools should be based on the best available evidence and be transparent in their algorithms and data sources.

Here we present a new "Green Score" tool designed for rapid assessment of PPCP ingredients and formulations based on knowledge at the fundamental molecular level, coupled with life cycle sourcing and end-of-life considerations. The 12 principles of green chemistry provide a hazard-based perspective on ingredients. While The Estée Lauder Companies (ELC) uses a risk-based safety assessment framework, our Green Score tool incorporates several of the 12 principles of green chemistry as a complement:

- Principle 4: Designing safer chemicals
- Principle 6: Design for energy efficiency
- Principle 7: Use of renewable feedstocks
- Principle 11: Real-time analysis for pollution prevention

The Green Score tool evaluates green chemistry principles and chemical hazards in 3 distinct categories: human health (HH), ecosystem health (ECO), and environmental impact (ENV). The tool has been applied to chemical components of ingredients for formulations within ELC and provides useful insights into the effect of specific ingredient choices (singly or in combination) on a product formulation's overall Green Score.

Notably, the tool includes several important features: (1) a balance between assessing inherent chemical and supply chain hazards, (2) a disincentive to use raw materials with low scores or lack of data by weighting their impact to reduce the score further, and (3) a certainty score to provide insight on the level of confidence in the Green Score for a given ingredient or chemical component.

Beyond presenting the methodology used to develop the novel Green Score tool, this report also demonstrates how results can be interpreted and applied. At the product category level, statistical analyses can provide an overall baseline and Green Score comparisons across products within a category and by product form (*e.g.*, solid or liquid). Descriptive statistical analyses within ingredient or product groupings can also be used to identify low Green Score outliers through best Green Score performers. This can be integrated and displayed at the formulator bench level, using data dashboards allowing for rapid identification of alternative ingredients. Before and after scoring for specific reformulations allows for quantitative analyses of PPCP improvements using green chemistry–based optimization. In concert with continuing advances in green chemistry and engineering, the Green Score tool is being leveraged to prioritize innovation for individual ingredients and raw material classes.

## Methods

The Green Score is applied in a sequence of steps, each of which is detailed in the sections that follow. The scoring framework has a nested structure: individual chemicals are combined to make ingredients, and ingredients are combined to make formulas. In Step I, the chemical composition of each ingredient is established from internal registration records, each of the 2300 + unique components is linked to internal and external chemical data sets, and water components are removed from the scoring. In Step II, each ingredient is scored on metrics covering HH, ECO, and ENV categories. HH and ECO metrics are based on inherent chemical properties and carried out at the component level, while ENV metrics are largely applied at the ingredient level. Each of these metric scores has an associated data quality rating based on a tiered system of data source preferences. In Step III, numeric penalties (*i.e.*, disincentives) are applied to any component or ingredient that receives the lowest score (1) for any metric. In Step IV, all metric and category scoring is mass averaged up to the ingredient level and a final Green Score is calculated. Finally, in Step V, ingredient scores are mass averaged up to the formula level and evaluated against benchmarks.

### Step I: ingredient data and calibration

Chemical composition and sourcing details for each ingredient and its components are obtained from suppliers during the registration process. Each chemical component is then matched to its Chemical Abstracts Service (CAS) registry number and European Inventory of Existing Commercial Chemical Substances (EINECS) number, either directly from supplier data or through manual matching. Each ingredient also has a code internal to ELC. These numerical identifiers are used throughout the scoring process.

Ingredient compositions are then adjusted to remove water, with all ingredient proportions rescaled as shown in the following equation:

$$P'_{ij} = \frac{P_{ij}}{1 - w_j}$$

where $P'_{ij}$ = adjusted proportion of component $i$ in ingredient $j$, $P_{ij}$ = original proportion of component $i$ in ingredient $j$, and $w_j$ = proportion of water in ingredient $j$.

Ingredient sourcing details considered are feedstock source (plant, mineral, petroleum, *etc.*), country of origin, and existing sustainability certifications, such as certified organic and Roundtable on Sustainable Palm Oil (RSPO). In addition, a separate GHG emissions survey is sent out to all ELC suppliers

to collect data on Scopes 1 and 2 emissions (according to the GHG protocol) per kilogram of manufactured ingredient delivered to ELC.

### Step II: scoring individual ingredients

Each ingredient is evaluated using available data on HH, ECO, and ENV endpoints. Fig. 1 shows the metrics for each endpoint by category and the scoring rubric used for evaluation on a scale of 1 to 5. Endpoints were selected based on (1) the principles of green chemistry, (2) availability and completeness of data, (3) the desire to enable transparency and traceability in the ingredient scoring process, and (4) the specific needs of the cosmetic industry. The intention of the scoring framework is to use the most complete, conservative, and traceable data

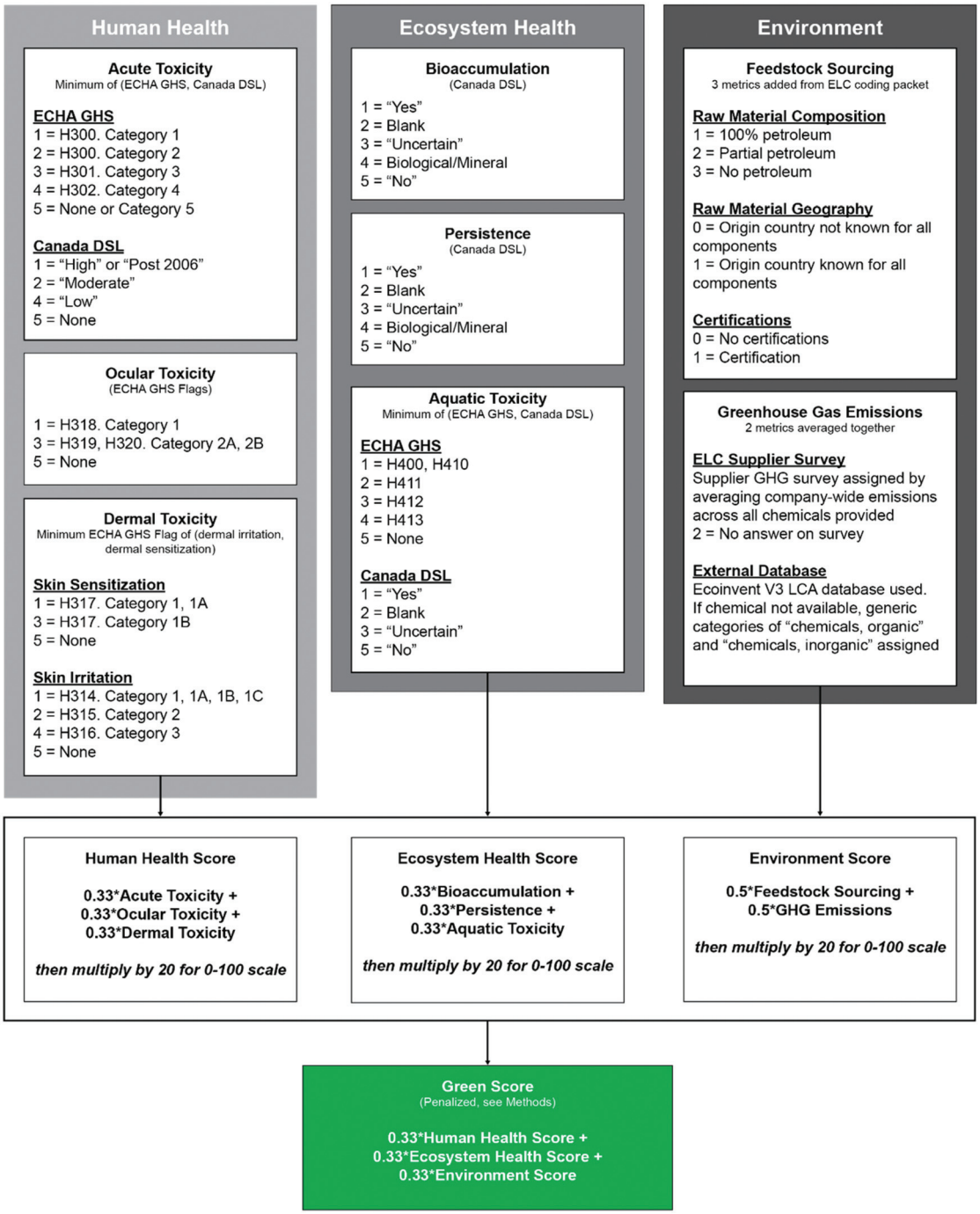



**Fig. 1** Green Score ingredient scoring rubric. Vertical columns represent human health, ecosystem health, and environmental categories, with each endpoint metric scored from 1 to 5. DSL, Domestic Substance List; ECHA, European Chemical Agency; ELC, The Estée Lauder Companies; GHG, greenhouse gas; GHS, Globally Harmonized System of Classification and Labelling of Chemicals.

available. Endpoints corresponding to green chemistry principles for which reliable data could not be comprehensively gathered across all ingredients or suppliers, such as hazardous chemicals used in upstream processes, are seen as important targets for future efforts.

**Human health.** Acute, ocular, and dermal toxicity endpoints are scored at the chemical component level based on the hazard codes from the Globally Harmonized System of Classification and Labelling of Chemicals (GHS), which are gathered from the European Chemical Agency (ECHA) Classification and Labelling (C&L) Inventory by matching CAS and EINECS numbers. Additional hazard classifications from the Canadian Domestic Substance List (DSL) for HH priorities are used in the absence of ECHA data or if DSL hazard data are more conservative. Chemical components listed in the ECHA C&L Inventory and/or Canadian DSL that do not have any associated hazard codes are assumed to be benign. Table S1† shows the detailed scoring rubric for each of the 3 HH endpoints.

**Ecosystem health.** For bioaccumulation, persistence, and aquatic toxicity endpoints, a score is assigned to each ingredient component based on feedstock source data provided by the Canadian DSL labelling and the supplier. In the case of aquatic toxicity data, hazard classifications according to the UN Globally Harmonized System of Classification and Labelling of Chemicals (GHS) were also reviewed, and the worst case of these and the Canadian DSL labelling was applied. Considering bioaccumulation and persistence, natural components that are not present in the DSL categorization and are wholly biological or mineral are treated separately than those from other chemical feedstock sources such as petroleum-based chemicals. For example, for bioaccumulation, if a component is not present in the DSL and its chemical feedstock source is wholly biological, then a score of 4 is assigned. Table S2† shows the detailed scoring rubric for each of the 3 ECO endpoints.

**Environment.** Feedstock source and GHG emissions endpoints make up the ENV category of the Green Score. Feedstock source, as an endpoint, is a composite metric that considers the physical source of the raw material, traceability, and any existing sustainable sourcing environmental certifications. An ingredient is given a score of 1 if it is 100% petroleum derived, 2 if partially of petroleum origin, and 3 if wholly biological or mineral. This rubric is designed to disincentivize the use of petrochemical ingredients in ELC formulations. One additional point is added if all components of an ingredient can be traced to a specific country of origin and/or if the ingredient has certified organic or RSPO certifications, for a maximum score of 5. Upstream organizational health & safety and labor practices throughout the value chain are central to the organization's supply chain practices but are managed through mechanisms and standard operating procedures outside of the Green Score tool.

GHG emissions are also scored as a composite metric averaged from 2 distinct data sources. Emissions performance of suppliers is represented by their reported Scopes 1 and 2 emissions per kilogram of ingredient produced, collected through supplier surveys. These data represent raw material manufacturing operations but do not account for upstream Scope 3 emissions of chemical ingredients. Emissions performance of the upstream supply chain is represented by modelling each of the 2300+ chemical ingredients using embodied cradle-to-gate GHG emissions data gathered from the ecoinvent life cycle inventory database, maintained by the Swiss Centre for Life Cycle Inventories (http://www.ecoinvent.org). GHG emissions from transporting raw materials to ELC manufacturing sites are not currently included, as production locations shift depending on demand and capacity constraints. These ingredient-based GHG emissions are then mass averaged up to the raw material level. Both GHG emissions are first recorded in absolute units of kilogram of carbon dioxide equivalents (kg $CO_2$ eq) per kilogram of ingredient, and then rescaled to the 1 to 5 scale used for other Green Score endpoints. Because GHG emissions for chemicals can span several orders of magnitude, a logarithmic scale is applied for rescaling. Ingredients with scores that fall above a threshold of 1000 kg $CO_2$ eq per kilogram are assigned the worst value of 1, as are ingredients from the supplier survey that are reported as having zero GHG emissions, likely indicating that the survey was not completed correctly. A conservative score of 2 is assigned if no emissions data were provided by the supplier.

Table S3† outlines the full scoring assignment rubric for the 2 ENV endpoints.

**Default scoring.** For HH and ECO endpoint metrics where data are incomplete or not available, a system of classification is applied and default values assigned by chemical component class so that scoring can proceed up to the ingredient level. Table S4† details the different chemical component classes and 1 to 5 scoring defaults that are used in these cases when no data are available. The default values are based on generally accepted environmental and toxicologic rules of thumb, for example, that high-molecular-weight polymers will not bioaccumulate,[13] or that fluorinated polymers will persist in the environment, or that components derived from edible plants are likely to be nontoxic. For a small number (<1%) of components, expert judgement is also used to establish read-across proxies where ingredients are closely related (*e.g.*, PEG-40 and PEG-42). If, after chemical class default and read-across scoring, there are still endpoints with missing data, these are assigned a conservative value of 2.

**Data availability and certainty.** Accompanying each of the HH, ECO, and ENV metrics is a certainty score, which indicates the robustness of each metric's data source and reflects the complexity and evolution of the science and scientific models underlying the scoring. The certainty scores are applied at the same level (*e.g.*, chemical component) as their corresponding metrics and mass averaged up to the ingredient level using the same 1 to 5 scale. The certainty scores are intended to give the user insight into the quality and robustness of the science underlying the data, and as such, help to inform their decision-making process. Certainty scores can

also be used to prioritize improvement of the Green Score overall by indicating where data quality needs to be improved. The rubric for assigning certainty scores is presented in Table S5.†

**Step III: penalty score analysis**

To further discourage the use of ingredients with poor performance, a numeric penalty system is used to reduce the score. A deduction of 0.1 is applied for every individual metric that receives a score of 1 (lowest possible initial score). As with the baseline scoring, for HH and ECO metrics, these penalties are assessed at the component level, while for ENV metrics, they are assessed at the ingredient level. To evaluate a variety of penalty scoring schemes, simulations were performed on a sample set of 14 formulations to evaluate how the spread (*i.e.*, standard deviation) was affected under various penalty scenarios. The 2 key variables evaluated to affect the spread were criteria, the value to begin penalization (*e.g.*, score of 1 or 2), and penalty, how much the score should be reduced for each criteria violation (*e.g.*, deduction of 0.1 or 0.2). The penalty weights can be adjusted but the initial value of 0.1 was determined through the simulations. Statistical analysis with and without the penalty demonstrated that the penalty system aids in providing differentiation among ingredient choices. The application of penalties shifts the scoring from a scale of 1 to 5 to a scale of 0 to 5. If the penalties, when applied, would bring an endpoint score to a negative number, a score of zero is assigned.

As the criteria to trigger a penalty increases, there are an increasing number of raw materials that will be penalized, particularly for the ENV endpoint, where average raw material scores tend to be higher. The effect of the penalty scheme on the overall Green Score is nonlinear as the criteria and penalty values change since the 14 formulations are composed of different raw materials at different percentages, each with their own attributes that may trigger different penalties. The selected criteria value of 1 avoids dramatic jumps associated with certain categories (*e.g.*, 500% increase in penalties for multiple formulations in the ENV category as the criteria value increases to 2).

Penalty calculations were evaluated using 14 different formulations (7 moisturizers and 7 foundations) to simulate spread using a variety of thresholds (*e.g.*, score of 1 or 2) and penalties (*e.g.*, deduction of 0.1 or 0.2). The standard deviation for each simulation (*i.e.*, penalized Green Score vector) was calculated by applying a matrix of penalties (0.0–0.25 in 0.05 increments) and thresholds (0–2 in 0.25 increments). This process was repeated separately for moisturizers and foundations. The penalty scheme of criteria ≤1 and penalty value of 0.10 per exceedance results in an increased standard deviation of ~100% for foundations (0.24–0.5) and ~200% for moisturizers (0.24–0.73) due to a higher number of low-scoring components in the moisturizer ingredients. For the ENV category, the scores were similar between these product categories, with foundations having a much smaller interquartile range.

**Step IV: calculation of ingredient Green Score**

The final ingredient Green Score is calculated as follows: first, metrics assessed at the chemical component level are mass averaged to the ingredient level (adjusted for removal of water from the ingredient):

$$I_{jk} = \sum_{i=1}^{n} C_{ijk} \cdot P'_{ij}$$

where $I_{jk}$ = ingredient-level score for ingredient $j$ on metric $k$, $C_{ijk}$ = component-level score for component $i$ in ingredient $j$ on metric $k$, $P'_{ij}$ = adjusted proportion of component $i$ in ingredient $j$, and $n$ = number of components in ingredient $j$.

Then, HH, ECO, and ENV category scores are calculated through simple averaging of the ingredient-level metrics $k$ in each category. Finally, the overall ingredient Green Score $I_j$ for each ingredient $j$ is obtained by simple averaging of the category scores. For ease of interpretation by product developers, the overall Green Score is rescaled from a scale of 0 to 5 to a scale of 0 to 100 (100 being best).

**Step V: establishing formula Green Score and benchmarks**

Formula Green Scores are calculated by the mass-averaged ingredient Green Scores based on ELC formula composition. As at the ingredient level, water in the formula is excluded from calculations:

$$F_I = \sum_{j=1}^{r} \frac{I_{jI} \cdot P_{jI}}{100 \cdot \left(1 - \left(\frac{w_I}{100}\right)\right)}$$

where $F_I$ = formula Green Score for formula $I$, $I_{jI}$ = ingredient Green Score for ingredient $j$ in formula $I$, $P_{jI}$ = percentage of ingredient $j$ in formula $I$, $w_I$ = percentage of water in formula $I$, and $r$ = number of ingredients in formula $I$.

Product category Green Score benchmarks are calculated by grouping all like active formulas in each product category (*e.g.*, haircare, skincare, make-up) and subcategory (*e.g.*, serums, waterproof mascara, conditioners, solid perfumes). The initial benchmark is set as the mean formula Green Score for that category/subcategory.

**Formula-level statistical analyses.** ELC provided a scored formula-level data set ($n$ = 11 030) to evaluate the performance tool at the formulation level, which included grouping variables such as product category ($n$ = 25; *e.g.*, lip products, facial cosmetics) and product form ($n$ = 28; *e.g.*, gel, solid, stick). Prior to statistical analysis, certain data were removed to reduce the number of grouping variables and improve the identification of trends in the data set. Specifically, product categories and product forms with <100 observations were removed. All formulations before 2005 were also removed. After cleaning, the data set was reduced by 8%, with >30 factors removed (from $n$ = 11 030 to $n$ = 10 138).

A variety of exploratory data analyses were used prior to statistical analysis in *R: A language and environment for statistical computing*, including visual tests (*e.g.*, boxplots, histograms) and statistical tests. Assumptions of normality were evaluated

using the Anderson-Darling test in *R*.[14,15] A variety of transformations were applied to the Green Score vector, then reevaluated for normality, including natural log, square root, logarithm base 10, inverse, sqrt[max($x$ + 1) − $x$], $\log_{10}$[max($x$ + 1) − $x$], and 1/[max($x$ + 1) − $x$]. A selected statistical significance value of $\alpha$ = 0.05 was used for evaluation. All tested transformations failed normality tests at extremely significant $P$ values (*e.g.*, $>10 \times 10^{-16}$).

The failure of normality assumptions indicates that parametric tests, such as the $t$ test, $z$ score, and analysis of variance (ANOVA), should not be used to analyse the Green Score vector between grouping factors. Instead, to evaluate multiple groups, the Kruskal–Wallis test, a nonparametric equivalent of ANOVA, was used. If a significant difference between groups was found, the Dunn test of multiple pairwise comparisons, a nonparametric equivalent of the Tukey honestly significant difference test, was used.[16,17]

# Results

### Descriptive statistics of ingredient Green Scores

To explore the use and effectiveness of the Green Score tool, statistical boxplots of results for all ingredients by endpoint and for overall Green Score were developed (Fig. 2). First, evaluation of an ingredient portfolio ($n$ = 4345) showed a mean Green Score of 72.95 with a standard deviation of 9.64 (Fig. 2a). These Green Score values are an aggregation of the 8 individual endpoints (3 HH, 3 ECO, 2 ENV), with each contributing to the overall Green Score distributions (Fig. 2a). The ECO endpoint of bioaccumulation has the highest median, while the ENV endpoint of sourcing has the lowest. It is also interesting to note that the ENV endpoints (sourcing and GHG emissions) and ECO endpoints (acute toxicity and bioaccumulation) have tighter interquartile ranges than several of the other endpoints, including ocular toxicity (HH), dermal toxicity (HH), persistence (ECO), and aquatic toxicity (ECO), largely due to differences in data availability. The distribution of Green Score values across raw materials is skewed toward higher scores and the HH scores are the highest. (Fig. 2b).

### Comparing Green Scores by functional class, product category, and product form

Further richness in the Green Score tool can be found by exploring the resulting scores in ECO, ENV, and HH categories for ingredients by functional class (Fig. S2a†), and for formulations by product category (Fig. S2b†) and product form (Fig. S2c†). It can be observed that certain categories have lower scores, indicative of opportunities for raw material substitution, product reformulation, or preferred product forms. In the functional class analysis (Fig. S2a†), certain functional classes score higher on ECO endpoints (*e.g.*, essential oils, emulsifiers) than others (*e.g.*, chelating agents, colorants). Similarly, certain functional classes perform better on ENV endpoints (*e.g.*, antioxidants, chelating agents) than others (*e.g.*, lubricants, suspending agents, solvents). In terms of the overall distribution of scores, the range is largest for colorants and conditioners, likely driven by low outliers not observed in other functional classes.

When considering Green Scores by product category (Fig. S2b†), differences in performance are observed. For example, lip care products score relatively well across endpoints, whereas haircare products tend to score well for the ENV endpoint. It is interesting to note that formulations in the haircare category have a lower mean HH score than all the other product categories. This can largely be attributed to the presence of solvents and colorants (lower-scoring ingredients) in hair dye, which affect the overall mean.

When considering product form (Fig. S2c†), it is unequivocally clear that liquids and emulsions score lower across all 3 endpoints, while sticks, solids, powders, and anhydrous products tend to score higher. Again, this can be attributed to the nature of the ingredients required for the various product forms and the need to use lower-scoring functional classes (*e.g.*, solvents, suspending agents) to formulate liquid and emulsion product forms.

Investigating combinations of attributes is also useful. A robust approach to greener formulation would be the simultaneous consideration of a product category, such as lip care, and product form (*e.g.*, gel, liquid, or stick; Fig. S3† and Table 1). It is interesting to note that within the lip care product category, the mean Green Score is significantly higher for sticks than for liquids and gels, suggesting that a focus on greener ingredient innovations meeting the unique technical needs of certain physical forms could be beneficial for the development of greener product lines.

### Leveraging Green Scores for product reformulation

As a critical tool to guide product reformulation, the Green Score can be leveraged to identify low-scoring ingredients and proactively design formulations that score higher. Fig. 3 shows an example dashboard for formulators to choose among available ELC ingredients that provide a specific function, in this case, waxes. Ingredients are ranked by total Green Score and the dashboard visually shows which hazard endpoints are most problematic for each ingredient option.

Examining the ingredient options in Fig. 3, solid natural beeswax (row 1) received the highest score, while the wax version of the same product (row 4) received a lower score. This difference stems entirely from the ENV category, where the solid version is certified organic from a supplier with relatively low reported facility GHG emissions (and thus a higher ENV GHG score). In contrast, the wax version is not certified organic and is from a supplier with higher reported GHG emissions. This heterogeneity in classifications and Green Score values for related chemical compounds underscores the importance of using substance-specific hazard data. Exploring further how specific hazard data can influence the overall green score, synthetic beeswax (row 6) is listed by the DSL as an aquatic toxicity concern and therefore has one of the lowest scores, so is penalized in the ECO category. The other petroleum waxes, such as the microcrystalline form (rows 16–18),

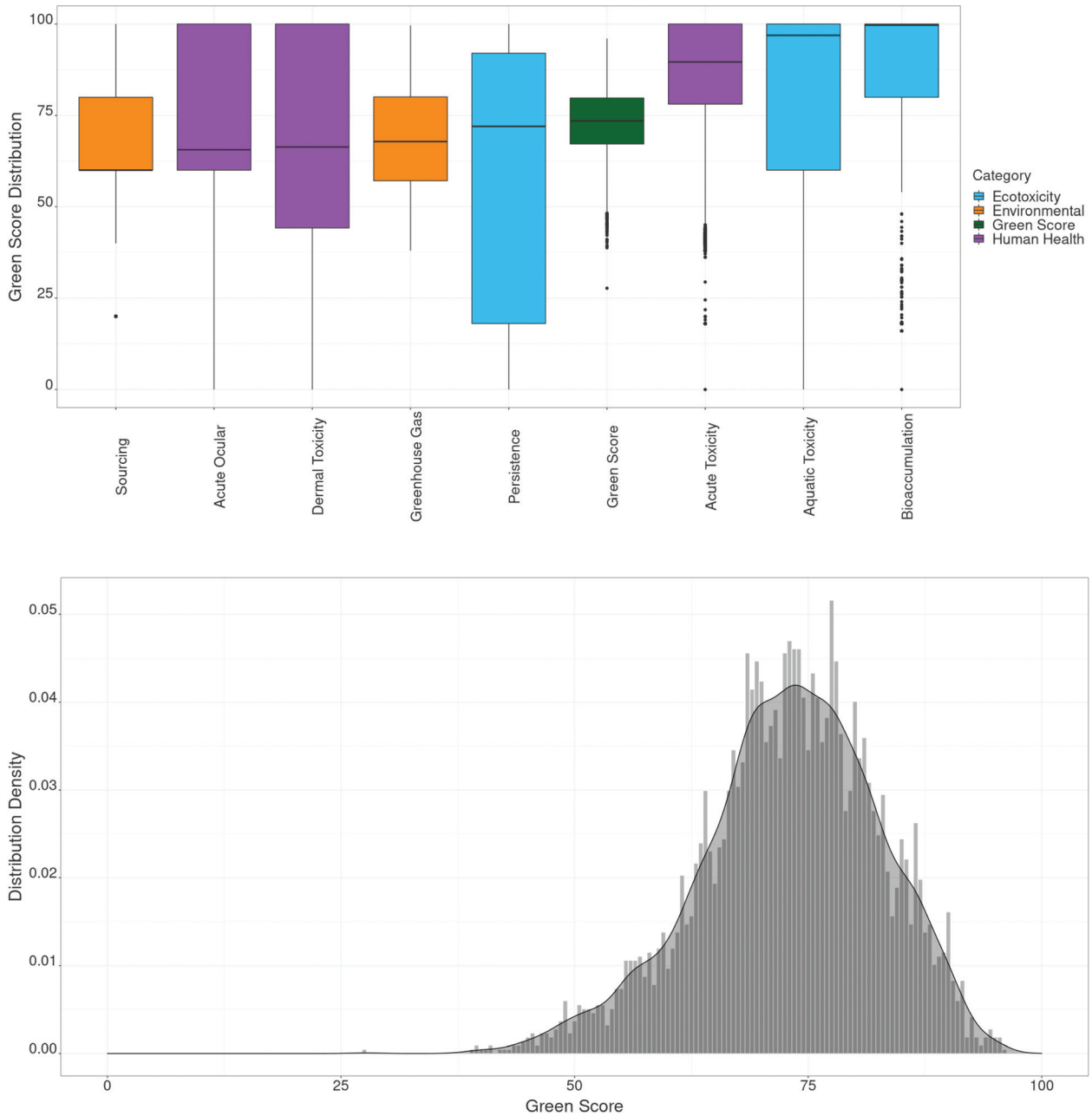



**Fig. 2** (A) Green Score distributions for individual metrics ranked by median score over the data set ($n$ = 4345). (B) Overall Green Score distribution histogram with density curve overlay across the data set ($n$ = 4345).

**Table 1** Green score comparisons for various forms of lip products

| Category | Product form | No. of products considered | Minimum score | Average score | Maximum score |
|---|---|---|---|---|---|
| Lip product | Gel | 161 | 63.9 | 71.9 | 80.5 |
| Lip product | Liquid | 63 | 65.7 | 72.8 | 79.3 |
| Lip product | Stick | 1356 | 55.3 | 76.3 | 85.1 |

are faced with the challenge of a persistent classification also resulting in a lower score. Rose floral wax (row 19) contains essential oils with GHS flags that further drive a lower score relative to most other waxes.

Evaluating the data set, 10 substances with the lowest Green Scores were identified (Table 2). As a proof point for the output of the tool, it is noteworthy that of these raw materials, many are silicone and silicone-like compounds that are cur-

| | | | HUMAN HEALTH HAZARDS | | | ECOSYSTEM HEALTH HAZARDS | | | ENVIRONMENTAL HAZARDS | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| INCI NAME | INGREDIENT FORM | SOURCE | HH ACUTE | HH OCULAR | HH DERMAL | ECO BIOACC | ECO PERS | ECO AQTOX | ENV SOURCE | ENV GHG | GREEN SCORE |
| 1. BEESWAX | Solid | Animal related | 100 | 100 | 100 | 80 | 80 | 100 | 100 | 82 | 93 |
| 2. CITRUS AURANTIUM DULCIS (ORANGE) PEEL WAX | Wax | Plant | 100 | 100 | 100 | 80 | 80 | 100 | 80 | 74 | 88 |
| 3. POLYGLYCERYL-3 BEESWAX | Granular | Animal related | 100 | 100 | 100 | 80 | 80 | 100 | 80 | 72 | 88 |
| 4. BEESWAX | Wax | Animal related | 100 | 100 | 100 | 80 | 80 | 100 | 80 | 69 | 87 |
| 5. JOJOBA ESTERS | Powder | Plant | 100 | 100 | 100 | 60 | 60 | 100 | 80 | 83 | 85 |
| 6. SYNTHETIC BEESWAX | Solid | Plant | 100 | 100 | 100 | 100 | 100 | 18 | 80 | 79 | 84 |
| 7. SYNTHETIC WAX | Liquid | Petroleum | 80 | 100 | 100 | 100 | 100 | 100 | 40 | 78 | 83 |
| 8. PARAFFIN | Solid | Petroleum | 80 | 100 | 100 | 100 | 98 | 100 | 40 | 74 | 83 |
| 9. SYNTHETIC BEESWAX | Solid | Plant-Petroleum | 100 | 100 | 100 | 100 | 100 | 18 | 80 | 57 | 80 |
| 10. PARAFFIN | Wax | Petroleum | 80 | 100 | 100 | 100 | 100 | 98 | 40 | 57 | 80 |
| 11. LAVANDULA ANGUSTIFOLIA (LAVENDER) FLOWER WAX | Wax | Plant | 100 | 60 | 60 | 80 | 80 | 100 | 60 | 88 | 78 |
| 12. CARNAUBA | Flakes | Plant | 100 | 60 | 100 | 60 | 60 | 100 | 80 | 65 | 78 |
| 13. ORYZA SATIVA (RICE) BRAN WAX | Solid | Plant | 100 | 60 | 60 | 60 | 60 | 100 | 80 | 88 | 77 |
| 14. OZOKERITE | Solid | Petroleum | 100 | 100 | 100 | 60 | 60 | 100 | 40 | 62 | 74 |
| 15. EUPHORBIA CERIFERA (CANDELILLA) WAX | Granular | Plant | 80 | 60 | 60 | 60 | 60 | 100 | 80 | 65 | 71 |
| 16. POLYETHYLENE/MICROCRYSTALLINE WAX | Solid | Petroleum | 70 | 90 | 100 | 100 | 16 | 100 | 40 | 57 | 68 |
| 17. MICROCRYSTALLINE WAX | Wax | Petroleum | 40 | 60 | 100 | 100 | 18 | 98 | 60 | 68 | 67 |
| 18. MICROCRYSTALLINE WAX | Pellets | Petroleum | 40 | 60 | 100 | 100 | 18 | 98 | 40 | 73 | 64 |
| 19. ROSA CENTIFOLIA/DAMASCENA FLOWER WAX | Wax | Plant | 90 | 16 | 40 | 70 | 70 | 28 | 80 | 65 | 59 |
| 20. CERESIN | Wax | Petroleum | 40 | 40 | 40 | 60 | 60 | 60 | 40 | 55 | 49 |

**Fig. 3** Green Score formulator dashboard for waxes by endpoint and ranked by overall score, with a target threshold of 70. AQTOX, aquatic toxicity; BIOACC, bioaccumulation; ECO, ecosystem health; ENV, environment; GHG, greenhouse gas; HH, human health; INCI, International Nomenclature Cosmetic Ingredient; PERS, persistence.

**Table 2** Ten lowest Green Score values among ingredients

| | | | | Element of Green Score | | |
|---|---|---|---|---|---|---|
| | Ingredient | Common function | Green Score | HH | ECO | ENV |
| 1 | Cyclopentasiloxane | Emollient | 38.8 | 53.3 | 18.0 | 45.1 |
| 2 | Cyclopentasiloxane | Emollient | 39.3 | 56.4 | 20.8 | 40.6 |
| 3 | Alcohol denatured | Solvent | 39.4 | 25.3 | 45.3 | 47.5 |
| 4 | Cyclopentasiloxane | Emollient | 40.2 | 55.7 | 20.0 | 45.0 |
| 5 | Red 17 (CI 26100) | Colorant | 40.7 | 52.7 | 18.0 | 51.4 |
| 6 | Cyclopentasiloxane | Emollient | 41.0 | 60.3 | 20.7 | 42.0 |
| 7 | Phenyl trimethicone | Skin conditioning | 41.1 | 46.7 | 38.7 | 38.1 |
| 8 | Cyclopentasiloxane | Emollient | 42.2 | 57.0 | 21.2 | 48.3 |
| 9 | Zinc oxide | Skin protectant | 42.7 | 57.1 | 26.4 | 43.2 |
| 10 | Red 28 Lake (CI 45410) | Colorant | 43.2 | 66.4 | 16.8 | 46.4 |

ECO, ecosystem health; ENV, environment; HH, human health.

rently restricted for use in certain products in the European Union and Canada, while 2 are colorants that have regulatory restrictions on use in Canada. In this way, the tool can be used to elevate and prioritize raw materials based on the scoring of endpoints related to ECO, HH, and/or ENV impact.

In order to demonstrate the effectiveness of the Green Score tool in guiding future formulations, a case study for 3 products containing decamethylcyclopentasiloxane (D5) is presented in Table 3. This organosilicon compound has recently garnered attention because of its potential to persist and bioaccumulate

Table 3 Green score changes for formulas with D5 removed

| Product type | Overview of key changes | Green Score (before) | Green Score (after) | Absolute change | Percent change |
|---|---|---|---|---|---|
| Make-up remover | Various silicones (including D5 at 17%) replaced with combination of petroleum and plant-based emollients | 68.7 | 70.4 | 1.6 | +2.3% |
| Moisturizer | Several ingredient changes made. D5 (at 5%) replaced with dimethicone (5% total) | 70.1 | 71.9 | 1.8 | +2.6% |
| Liquid foundation | D5 (at 38%) replaced with alternative silicones | 57.8 | 67.0 | 9.2 | +15.9% |

D5, decamethylcyclopentasiloxane.

in the environment.[18–21] Given these concerns, many of the proposed and enacted restrictions on D5 are related to their use in wash-off product formulations. In being proactive, ELC began to reformulate wash-off products to eliminate the use of D5 and replace the functionality it provided with greener alternatives. Three products reformulated to eliminate D5, including a make-up remover, moisturizer, and liquid foundation, all yielded higher Green Scores, with score improvements ranging from 1.6 to 9.2 in absolute terms representing between 2.3% and 15.9% (Table 3). These results highlight how the Green Score tool can identify emerging chemicals of concern and guide substitution with greener alternatives.

### Leveraging Green Scores to incentivize raw materials from non-petroleum feedstocks

Principle 7 of the green chemistry principles – as reflected in the ENV sourcing score – states that "a raw material or feedstock should be renewable rather than depleting whenever technically and economically practicable."[12] Here, we evaluated whether the Green Score tool yielded statistically significant results for non-petroleum raw materials *versus* other sources (*e.g.*, petroleum and mixed petroleum/non-petroleum) with an initial focus on functional class (Fig. S1†) followed by a Kruskal–Wallis test (Table 4). The Kruskal–Wallis test between the three types of feedstock sources rejected the null hypothesis (Kruskal–Wallis chi-squared = 464.37; df = 2; $P$ value $<2.2 \times 10^{-16}$), indicating that the scores for the non-petroleum raw materials were significantly different (and higher) than those for the petroleum and mixed petroleum/non-petroleum raw materials. While the subsequent results of the Dunn test indicated no significant difference between mixed and petroleum raw materials at the Green Score level, the tool does reward the use of non-petroleum sourced raw materials.

### Role of uncertainty in the Green Score tool

To explore the relationship between data quality (*i.e.*, certainty), the ECO, ENV, HH category endpoints and the overall Green Score values, statistical analyses were performed by functional class of raw materials (Fig. S4a†), product category (Fig. S4b†), and product form (Fig. S4c†). As expected, some functional classes had lower certainty scores (*e.g.*, preservatives, chelating agents) compared with well-studied commodity ingredients, such as solvents (Fig. S4a†). Notably, certainty scores were highest for eyeshadow product types (Fig. S4b†) and powder product forms (Fig. S4c†), which may be attributable to a robust data set demonstrating low bioavailability of solids to the skin.[22]

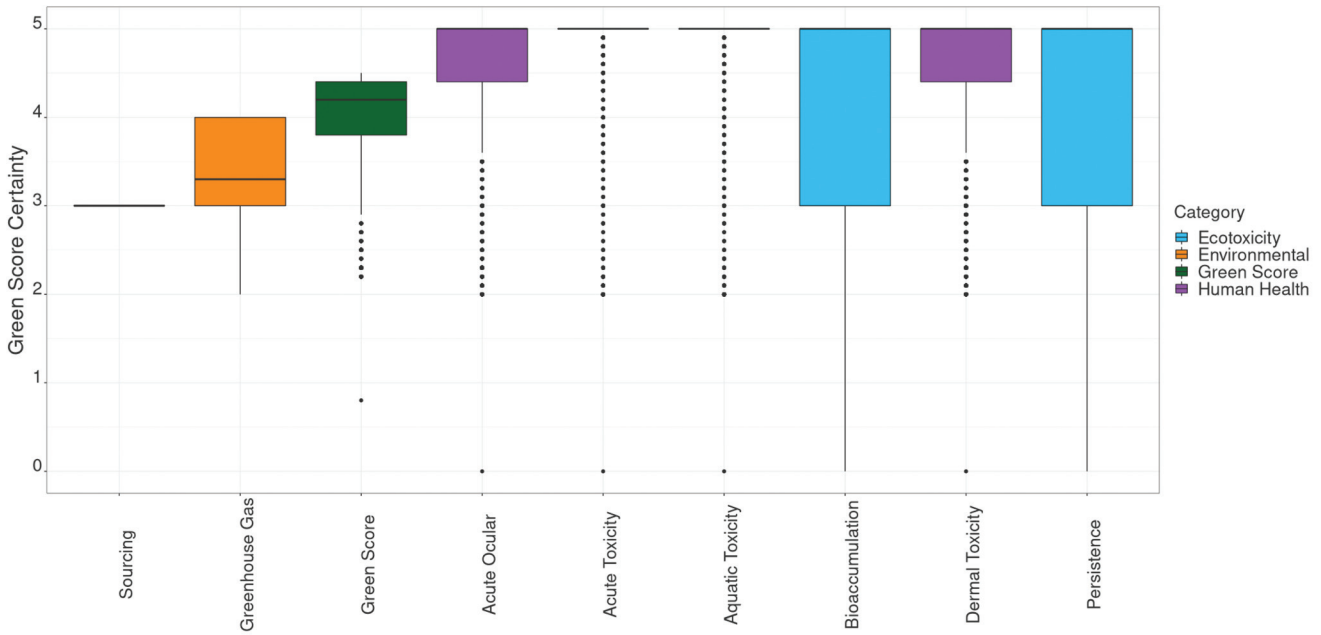
Comparing across hazard categories (boxplots shown in Fig. 4), ENV endpoints demonstrated the lowest median certainty scores and among the smallest range of uncertainty compared with the ECO and HH endpoints. This is to be expected based on the scoring methodology for ECO and HH, which depends on empirically studied endpoints of inherent hazard (providing higher-quality data) from multiple data sets (providing a larger data quality range), *versus* the scoring methodology for ENV, which depends on self-reported raw data from suppliers and modelled results. This provides a clear indication that improved confidence in the overall Green Score could be readily achieved through a more robust process for supplier-provided data, including training for small- and medium-sized enterprises that may lack the staff to perform the necessary calculations. The presence of outliers in the boxplots for HH and ECO endpoints signals that the data quality is not universally high for these categories and should ideally be improved for some chemicals as more robust data become available.

### Implications and limitations

The Green Score tool was designed to assess ingredient portfolios for ELC and unlock critical insights needed to guide future innovation. As a hazard-based scoring system, the tool

Table 4 Statistical comparisons between different compositions

| Comparison | Dunn test $z$ score | Adjusted $P$ value | Significant |
|---|---|---|---|
| Mixed *vs.* non-petroleum | −17.16 | $2.66\times10^{-66}$ | * |
| Non-petroleum *vs.* Petroleum | 16.16 | $4.98\times10^{-59}$ | * |
| Mixed *vs.* petroleum | 1.412 | 0.079 | |




| | | | HUMAN HEALTH HAZARDS: DATA CERTAINTY | | | ECOSYSTEM HEALTH HAZARDS: DATA CERTAINTY | | | ENVIRONMENTAL HAZARDS: DATA CERTAINTY | |
|---|---|---|---|---|---|---|---|---|---|---|
| INCI NAME | INGREDIENT FORM | SOURCE | HH ACUTE | HH OCULAR | HH DERMAL | ECO BIOACC | ECO PERS | ECO AQTOX | ENV SOURCE | ENV GHG |
| 1. BEESWAX | Solid | Animal related | 5 | 5 | 5 | 2 | 2 | 5 | 3 | 4 |
| 2. CITRUS AURANTIUM DULCIS (ORANGE) PEEL WAX | Wax | Plant | 5 | 5 | 5 | 2 | 2 | 5 | 3 | 4 |
| 3. POLYGLYCERYL-3 BEESWAX | Granular | Animal related | 5 | 5 | 5 | 2 | 2 | 5 | 3 | 4 |
| 4. BEESWAX | Wax | Animal related | 5 | 5 | 5 | 2 | 2 | 5 | 3 | 4 |
| 5. JOJOBA ESTERS | Powder | Plant | 5 | 5 | 5 | 5 | 5 | 5 | 3 | 4 |
| 6. SYNTHETIC BEESWAX | Solid | Plant | 5 | 3 | 3 | 5 | 5 | 5 | 3 | 3 |
| 7. SYNTHETIC WAX | Liquid | Petroleum | 5 | 5 | 5 | 5 | 5 | 5 | 3 | 3 |
| 8. PARAFFIN | Solid | Petroleum | 5 | 5 | 5 | 5 | 5 | 5 | 3 | 3 |
| 9. SYNTHETIC BEESWAX | Solid | Plant-Petroleum | 5 | 3 | 3 | 5 | 5 | 5 | 3 | 2 |
| 10. PARAFFIN | Wax | Petroleum | 5 | 5 | 5 | 5 | 5 | 5 | 3 | 2 |
| 11. LAVANDULA ANGUSTIFOLIA (LAVENDER) FLOWER WAX | Wax | Plant | 2 | 2 | 2 | 2 | 2 | 2 | 3 | 4 |
| 12. CARNAUBA | Flakes | Plant | 5 | 5 | 5 | 5 | 5 | 5 | 3 | 3 |
| 13. ORYZA SATIVA (RICE) BRAN WAX | Solid | Plant | 5 | 2 | 2 | 5 | 5 | 5 | 3 | 4 |
| 14. OZOKERITE | Solid | Petroleum | 5 | 5 | 5 | 2 | 2 | 5 | 3 | 3 |
| 15. EUPHORBIA CERIFERA (CANDELILLA) WAX | Granular | Plant | 5 | 2 | 2 | 5 | 5 | 5 | 3 | 3 |
| 16. POLYETHYLENE/MICROCRYSTALLINE WAX | Solid | Petroleum | 5 | 5 | 5 | 5 | 5 | 5 | 3 | 3 |
| 17. MICROCRYSTALLINE WAX | Wax | Petroleum | 5 | 5 | 5 | 5 | 5 | 5 | 3 | 3 |
| 18. MICROCRYSTALLINE WAX | Pellets | Petroleum | 5 | 5 | 5 | 5 | 5 | 5 | 3 | 3 |
| 19. ROSA CENTIFOLIA/DAMASCENA FLOWER WAX | Wax | Plant | 5 | 5 | 5 | 4 | 4 | 5 | 3 | 3 |
| 20. CERESIN | Wax | Petroleum | 5 | 2 | 2 | 5 | 5 | 5 | 3 | 3 |

**Fig. 4** (A) Green Score certainty distributions for individual metrics ranked by median score over the data set ($n$ = 4345). (B) Certainty scores by individual metrics for waxes. Areas for data certainty improvement are highlighted in orange. AQTOX, aquatic toxicity; BIOACC, bioaccumulation; ECO, ecosystem health; ENV, environment; GHG, greenhouse gas; HH, human health; INCI, International Nomenclature Cosmetic Ingredient; PERS, persistence.

complements the company's existing risk-based safety program to ensure a conservative view on ingredient selection and formula creation. By utilizing the tool throughout the product development cycle (*e.g.*, from ingredient functional class to product category and product form) the framework provides formulators with the technical agility required to shape formulation decisions and embeds green chemistry into decision-making. As data are captured using this metric, knowledge and best practices can be leveraged and codified to inform green chemistry formulation across the cosmetics community for various product types.

The credibility behind the continuous improvement of formulas – as driven by the Green Score tool – can be assured, as existing and new raw materials are rigorously assessed on an ongoing basis. In addition, the approach presented here clearly indicates that certain product forms score higher than others across a variety of product categories, enabling formulators to readily focus on key innovation opportunities. With every improvement made by the Green Score, the tool itself will also be updated to further incentivize substitution by modifying default scores as well as the criteria set for penalties. By taking such a dynamic approach in evolving the tool, we can ensure that feedback loops are in place to improve scores across the entire product portfolio while staying ahead of and pre-empting reactive reformulation triggered by regulatory action.

The current framework strives for data transparency and verifiability, and so does not include all possible HH, ECO, or ENV endpoints of concern where only limited data are currently available. For the same reasons, not all of the 12 principles of green chemistry are currently accounted for in the tool. However, with improvements in testing and modelling methods, data availability, and broader regulatory review, additional HH or ECO endpoints such as endocrine disruption could be added. For supplier data, future development and standardization of supply chain reporting and frameworks may allow for inclusion of additional ENV data, such as manufacturing waste generation and use of hazardous process chemicals.

The current approach advances the organization's sustainability goals in a way that can be transparently measured, tracked, and validated. With this data-driven approach comes the opportunity to proactively guide the supply chain and strengthen green-chemistry-inspired formulation above and beyond regulations. While the Green Score will be continuously improved to incorporate new data from regulators and suppliers, the current version is a transparent and robust tool to inform formulator decision-making, communicate expectations with suppliers, and prioritize raw materials, product types, and product forms for reformulation. ELC will use the Green Score across its operations to guide future innovation for greener alternatives.

## Conflicts of interest

Eva Thompson, Paul Scott, Maryann McKeever-Alfieri, Paul Cavanaugh, and George Daher are employees of and may hold stock in The Estée Lauder Companies.

Paul Anastas, Matthew Eckelman, Matthew Moroney, and Julie Zimmerman have served as consultants to The Estée Lauder Companies.

## Acknowledgements

Editorial assistance was provided with support under the direction of the authors by Samantha Agron, MD, and Erin Reineck, ELS, of MedThink SciCom.

## References

1 P. G. Jessop, F. Ahmadpour, M. A. Buczynski, T. J. Burns, N. B. Green II, R. Korwin, D. Long, S. K. Massad, J. B. Manley, N. Omidbakhsh, R. Pearl, S. Pereira, R. A. Predale, P. G. Sliva, H. VanderBilt, S. Weller and M. H. Wolf, *Green Chem.*, 2015, **17**, 2664–2678.
2 S. Bom, J. Jorge, H. M. Ribeiro and J. Marto, *J. Cleaner Prod.*, 2019, **225**, 270–290.
3 P. T. Anastas and J. B. Zimmerman, *Chem*, 2016, **1**, 10–12.
4 P. Coish, E. McGovern, J. B. Zimmerman and P. T. Anastas, in *Green Chemistry*, ed. B. Török and T. Dransfield, Elsevier, 2018, pp. 981–998. DOI: 10.1016/B978-0-12-809270-5.00033-9.
5 J. B. Zimmerman, P. T. Anastas, H. C. Erythropel and W. Leitner, *Science*, 2020, **367**, 397–400.
6 P. T. Anastas and J. B. Zimmerman, *Curr. Opin. Green Sustain. Chem.*, 2018, **13**, 150–153.
7 United Nations General Assembly, Transforming our world: the 2030 Agenda for Sustainable Development, https://www.refworld.org/docid/57b6e3e44.html, Accessed 4/29/2021.
8 P. Withisuphakorn, I. Batra, N. Parameswar and S. Dhir, *J. Bus. Retail. Manag. Res.*, 2019, **13**, (Special Edition), 35–47.
9 L. Leseurre, C. Merea, S. Duprat de Paule and A. Pinchart, *Green Chem.*, 2014, **16**, 1139–1148.
10 F. Johnson, *Harv. Bus. Rev.*, 2015, 33–36.
11 R. A. Sheldon, *ACS Sustainable Chem. Eng.*, 2018, **6**, 32–48.
12 P. T. Anastas and J. C. Warner, *Green Chemistry: Theory and Practice*, Oxford University Press, 1998.
13 B. J. Henry, J. P. Carlin, J. A. Hammerschmidt, R. C. Buck, L. W. Buxton, H. Fiedler, J. Seed and O. Hernandez, *Integr. Environ. Assess. Manage.*, 2018, **14**, 316–334.
14 T. Anderson and D. Darling, *Ann. Math. Stat.*, 1952, **23**, 193–212.
15 *R Foundation for Statistical Computing*, Vienna, Austria, 2013.
16 W. H. Kruskal and W. A. Wallis, *J. Am. Stat. Assoc.*, 1952, **47**, 583–621.
17 O. J. Dunn, *Technometrics*, 1964, **6**, 241–252.
18 D. Mackay, *Environ. Toxicol. Chem.*, 2015, **34**, 2687–2688.
19 European Chemicals Agency, *Annex XV Restriction Report: Proposal for Restriction: Substance Name: Decamethylcyclopentasiloxane (D5)*, Helsinki, Finland, 2019.
20 A. Fairbrother, G. A. Burton, S. J. Klaine, D. E. Powell, C. A. Staples, E. M. Mihaich, K. B. Woodburn and

F. A. P. C. Gobas, *Environ. Toxicol. Chem.*, 2015, **34**, 2715–2722.

21 European Chmicals Agency, *Background Document to the Opinion on the Annex XV dossier proposing restrictions on Octamethylcyclotetrasiloxane (D4) and Decamethylcyclopentasiloxane (D5)*, Helsinki, Finland, 2016.

22 A. M. Voutchkova, T. G. Osimitz and P. T. Anastas, *Chem. Rev.*, 2010, **110**, 5845–5882.