Improved iGAL 2.0 Metric Empowers Pharmaceutical Scientists to Make Meaningful Contributions to United Nations Sustainable Development Goal 12

Roschangar, Frank; Li, Jun; Zhou, Yanyan; Aelterman, Wim; Borovika, Alina; Colberg, Juan; Dickson, David P.; Gallou, Fabrice; Sheldon, Roger A.; More Authors

DOI
10.1021/acssuschemeng.1c01940

Publication date
2021

Document Version
Final published version

Published in
ACS Sustainable Chemistry and Engineering

Citation (APA)

Important note
To cite this publication, please use the final published version (if applicable). Please check the document version above.
Improved iGAL 2.0 Metric Empowers Pharmaceutical Scientists to Make Meaningful Contributions to United Nations Sustainable Development Goal 12


ABSTRACT: The large and steadily growing demand for medicines combined with their inherent resource-intensive manufacturing necessitates a relentless push for their sustainable production. Pharmaceutical companies are constantly seeking to perform reliable life cycle assessments of their medicinal products and assess the true value of their sustainable development achievements; however, they find themselves impeded by the lack of a universal metric system that allows for objective quantification of the underlying core denominators. Guided by the unambivalent purpose of the United Nations Sustainable Development Goal 12, which aims at substantially reducing production waste by 2030, and driven by a vision to catalyze greener active pharmaceutical ingredient (API) manufacturing around the globe, the authors set out to overcome current obstacles by defining an improved model for the metric named innovation green aspiration level, iGAL 2.0. We propose yield and convergence as new key sustainability indicators and include a new formula for convergence with potential applicability in computer assisted synthesis planning (CASP) algorithms. The improved statistical model of iGAL 2.0 represents a valuable extension to the common API process waste metrics, process mass intensity (PMI) and complete E factor (cEF), by putting those measures into perspective: iGAL 2.0 enables determination of relative process greenness (RPG) to identify potentially underperforming and environmentally concerning processes early and thereby deliver environmental value. At the same time, iGAL 2.0 generates economic value since reduced waste correlates to lower API production costs. The metric is complemented by its scorecard companion to highlight the impact of innovation on reductions of API manufacturing waste, enabling scientists to readily communicate the value of their work to their peers, managers, and the general public. We believe that iGAL 2.0 can readily be adopted by pharmaceutical firms around the globe and thereby empower and inspire their scientists to make meaningful and significant contributions to global sustainability.

KEYWORDS: Green chemistry metrics, Innovation green aspiration level (iGAL), Relative process greenness (RPG), Complete E factor (cEF), Convergence, United Nations Sustainable Development Goal (UN SDG), Active pharmaceutical ingredient (API), Scorecard, Life cycle assessment (LCA), Computer assisted synthesis planning (CASP)

1. INTRODUCTION

There is an urgent need for businesses to balance short- and long-term priorities and integrate environmental, social, and governance (ESG) principles into their strategy to mitigate risk and drive profitable growth. For this reason, pharmaceutical companies aim to balance growth with sustainability while advancing human and animal health and tie their ambitions to the United Nations sustainable development goals (SDG), which address key global challenges and form the cornerstone for a sustainable future, good health, and worldwide well-being. Among those, SDG 12 focuses on responsible and environmentally sustainable production, and targets the substantial reduction of waste generation by 2030. SDG 12 is thereby closely linked to the principles of green chemistry and engineering.

As first highlighted by Sheldon in 1992, the pharmaceutical industry’s environmental footprint, expressed as the E factor...
via kg of generated waste per kg of final product, is substantial: the manufacture of just 1 kg of commercial active pharmaceutical ingredient (API) of a medicine generates an estimated average 182 kg of waste.8 Put differently, fewer than 1% of all raw materials are incorporated into the API. Moreover, compared to small molecule drugs, the production mass efficiency for biological drugs is considerably less favorable.7 This represents both an environmental and an economic problem, as waste correlates to API production costs.8,9 Given that the demand for medicines is large and steadily growing,10 it is apparent that their sustainable production is critical. However, the absence of clearly defined standards and lack of published data has rendered it challenging for companies to perform consistent API life cycle assessments (LCA)11,12 to assess the true value of their sustainable development efforts and measure the degree of relative process greenness (RPG) of their API manufacturing processes, as defined below.

This fundamentally changed in 2018, when this team of authors, having collaborated across the International Consortium for Innovation & Quality in Pharmaceutical Development (IQ)13 and the American Chemical Society Green Chemistry Institute Pharmaceutical Roundtable (ACS GCI PR),14 introduced the innovation green aspiration level (iGAL).15 This metric filled the void and not only harmonized API process analysis via the iGAL rule but also enhanced the value of waste as an API LCA indicator by means of comparability with industry norms as derived from our own data set: iGAL empowered the determination of RPG. This enabled companies, for the first time, to report meaningful API manufacturing waste reduction figures toward SDG 12. It also provided an opportunity to identify potentially problematic API processes and prioritize their improvements with respect to sustainability and costs.16 Even though iGAL was a big step forward, we inferred from users that the sustainability indicators of iGAL’s statistical model, complexity and ideality, were not commonly used in the context of API process optimization and therefore not sufficiently intuitive. This presented a barrier to swift industry-wide embracement of iGAL.

Guided by the unambivalent purpose of SDG 12 and driven by our vision to motivate greener API manufacturing around the globe, we herein disclose the evolution of iGAL 2.0 and its considerable improvements. It deploys the more intuitive and tangible process sustainability indicators convergence (CV) and yield (YD) and substantially enhances the fit of our statistical model. Furthermore, the first quantitative formula for CV may find material utility in the synthetic organic chemistry community and with computer assisted synthesis planning (CASP) applications.16–19 We believe that the user-centric enhancements will help with broader adoption of iGAL 2.0 and thereby empower and motivate pharmaceutical scientists around the globe to make meaningful contributions to SDG 12.

2. METHODS

2.1. Scope. With respect to scope, our methodology is best explained in the context of the life cycle assessment (LCA) of a pharmaceutical product (Figure 1).54 A full LCA includes (1) upstream production of the starting materials, (2) core production of the packaged pharmaceutical or combination product from the starting materials via the API and drug product, and (3) downstream activities such as product distribution, use, and disposal. Herein, we focus on API manufacturing efficiency and therefore include the waste from the core production of the API from its starting materials, representing a gate-to-gate evaluation. The starting materials must be aligned with the iGAL rule and not exceed a $100/mole threshold20 (SI Chapter 1.1), which requires inclusion of external supply chain manufacturing steps. This rule forms the cornerstone of our standardized API process analysis.

Figure 1. Scope of iGAL 2.0 in the context of life cycle assessment for the manufacture of a pharmaceutical product encompassing API core production waste.
upstream and core production. To streamline assessment, in particular for early development processes with less available information, we treat reagents and catalysts, which do not contribute to the API structure, as commodities, i.e., we exclude their synthesis. Lastly, we include reaction workup but exclude waste from ancillary process operations such as analytical activities as well as reactor and equipment cleaning. In essence, all materials that are used in core production during processing and workup, including gaseous streams, but that do not get incorporated into the API structure, are considered process waste.

While developed for pharmaceutical process scientists in the context of overall API route design and development, we encourage students and members of the academic community to utilize iGAL 2.0. Our metric can help guide sustainable academic research by quantifying efficiency improvements and rating greenness of new and disruptive synthetic strategies and methods, such as multi-component reactions and biocatalysis, in the context of complex molecule synthesis.

2.2. Progressing from Complexity and Ideality to Convergence and Yield. In prior work, we had introduced iGAL as a unifying metric to enable meaningful and comparative evaluation of API manufacturing waste. Pharmaceutical waste is commonly reported as complete E factor (cEF) or process mass intensity (PMI = cEF + 1) and lets scientists capture improvements during development of a specific API process. iGAL provided more context to waste by establishing a realistic cEF-based waste goal for any API based on (1) the complexity of its molecular structure as denoted by its free molecular weight (FMW), and (2) an API complexity-adjusted pharmaceutical waste index mGAL. mGAL is a mathematical constant with a value of 0.403 that represents the adjusted pharmaceutical waste index mGAL. 27 mGAL is a unifying metric to enable meaningful and comparative evaluation of API process against industry averages and derive a quantitative and meaningful assessment of its greenness expressed as relative process greenness (eq 2).

\[
RPG = \frac{\text{iGAL}}{\text{cEF}} \times 100\%
\]

Hence, iGAL represented the first comparative metric where one could evaluate an API process against industry averages and derive a quantitative and meaningful assessment of its greenness expressed as relative process greenness (eq 2).

\[
RPG = \frac{\text{iGAL}}{\text{cEF}} \times 100\%
\]

By comparing the RPG score of an optimized API process with an earlier iteration, one could quantify the sustainable innovation impact achieved by the process development scientists (eq 3).

\[
\Delta \text{RPG} = \text{RPG (current process)} - \text{RPG (1st development process)}
\]

Since iGAL is a metric that provides a fixed waste target for each API based on its FMW, the key of upgrading RPG is to reduce waste by improving process efficiency parameters that strongly correlate to waste. We designate these parameters as key sustainability indicators (KSI). We found that the KSI complexity (CP = \( \sum \) construction steps) and ideality (I = CP/total steps \times 100%) were significant. Their correlation to waste could be expressed by a logarithmic regression equation, that exhibited a moderate 34% goodness-of-fit.

However, our goal to enable reporting of meaningful pharmaceutical contributions to SDG 12 through broad adoption of the iGAL methodology remained elusive following its rollout in 2018. Feedback from users revealed that the selected KSI were not optimal. They did not register as the right measures and consequently did not support our strategic intent. Consequently, we set out to replace the existing KSI with more relatable indicators. We subsequently adopted yield (YD) and convergence (CV), both of which are commonly used by industrial and academic scientists in the context of process efficiency and optimization. Surprisingly, CV has been a vague qualitative concept. Prior to our work, there was no substantiated quantitative formula for synthesis convergence, so it could not be calculated.

2.2.1. Process Yield (YD). Yield is a measure for step productivity and is based on the molar limiting starting material for a step. We define YD as the overall yield of the longest linear step sequence (LS) from an iGAL rule-aligned starting material to the final API. For processes with two or more LS starting materials with the same step distance to the API, the starting material with the largest contribution to the API structure, or largest atom economical molecular weight (MWAE), is prioritized. Yield reflects the cumulative product of yields LS across steps k (eq 4).

\[
YD = \prod_{k=1}^{LS} \text{yield}_k
\]

Special attention must be paid to steps where the limiting material is not part of the longest linear step sequence of the process (LS). In Figure 2, starting material 2 is not part of LS. If it is limiting and 2.0 equiv of intermediate 1 are used relative to starting material 2, then the yield of step 2 must be adjusted to 40% (=80% /2.0). As a result, YD = 32% (=80% \times 40%).

While there is no single measure for the efficiency of process design, YD is commonly used to gauge overall process productivity. However, YD does not provide information on the productivity of process steps outside of LS and the quality of the overall process design. This is where convergence as a complementary KSI comes into play.

2.2.2. Process Convergence (CV). A convergent step is one that combines two or more starting materials or intermediates. CV indicates how directly the starting materials are assembled into the API and therefore reflects the efficiency of API process design. The appeal of using CV in combination with YD is that they are orthogonal and pertain to two complementary dimensions of process efficiency: design efficiency (CV) and productivity (YD). We expected to encounter ample scientific literature for the required mathematical descriptor of the commonly used convergence concept but were surprised to uncover the opposite. We could not identify a fit-for-purpose solution, and so we decided to close this gap and develop a new formula for CV.

2.3. Deriving a New Quantitative Formula for Process Convergence. Over 40 years ago, Hendrickson was one of few researchers who mathematically characterized convergence in the context of synthesis design. With convergence, he broadened the scope of synthetic efficiency beyond the longest linear step sequence: Hendrickson used the sum of subprocess steps (SSS) to account for the step sequences of all starting materials i to the product, which are termed subprocess steps (SSi). He ascribed SSS as the extent or index of convergence (herein termed CV\textsubscript{Hendrickson}, eq 5).

\[
CV_{\text{Hendrickson}} = SSS = \sum_{i=1}^{SSS} S_{SI}
\]

We exemplify determination of SSS with the process shown in Figure 3.
SSS permitted evaluation of different processes with the same number of starting materials (SM). For a given SM, the process with lower SSS indicates a higher degree of convergence. We consider $CV_{\text{Hendrickson}}$ a valuable heuristic but are guided by the need to evaluate processes for different APIs with varying numbers of starting materials. For example, we can have two processes, A and B, with the same SSS but quite different extents of convergence (Figure 4). With $SSS = 3$ for both processes, it is not clear which process would be preferred. However, from a strategic perspective, process A is more preferred. In 2018, we established iGAL as a yardstick to empower and inspire scientists to design efficient and mass-efficient API manufacturing processes. We showed that waste reduction was correlated to improvements to the key sustainability indicators complexity and ideality, according to the logarithmic equation $\ln(CEF) = 5.789 + (0.1437 \times CP) - (1.725 \times I)$.15,35

2.3.1. Weighted Convergence Option. When evaluating alternative approaches to convergence, we considered weighted convergence ($CV_{\text{w}}$) because it addresses a potential shortcoming of $CV_{\text{GAL}}$. Which is the assumption that all starting materials contribute equally to the API structure. In actuality, starting materials are quite diverse and contribute to varying degrees, where starting materials with higher structural contributions may have a higher impact on convergence. We derive the formula for $CV_{\text{w}}$ in SI Chapter 3 and arrive at $CV_{\text{w}} = \sqrt{SSS} \times \frac{SM}{SSS}$ (eq 5). When applying the formula to the Viagra API, it turns out that $CV_{\text{w}}$ (21%) is higher than its $CV_{\text{GAL}}$ (16%). This result is driven by incorporation of large starting material 3, citric acid, in the last step of the synthesis to make the citrate salt of the API. We ultimately chose to incorporate $CV_{\text{GAL}}$ into iGAL 2.0, instead of $CV_{\text{w}}$ for reasons discussed in section 3.2.1 vide infra.

2.4. Statistical Model of iGAL 2.0. In 2018, we established iGAL as a way to quantify how financial outcomes as waste correlates to API manufacturing processes. We showed that waste reduction was correlated to improvements to the key sustainability indicators complexity and ideality, according to the logarithmic equation $\ln(CEF) = 5.789 + (0.1437 \times CP) - (1.725 \times I)$.35,38

By substituting CP and I with the KSI convergence and yield, which we believe to be more intuitive, we arrive at the new statistical model iGAL 2.0 that also turns out better at predicting waste. We used Minitab 18.1 statistical software to perform the regression analysis on our expanded data set of 100 API manufacturing processes, as detailed in SI Chapter 4.

2.5. Updating the Green Chemistry Innovation Scorecard. We created the green chemistry innovation scorecard to supplement the iGAL metric and support our strategic intent to reduce global API manufacturing waste.35 An important motivational component was the inclusion of key sustainability indicators, which explain how the goal of reducing waste is achieved. Our improved scorecard version is adapted to the new KSI under the new name iGAL 2.0 scorecard (Figure 5).

The scorecard is central to tracking and communicating the value-added of process scientists to all stakeholders, including the scientists themselves, their managers, and the public. Our hope is that it will inspire scientists in two ways: (1) by quantifying their achievements and innovation over the course of process development in terms of waste and RPG improvements and (2) by assessing the comparative performance of their process to industry via our standardized RPG rating matrix ranging from excellent to good, average, and below average. Concurrently, we hope that the scorecard will inspire managers via assessment of the impact of their scientist teams on innovation and on the public’s desire for a healthier environment with less waste. Regular scorecard updates and reviews can trigger early identification and mitigation of resource-inefficient API processes and, in turn deliver improved financial outcomes as waste correlates to API production costs.9

The updates to the iGAL 2.0 methodology necessitate adjustments of the four quadrants of the scorecard. The western quadrant displays the current status of the process including development phase of the API and now includes information on convergence, yield, waste as expressed by the complete E factor, and the calculated iGAL waste goal (eq 1a; for its derivation see SI Chapter 5).

Figure 3. Process where the orange node is the API, gray nodes are intermediates, and black nodes are starting materials. When counting the subprocess steps (SS) of the three starting materials 1 to 3 to the API, we obtain $SS_1 = 3$ steps, $SS_2 = 2$ steps, and $SS_3 = 1$ step. SSS represents the sum of the subprocess steps, so $SSS = SS_1 + SS_2 + SS_3 = 3 + 2 + 1 = 6$.

Figure 4. Two processes A and B with SSS = 3 and opposing degrees of convergence, where the orange node is the API, the gray nodes are intermediates, and the black nodes are starting materials.
The northern quadrant features the process performance rating compared to industry averages with consideration for development phase and API complexity. The underlying RPG rating matrix is shown in Table 5 vide infra. To render the rating more descriptive, we include information on how much more or less waste is generated relative to the industry average of processes in the same development phase. We determine this waste ratio according to SI Chapter 6.1, eqs S9 and S10. For example, if the RPG of the current process is higher than the industry average, then less waste is generated relative to an industry average process.

Shifting to the innovation impact quadrant on the eastern side of the scorecard, improvements to the KSI of an API manufacturing process, and their overall impact on RPG are displayed. To assess the full value-added of their contributions, users should derive the baseline RPG from the first process in early development. Under iGAL, it is not permitted to use the medicinal chemistry predecessor as a reference process, since divergent objectives render such comparisons meaningless. The sustainable innovation impact (ΔRPG) is determined per eq 3. Improvements to the KSI convergence and yield are reported via their discrete calculated impact on RPG (SI Chapter 6.2).

The southern waste reduction section shows both the absolute amount and relative improvement of waste reduction achieved through process innovation when comparing the results of the most recent API process to its reference. Scientists will find it easy to access the new iGAL 2.0 scorecard calculator, free-of-charge and without registration, on the ACS GCI PR Web site. The concise and purposeful graphical output can simply be copied and pasted into reports, presentations, and publications.

### Scheme 1. iGAL Rule-Aligned Viagra API Manufacturing Process

![Scheme 1](image)

### Scheme 2. Arrow Scheme Corresponding to Viagra API Process

![Scheme 2](image)

### Table 1. Determination of SS, for Starting Materials i from Viagra API Arrow Scheme

<table>
<thead>
<tr>
<th>i</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS&lt;sub&gt;i&lt;/sub&gt;</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

“For the Viagra API, we determine SSS = \( \sum_{i=1}^{10} SS_i = 63 \), and CV<sub>iGAL</sub> = SM/SSS \times 100% = 10/63 \times 100% = 16%.

iGAL = 0.403 \times \text{FMW} \quad (1a)

The northern quadrant features the process performance rating compared to industry averages with consideration for development phase and API complexity. The underlying RPG rating matrix is shown in Table 5 vide infra. To render the rating more descriptive, we include information on how much more or less waste is generated relative to the industry average of processes in the same development phase. We determine this waste ratio according to SI Chapter 6.1, eqs S9 and S10. For example, if the RPG of the current process is higher than the industry average, then less waste is generated relative to an industry average process.

Shifting to the innovation impact quadrant on the eastern side of the scorecard, improvements to the KSI of an API manufacturing process, and their overall impact on RPG are displayed. To assess the full value-added of their contributions, users should derive the baseline RPG from the first process in early development. Under iGAL, it is not permitted to use the medicinal chemistry predecessor as a reference process, since divergent objectives render such comparisons meaningless. The sustainable innovation impact (ΔRPG) is determined per eq 3. Improvements to the KSI convergence and yield are reported via their discrete calculated impact on RPG (SI Chapter 6.2).

The southern waste reduction section shows both the absolute amount and relative improvement of waste reduction achieved through process innovation when comparing the results of the most recent API process to its reference. Scientists will find it easy to access the new iGAL 2.0 scorecard calculator, free-of-charge and without registration, on the ACS GCI PR Web site. The concise and purposeful graphical output can simply be copied and pasted into reports, presentations, and publications.

### 2.6. Applicability of Convergence to CASP Tools

Devising an efficient API manufacturing process is challenging due to compressed pharmaceutical development timelines and high molecular complexity of the APIs. To overcome these challenges, process scientists rely on retrosynthetic analysis, which helps identify efficient process options by reverse-engineering the API structure via iterative bond disconnections that optimally reduce complexity all the way back to the starting materials. To facilitate retrosynthetic planning, artificial intelligence (AI) in computer assisted synthesis planning (CASP) is increasingly being deployed, and three principal methods have emerged based on algorithms by Jensen.
Grzybowski (Chematica/Synthia),\textsuperscript{44} and Segler.\textsuperscript{45} However, while the merit of convergence in synthetic planning has long been recognized,\textsuperscript{46} the concept could not be broadly adapted into CASP algorithms due to the absence of substantiated mathematical descriptors.\textsuperscript{35}

We have now developed a useful formula for convergence and demonstrated its value as a fundamental concept for assessing process efficiency \textit{vide infra}. While iGAL 2.0 methodology works “retrospectively”, i.e., it evaluates process performance of completed API campaigns, we believe that the convergence KSI can be deployed in a “forward” direction for process design. For example, CV\textsubscript{iGAL} may find utility in CASP applications by means of integrating predicted waste economics as selection criterion and thereby help narrow down the multitude of conceivable synthesis routes to a more manageable level. We provide an example in section 3.4.

3. RESULTS AND DISCUSSION

We generated the underlying body of data for our evaluation from the iGAL rule-aligned API manufacturing processes contributed by 13 pharmaceutical companies, and the application of the KSI formulas for convergence and yield described in the Methods section. The detailed data set encompassing 100 process analyses, derived from the arrow schemes shown in SI Chapter 2, is tabulated in SI Chapter 4, and averages of key sustainability measures and indicators derived therefrom, grouped by API phase, are displayed in Table 2.

![Figure 5. Upgraded contents and updated visuals in iGAL 2.0 scorecard. Integrating convergence and yield as new KSI while emphasizing contributions to waste reduction in the context of the United Nations SDG 12. This shall not imply or suggest the UN’s endorsement or seal of approval of the metric. Reprinted (adapted or reprinted in part) with permission from the ACS Green Chemistry Institute Pharmaceutical Roundtable and the International Consortium for Innovation & Quality in Pharmaceutical Development.](https://doi.org/10.1021/acssuschemeng.1c01940)

Table 2. Averages for Key Sustainability Measures and Indicators Derived from Our Dataset of 97\textsuperscript{a} API Manufacturing Processes

<table>
<thead>
<tr>
<th>API phase</th>
<th>N\textsuperscript{b}</th>
<th>FMW\textsuperscript{c} [g/mol]</th>
<th>CV\textsubscript{iGAL} [%]</th>
<th>YD [%]</th>
<th>cEF\textsuperscript{d} [kg/kg API]</th>
<th>iGAL\textsuperscript{e} [kg/kg API]</th>
<th>RPG\textsuperscript{f} [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>early development\textsuperscript{g}</td>
<td>35</td>
<td>458</td>
<td>15</td>
<td>20</td>
<td>632</td>
<td>185</td>
<td>39</td>
</tr>
<tr>
<td>late development\textsuperscript{h}</td>
<td>33</td>
<td>438</td>
<td>20</td>
<td>35</td>
<td>308</td>
<td>177</td>
<td>84</td>
</tr>
<tr>
<td>commercial\textsuperscript{i}</td>
<td>29</td>
<td>451</td>
<td>25</td>
<td>44</td>
<td>182</td>
<td>182</td>
<td>138</td>
</tr>
</tbody>
</table>

\textsuperscript{a}After removing three outliers (see SI Chapter 7.1). \textsuperscript{b}N = number of API manufacturing process data sets. \textsuperscript{c}FMW = free molecular weight of API.\textsuperscript{26} \textsuperscript{d}cEF = complete E factor, API manufacturing waste. \textsuperscript{e}iGAL = innovation green aspiration level; complexity-adjusted commercial API manufacturing waste goal. \textsuperscript{f}RPG = relative process greenness; indicator for sustainable performance based on generated waste relative to commercial averages; for an explanation of the average commercial RPG equaling 138%, refer to ref 15 and SI Chapter 4. \textsuperscript{g}Campaigns making API supplies for up to phase IIa/proof of concept clinical trials. \textsuperscript{h}API campaigns supporting phase IIb clinical trials up to registration. \textsuperscript{i}API campaigns providing market-scale supplies.
included both candidates for convergence, CV_{iGAL} and CV_{wt}, and for comparison purposes the KSI complexity and ideality from the prior iGAL model.

Table 3. Results of Simple Regression of KSI Predictors against ln(cEF) Response Shows Higher R-sq for YD and CV

<table>
<thead>
<tr>
<th>KSI</th>
<th>R-sq [%]</th>
<th>P-value</th>
<th>coef</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP</td>
<td>24</td>
<td>0.000</td>
<td>0.14</td>
</tr>
<tr>
<td>I</td>
<td>5.1</td>
<td>0.029</td>
<td>-1.16</td>
</tr>
<tr>
<td>YD</td>
<td>49</td>
<td>0.000</td>
<td>-2.92</td>
</tr>
<tr>
<td>CV_{iGAL}</td>
<td>34</td>
<td>0.000</td>
<td>-7.02</td>
</tr>
<tr>
<td>CV_{wt}</td>
<td>37</td>
<td>0.000</td>
<td>-7.43</td>
</tr>
</tbody>
</table>

“KSI = key sustainability indicator. CP = complexity. I = ideality. YD = overall yield of longest linear process step sequence. CV_{iGAL} = our new convergence metric. CV_{wt} = the new weighted convergence metric.”

We are pleased that the relationships between waste and the new KSI are statistically significant (P-value < 0.05). Also, and in line with expectation, the negative coefficients of convergence and yield denote that waste decreases as CV or YD increases. Encouragingly, YD and the two CV display substantially higher R-sq values than CP and I, signaling that the new indicators explain more of the variability in waste and indicating a better fit. Having herewith validated CV and YD as not only suitable but also improved iGAL KSI, we are now ready to proceed to the multiple regression model and evaluate the combined effect of CV and YD on waste reduction.

3.2. Improved Statistical Model of iGAL 2.0. After removing three outliers from our data set (SI Chapter 7.1), we ran the multiple regression analysis of CV_{iGAL} and YD with ln(cEF) as the response variable. The results are displayed in Table 4. The large P-value from the lack-of-fit test in Table 4a is a strong indication that the model fits adequately. The negative signs of the KSI coefficients in Table 4c indicate the proper direction of their relationship to waste. multicollinearity between the KSI is moderate (1 < VIF < 5) and the goodness-of-fit of the model is 55% (R-sq).

We arrive at the new iGAL 2.0 statistical model, where the relationship of waste to the KSI is expressed via regression eq 8.

**iGAL 2.0 Statistical Model:**

\[
\ln(cEF) = 6.913 - (3.130 \times CV_{iGAL}) - (1.987 \times YD)
\]

The formula provides an indication for the impact of convergence and yield improvements on API process waste as illustrated in Figure 6.

3.2.1. CV_{iGAL} Prevails Over Weighted Convergence Alternative. Next, we gauged weighted convergence as an alternative to CV_{iGAL} by running the multiple regression analysis of ln(cEF) vs CV_{wt} and YD. We obtained a valid statistical model with a comparable R-sq of 54% (SI Chapter 7.2).

\[
\ln(cEF) = 6.867 - (2.770 \times CV_{wt}) - (2.077 \times YD)
\]

We regarded it important to keep the KSI simple yet meaningful in order to catalyze industry adoption of iGAL 2.0. Since CV_{wt} does not provide an improved goodness-of-fit (R-sq of 54% vs 55% from CV_{iGAL}) and requires the extra steps of determining MW_{AE} and then SS_{AE} for the starting materials, we chose CV_{iGAL} as the more user-friendly convergence KSI for iGAL 2.0.

Finally, we examined how the new iGAL 2.0 model fared when compared to its prior complexity- and ideality-based version (iGAL 1.0) by running the multiple regression analysis of ln(cEF) vs CP and I using the new expanded data set. We

Table 4. Results of Regression Analysis of ln(cEF) Response vs CV_{iGAL} and YD Predictors Using Minitab 18.1

<table>
<thead>
<tr>
<th>source</th>
<th>DF</th>
<th>seq SS (%)</th>
<th>adj SS (%)</th>
<th>adj MS</th>
<th>F-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>regression</td>
<td>2</td>
<td>36.1474</td>
<td>36.1474</td>
<td>18.0073</td>
<td>58.20</td>
<td>0.000</td>
</tr>
<tr>
<td>CV_{iGAL}</td>
<td>1</td>
<td>27.7710</td>
<td>24.4522</td>
<td>27.7110</td>
<td>89.75</td>
<td>0.000</td>
</tr>
<tr>
<td>YD</td>
<td>1</td>
<td>8.2437</td>
<td>8.2437</td>
<td>8.2437</td>
<td>26.64</td>
<td>0.000</td>
</tr>
<tr>
<td>error</td>
<td>94</td>
<td>29.0865</td>
<td>29.0865</td>
<td>29.0865</td>
<td>30.934</td>
<td>0.91</td>
</tr>
<tr>
<td>lack-of-fit</td>
<td>93</td>
<td>28.7460</td>
<td>28.4760</td>
<td>28.4760</td>
<td>10.64</td>
<td>0.52</td>
</tr>
<tr>
<td>pure error</td>
<td>1</td>
<td>0.3405</td>
<td>0.3405</td>
<td>0.3405</td>
<td>0.3405</td>
<td>0.14</td>
</tr>
<tr>
<td>total</td>
<td>96</td>
<td>65.1012</td>
<td>65.1012</td>
<td>65.1012</td>
<td>65.1012</td>
<td>100.00</td>
</tr>
</tbody>
</table>

b. Model Summary

<table>
<thead>
<tr>
<th>S</th>
<th>R-sq [%]</th>
<th>R-sq(adj) [%]</th>
<th>PRESS</th>
<th>R-sq(pred) [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.556265</td>
<td>55.32</td>
<td>54.37</td>
<td>30.9901</td>
<td>52.40</td>
</tr>
</tbody>
</table>

c. Coefficients

<table>
<thead>
<tr>
<th>term</th>
<th>coef</th>
<th>SE coef</th>
<th>95% CI</th>
<th>T-value</th>
<th>P-value</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>6.913</td>
<td>0.166</td>
<td>(6.583, 7.243)</td>
<td>41.60</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>CV_{iGAL}</td>
<td>-3.13</td>
<td>1.11</td>
<td>(-5.33, -0.92)</td>
<td>-2.81</td>
<td>0.006</td>
<td>2.10</td>
</tr>
<tr>
<td>YD</td>
<td>-1.987</td>
<td>0.385</td>
<td>(-2.751, -1.222)</td>
<td>-5.16</td>
<td>0.000</td>
<td>2.10</td>
</tr>
</tbody>
</table>

“Tests use the sequential sums of squares. DF = degrees of freedom. seq SS = sequential sums of squares. adj SS = adjusted sums of squares. adj MS = adjusted mean squares. CV_{iGAL} = our new convergence metric. YD = overall yield of longest linear process step sequence. S = standard deviation of the distance between the data values and the fitted values. R-sq = R^2; percentage of variation in the response that is explained by the model. R-sq(adj) = adjusted R^2. PRESS = prediction error sum of squares. R-sq(pred) = predicted R^2. coef = coefficient. SE coef = standard error of the coefficient. CI = confidence interval for coefficient. VIF = variance inflation factor; indicates how much the variance of a coefficient is inflated due to the correlations among the predictors in the model.”
obtained a valid statistical model but with inferior R-sq = 40% vs the 55% of iGAL 2.0 (SI Chapter 7.3).

Our comparison has shown a 2-fold advantage of iGAL 2.0 (eq 8) over iGAL 1.0: (1) it is a better statistical descriptor of cogenerated API manufacturing waste, and (2) it is expected to be more intuitive to users via the KSI convergence and yield, which are common dimensions for assessing API process efficiency.

3.3. More Relevant iGAL 2.0 Scorecard. On the basis of the new data set and iGAL 2.0 model, we updated the RPG rating matrix required for process appraisal in the scorecard’s northern performance quadrant. We used the probability plot from Minitab 18.1 to estimate the 90, 70, and 40 percentiles for each of the three phases (Figure 7).

The top 10% of the RPG scores for each phase is classified “excellent”, the 70 percentile is ranked “good”, the 40 percentile “average”, and the bottom 40% rate “below average”. From here we deduced the iGAL 2.0 RPG rating matrix shown in Table 5.

We emphasize that the iGAL metric and RPG score for an API process are based on the commercial waste average (SI Chapter 5). For this reason, processes in an earlier development trend lower in RPG as less time and resources have been invested into their optimization. We remedy this through our phase-dependent RPG rating system, which provides a
mechanism for equitable assessment of API processes in any phase.

To exemplify the new iGAL 2.0 Scorecard, we took the Dabigatran API (FMW = 628; SI Chapter 6, Scheme S1).\textsuperscript{47} We used its baseline process from early development (Table S2-1, entry 18), compared it to an optimized commercial phase process (Tables S2-3, entry 85), and collected the required scorecard inputs for cEF, CViGAL, and YD (Table 6).

The new open-access iGAL 2.0 scorecard calculator, which can be used without registration on the ACS GCI PR Web site,\textsuperscript{40} allows one to enter the required inputs for the early development first generation reference process under Campaign 1, and for the commercial phase third generation process in the row labeled Campaign 2 (Figure 8).

The new open-access iGAL 2.0 scorecard calculator, which can be used without registration on the ACS GCI PR Web site,\textsuperscript{40} allows one to enter the required inputs for the early development first generation reference process under Campaign 1, and for the commercial phase third generation process in the row labeled Campaign 2 (Figure 8).

After clicking on the gray "Campaign 2" rectangle, we obtained the iGAL 2.0 scorecard output (Figure 9). This graphic can be conveniently copied and pasted as a high-quality image into any desired presentation to highlight the sustainability impact of process innovation.

### 3.4. Added Value of the New Convergence Algorithm to CASP Applications

After having validated the statistical relationship between waste, which is an indicator for process efficiency, and convergence (CV\textsubscript{iGAL} per eq 7 and CV\textsubscript{wt} per eq S5), we set out to exemplify the utility of a convergence algorithm for CASP applications. This was accomplished via an exercise with Umifenovir API, an investigational COVID-19 therapeutic (Figure 10). The added value of the convergence selection criterion lies in its ability to filter or rank the output from CASP applications for greater resource-efficiency and thus cost-effectiveness.\textsuperscript{48}

We illustrate application of the convergence algorithm to five CASP-derived Umifenovir API synthesis plans: two proposals developed by Cernak’s research group using the CASP application Synthia (SI Chapter 8.1, Scheme S2),\textsuperscript{49} and three ideas from our team using the ASKCOS platform (Scheme S3).\textsuperscript{50−52} To provide consistent convergence scores, we aligned the proposals with the $100/mol starting material iGAL rule. The results are shown in Table 7.

We observed that among the five candidates, proposals no. 1 and no. 3 have the highest convergence scores for both CV\textsubscript{iGAL} and its weighted variant CV\textsubscript{wt} and are therefore the most promising options for resource-efficient and economical API synthesis. Based on their estimated average waste of 186 and 192 kg per kg Umifenovir API, respectively, these synthesis plans are capable of achieving the iGAL 2.0 target of 192 kg waste and thus reaching RPG $\geq 100\%$. This example demonstrates how convergence can be used as an algorithm to focus a larger number of CASP-derived proposals to a smaller and more manageable number of synthesis plans with

### Table 5. iGAL 2.0 RPG Rating Matrix to Assess Sustainability Performance of API Processes by Phase

<table>
<thead>
<tr>
<th>Percentile (PCTL)</th>
<th>Color Code</th>
<th>Rating</th>
<th>Minimum RPG for</th>
<th>innovation impact (=\Delta\text{RPG} \ [%])</th>
<th>waste reduction/kg API</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>🔵</td>
<td>Excellent</td>
<td>69%</td>
<td>158%</td>
<td>245%</td>
</tr>
<tr>
<td>70%</td>
<td>🔴</td>
<td>Good</td>
<td>51%</td>
<td>114%</td>
<td>182%</td>
</tr>
<tr>
<td>40%</td>
<td>🟠</td>
<td>Average</td>
<td>33%</td>
<td>69%</td>
<td>117%</td>
</tr>
<tr>
<td>-</td>
<td>🟡</td>
<td>Below Average</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Mean Industry RPG\textsuperscript{a} 39\% 84\% 138\%

\textsuperscript{a}Derived from our data set (Tables S2-1 to S2-3).

### Table 6. Assessing Sustainability Performance, Innovation Impact, and Waste Reduction for Dabigatran API

<table>
<thead>
<tr>
<th>API phase</th>
<th>cEF [kg/kg]</th>
<th>CV\textsubscript{iGAL} [%]</th>
<th>YD [%]</th>
<th>RPG [%]</th>
<th>scorecard iGAL 2.0 rating</th>
<th>innovation impact (=\Delta\text{RPG} \ [%])</th>
<th>waste reduction/kg API</th>
</tr>
</thead>
<tbody>
<tr>
<td>early development</td>
<td>252</td>
<td>14</td>
<td>19</td>
<td>100</td>
<td>excellent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1st generation)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>commercial</td>
<td>89</td>
<td>21</td>
<td>57</td>
<td>284</td>
<td>excellent</td>
<td></td>
<td>184</td>
</tr>
<tr>
<td>(3rd generation)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>163 kg</td>
</tr>
</tbody>
</table>

Figure 8. Screenshot of the iGAL 2.0 scorecard input screen populated with information for the early development and commercial phase Dabigatran API processes.
improved predicted environmental impact and process economics.

Further in-depth evaluation of the leading proposals can be performed, for example, with the user-friendly PMI predictor, which takes into account the specific type of chemistry for each step. When applying the tool to the Synthia no. 1 and ASKCOS no. 3 proposals, we predict process waste of 331 and 157 kg, respectively (SI Chapter 8.2), suggesting that it may be prudent to prioritize evaluation of ASKCOS no. 3 for the Umifenovir API.54

Due to the simplicity of the new convergence formula as an indicator for resource-efficiency, we believe that CV_iGAL and CV_wt represent a valuable algorithm for CASP applications,
allowing for straightforward integration of sustainability and economic considerations into retrosynthesis planning.

4. CONCLUSION

4.1. Evaluation of Results. We have demonstrated that the new iGAL 2.0 methodology can serve as a tool to enable significant reduction of global API manufacturing waste. To make industry-wide adoption easier, we introduced two new key sustainability indicators: convergence and yield. For convergence, we established a simple new formula and demonstrated its value as a fundamental indicator for process efficiency and sustainability. Furthermore, we exemplified how this KSI may find utility in CASP applications by means of integrating waste economics as selection criterion during synthesis design.

Convergence and yield are pragmatic for scientists in their day-to-day activities as tools to judge the mass efficiency of their API processes and interpret their innovation impact. As a strong statistical indicator for waste, the concept of convergence is of high relevance to innovation in early development, before the API process gets locked, which typically occurs during Phase 2 for the manufacture of Phase 3 clinical trial supplies. By firmly integrating convergence into early process design, one could minimize the risk of locking a process with suboptimal convergence to prevent creation of a low ceiling for subsequent sustainability improvements.

The unique advantage of our methodology is that it secures meaningful process waste analysis in two ways: (1) via the iGAL rule, which clearly defines the process starting materials and thereby ensures consistent analysis, and (2) via the RPG score, which enables scientists to compare the sustainability performance of their process with industry averages. We envision two high-value-added applications:

1. Applying the iGAL rule to the largely inconsistent cradle-to-gate life cycle assessment of pharmaceutical products to secure streamlined and uniform analysis of API impact. The iGAL rule requires inclusion of external supply chain manufacturing steps and standardizes the “first gate” of the API process. While this approach would modestly underestimate the full environmental impact of pharmaceutical products due to the moderate $100/mol raw material cost ceiling, it would alleviate significant challenges and inconsistencies associated with obtaining life cycle inventory (LCI) data of materials needed for the upstream manufacture of simple commodity chemicals that are often not readily available.

2. Implementing iGAL 2.0 across the entire API portfolio of a pharmaceutical company provides an opportunity to a. report meaningful API manufacturing waste figures in the context of SDG 12 b. establish “target sustainability profiles” for API manufacturing and guide process development to

### Table 7. Convergence Scores and Predicted Resource Efficiency (cEF) for the Five CASP-Generated and iGAL Rule-Aligned Umifenovir Synthesis Options

<table>
<thead>
<tr>
<th>Synthesis Plan</th>
<th>Synthia #1</th>
<th>Synthia #2</th>
<th>ASKCOS #3</th>
<th>ASKCOS #4</th>
<th>ASKCOS #5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arrow Scheme</td>
<td><img src="image1.png" alt="Arrow Scheme" /></td>
<td><img src="image2.png" alt="Arrow Scheme" /></td>
<td><img src="image3.png" alt="Arrow Scheme" /></td>
<td><img src="image4.png" alt="Arrow Scheme" /></td>
<td><img src="image5.png" alt="Arrow Scheme" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LS (steps)</th>
<th>6</th>
<th>7</th>
<th>6</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calc. YD</td>
<td>41%</td>
<td>35%</td>
<td>41%</td>
<td>26%</td>
<td>22%</td>
</tr>
<tr>
<td>SS relationships</td>
<td>3.6</td>
<td>4.1</td>
<td>3.7</td>
<td>6.9</td>
<td>5.9</td>
</tr>
<tr>
<td>CVIgAL</td>
<td>28%</td>
<td>24%</td>
<td>27%</td>
<td>15%</td>
<td>17%</td>
</tr>
<tr>
<td>CVw</td>
<td>25%</td>
<td>23%</td>
<td>26%</td>
<td>15%</td>
<td>16%</td>
</tr>
<tr>
<td>Expected cEF, based on CVIgAL</td>
<td>186</td>
<td>234</td>
<td>192</td>
<td>382</td>
<td>378</td>
</tr>
<tr>
<td>[kg waste / kg API]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* YD = overall yield of longest linear process step sequence. ^b Calc. YD = 86%15; we assume 86% average commercial step yield as derived from our data set (Table S2-3). *SSavg = average subprocess steps from all iGAL rule-aligned starting material to the API. CVIgAL = our new convergence metric. CVw = the new weighted convergence metric.
identify potentially eco-underperforming processes via their RPG scores, and subsequently focus resources on their optimization to deliver environmental and inherent economic value.

c. reduce overall production costs

We also introduced the complementary iGAL 2.0 scorecard companion to display (a) progress toward reducing overall API manufacturing waste and improving sustainable process performance, and (b) the impact of the KSI on RPG as benchmark for process innovation. The scorecard is a communication tool intended to inspire scientists by highlighting their achievements to peers, managers, and the general public. For the latter, the scorecard demonstrates how the pharmaceutical scientists contribute to the wellbeing of the planet with respect to United Nations SDG 12.

We believe that implementation of iGAL 2.0 across the pharmaceutical industry will stimulate fruitful competition in the spirit of SDG 12 by allowing for industry comparability. This creates a win-win situation for process efficiency as well as sustainable development.

4.2. Limitations. While the iGAL 2.0 methodology enables objective and consistent assessment of waste by standardizing the “first gate” of the API process via the iGAL rule, which requires inclusion of external supply chain manufacturing steps, it is important to outline its limitations with respect to API class and method.

Regarding API class, iGAL 2.0 was established based on small molecule API with FMW ranging from 200 to 873 and averaging 449 (Table 2 and Tables S2-1 to S2-3). Since our model has not been scaled beyond this range, we cannot predict whether we require a new statistical model and iGAL waste target for medium-sized APIs such as peptides and oligonucleotides with MW ranging from 1 000 to 5 000, and for the even larger-sized biological drugs.

In terms of the method, the scope of iGAL 2.0 process analysis reflects an iGAL rule-aligned gate-to-gate API process waste assessment as outlined in section 2.1. We note that iGAL 2.0 considers solely the mass of the process waste but not its type. We need to bear in mind that the green credentials of the overall API process are impacted by additional qualitative factors that should be evaluated during its development. These include elemental sustainability, renewability, and the environmental and health impact of reagents, reactants, and solvents, which can be assessed, for example, with the metrics toolkit developed by the CHEM21 consortium.57

4.3. Outlook and Future Research. We envision to expand the scope of the methodology and include the environmental impact of starting materials for small molecule APIs via cradle-to-gate assessment.56 We intend to incorporate carbon footprint reporting (Figure 1) and establish a motivational carbon footprint goal for API manufacturing, iGAL(CO2e). This would extend our ability to integrate carbon footprint in "target sustainability profiles" for API manufacturing and stimulate meaningful contributions to SDG 13 to combat climate change. A particularly valuable extension of the iGAL 2.0 methodology may be the inclusion of environmental impact analysis for the other drug manufacturing activities: drug product, packaging, and combination devices. This would provide the complete picture for expected waste and carbon footprint of our medicinal products. Other future activities may include expansion of the methodology from small molecule APIs to mid-sized and large molecules, which may require establishing new compound-class specific iGAL waste targets and new key sustainability indicators.

Furthermore, we plan to reach out to software developers for potential integration of the new convergence score into CASP software to allow for its use as an important environmental and economical criterion for API synthesis route prioritization and AI decision making.

Since one of the pharmaceutical industry’s main goals is to ensure a reliable supply of affordable and life-saving medicines within a sustainable and economic and environmental framework while focusing on environmental impact, we also consider iGAL 2.0 to be an adequate starting point to approach the World Economic Forum58 with the purpose of reaching a broader audience. This will be aided by showcasing the value of iGAL 2.0 with real-world case studies from participating companies.

In summary, we have advanced and substantiated iGAL 2.0 and its scorecard as unique and powerful tools to empower and inspire pharmaceutical scientists to make meaningful contributions to SDG 12, which aligns with the UN’s ambition to foster a sustainable future for people, animals, and the planet.
Funding
No funding was received to assist with the preparation of this manuscript.

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS
This manuscript was developed with the support of the International Consortium for Innovation and Quality in Pharmaceutical Development (IQ, http://www.iqconsortium.org). IQ is a not-for-profit organization of pharmaceutical and biotechnology companies with a mission of advancing science and technology to augment the capability of member companies to develop transformational solutions that benefit patients, regulators, and the broader research and development community. The authors thank Christiana Briddell (ACS) for integrating the iGAL 2.0 Scorecard app into the ACS GCI Pharmaceutical Roundtable website. F. R. thanks Steve Han, Hao Wu, Bodo Betzemeier, and Markus Gläßig (Boehringer Ingelheim) for their assistance with the process analyses. E. O’B. thanks Wenli Liang, Aireal Jenkins, Erwin Irdam, and Donald Walker (Biogen). J.D.H. thanks Gregg Barcan, Lee Boulton, Sabine Fenner, John Lim, and Alastair Roberts (GSK) for their assistance with the process analyses. A.G.S. thanks Margaret Faul, Oliver Thiel, and Sharon Michalak (Amgen) for their support of this project and support of Amgen’s Green Chemistry initiative.

ABBREVIATIONS
ACS American Chemical Society
ACS GCI PR ACS Green Chemistry Institute Pharmaceutical Roundtable
AI artificial intelligence
API active pharmaceutical ingredient, drug substance
CASP computer assisted synthesis planning
CV API manufacturing process convergence
CViGAL novel descriptor of convergence selected as KSI for the iGAL 2.0 model
CVwt novel descriptor of convergence using MWAE-based weighting of starting materials. Not selected for iGAL 2.0 model but could be useful in CASP applications
cEF complete E factor, API manufacturing waste from iGAL-rule aligned process starting materials
FMW free molecular weight, molecular weight of the API excluding its salt, cocrystal and solvate components
iGAL innovation Green Aspiration Level; complexity-adjusted commercial API manufacturing waste goal
iGAL 1.0 statistical waste model based on the KSI complexity (CP) and ideality (I)
iGAL 2.0 improved statistical waste model based on the KSI convergence and yield
iGAL rule $100/mol lab catalog pricing threshold for API process starting materials. The iGAL rule requires inclusion of external supply chain manufacturing steps and standardizes the “first Gate” of the API process.
IQ International Consortium for Innovation & Quality in Pharmaceutical Development
kg CO2e kg carbon dioxide equivalent describing standardized carbon footprint
KSI key sustainability indicator; process efficiency parameters that strongly correlates to waste life cycle assessment
LCA
LS longest linear step sequence for an iGAL-rule aligned starting material to the API
MWAE atom economical molecular weight; portion of MW of a starting material that gets incorporated into API structure
PMI API manufacturing process mass intensity = cEF + 1
RPG relative process greenness; indicator for sustainable performance based on API manufacturing waste relative to commercial averages
ΔRPG sustainable innovation impact as defined by eq 3
SDG United Nations sustainable development goal
SM number of iGAL rule-aligned API starting materials
SSavg average of all subprocess steps from iGAL-aligned starting material to the API
SSi subprocess steps from an iGAL-aligned starting material i to the API
Sustainable Development Goals

Simplification in Life Cycle Assessments.

Sustainability Assessments of Pharmaceuticals: An Emerging Need for Barriers to Green Chemistry in the Pharmaceutical Industry – the Green Aspiration Level\textsuperscript{TM} Concept. Green Chem. 2015, 17 (2), 752–768.

CO\textsubscript{2}e = Carbon dioxide equivalent is a unit that describes the standardized carbon footprint associated with, for example, energy consumption and waste. It signifies the amount of CO\textsubscript{2} which would have the equivalent global warming impact.

In our prior disclosure (ref 15) we used \textit{overall process improvement} instead of \textit{sustainable innovation impact} to quantify the value-added of process scientists. We believe the latter is a better descriptor in the context of this work.

In our prior disclosure (ref 15) we used \textit{key process performance indicators (KPPI)} instead of \textit{key sustainability indicators (KSI)}. We believe KSI represents a better descriptor in the context of this work.

The goodness-of-fit of a regression model is expressed by R\textsuperscript{2}.

A limiting starting material limits the amount of product that can be formed in a step and forms the basis for determining step yield from the ratio of actual amount vs theoretical maximum of a step product.

ACS Sustainable Chemistry & Engineering

SSS sum of subprocess steps of all iGAL-rule aligned starting materials to the API

UN United Nations

YD API manufacturing process yield

\section*{REFERENCES}


(6) The 182 kg average commercial API manufacturing waste stems from iGAL rule-aligned analysis of commercial manufacturing processes of this work, see Table S2-3.


(9) Amgen described a clear correlation between API manufacturing waste and process costs, in: Tucker, J. L.; Faul, M. M. Industrial research: Drug companies must adopt green chemistry. Nature 2016, 534 (7605), 27−29.


(21) CO\textsubscript{2}e = Carbon dioxide equivalent is a unit that describes the standardized carbon footprint associated with, for example, energy consumption and waste. It signifies the amount of CO\textsubscript{2} which would have the equivalent global warming impact.


(26) FMW = free molecular weight of the API, which represents the molecular weight of the API excluding its salt, cocrystal, solvate and hydrate components.

(27) Based on the expanded data set, we herein update mGAL = 403 (SI Chapter 5).

(28) In our prior disclosure (ref 15) we used \textit{overall process improvement} instead of \textit{sustainable innovation impact} to quantify the value-added of process scientists. We believe the latter is a better descriptor in the context of this work.

(29) In our prior disclosure (ref 15) we used \textit{key process performance indicators (KPPI)} instead of \textit{key sustainability indicators (KSI)}. We believe KSI represents a better descriptor in the context of this work.


(31) The goodness-of-fit of a regression model is expressed by R\textsuperscript{2}.


(33) A limiting starting material limits the amount of product that can be formed in a step and forms the basis for determining step yield from the ratio of actual amount vs theoretical maximum of a step product.
(36) "The ideal synthesis... may be defined as one in which the target molecule is prepared from readily available starting materials in one simple, safe, environmentally-acceptable, and resource-effective operation that proceeds quickly and in quantitative yield." Wender, P. A.; Miller, B. L. Organic Synthesis: Theory and Applications; Hudlicky, T., Ed.; JAI Press: Greenwich, CT, 1993; Vol. 2, pp 27–66.


(38) In our prior work (ref 15 – SI Chapter 6) we showed that the cEF data followed an approximately log-normal distribution. Running a multiple linear regression with cEF as the dependent variable resulted in violations of the model assumptions. Both the normality assumption, which was assessed with a Kolmogorov–Smirnov test, and the homoscedasticity assumption, which was assessed with residual plots, were not met. As a result, the natural logarithm transformation of the dependent variable was used, and the assumptions for normality and homoscedasticity of the error terms were reassessed. Both assumptions were now met. The final regression model was run with the transformed dependent variable: the natural logarithm of cEF. All effect interpretations were back-transformed.

(39) Comparing a development process to a medicinal chemistry synthesis is common practice to highlight process improvements for a specific API. However, these improvements are not meaningful since the goal of a medicinal chemistry synthesis is to create diversity and not an efficient process to a specific API. Thus, a medicinal chemistry synthesis should not be chosen as reference.


(47) The manufacturing processes of the Dabigatran API are shown in Figure 2 of ref 15.

(48) We note that CASP applications do not capture all operations that are required to meet the high-quality control demands of an API process, such as crystallizations, distillations, and salt formations. Therefore, the predicted API manufacturing waste derived from these applications is most likely underestimated.


(54) The large variation in waste from the PMI Predictor compared to the iGAL 2.0 model is primarily due to the resource-intensive nitration of Syntha no. 1 at the beginning of the longest linear step sequence.


(56) A streamlined tool for estimating cradle-to-gate LCA indicators for pharmaceutical APIs is under development by the ACS GCI PR and will make use of the iGAL rule for defining boundary conditions to ensure consistent application.
