



Cite this: *Green Chem.*, 2025, **27**, 6699

How to correctly evaluate greenness, whiteness and other “colours”? Introducing general rules of a good evaluation practice

Paweł Mateusz Nowak 

In analytical chemistry, the use of dedicated metrics for assessing greenness, whiteness and other “colours” of new methods is becoming very popular. However, does this entail an increase in the overall scientific value? In this article, I explain why the correct answer is “not always”. In fact, one can have an impression that the assessments made currently may deliver additional information that nicely complements analytical validation, but sometimes, it only creates unnecessary confusion. Is the vision of easy profit in the form of publishing a greenness-oriented article so tempting? Or maybe the reason is the lack of clear guidelines and appropriate education? Whatever the answer is, the situation should be changed. I am trying to remedy this situation by proposing the five general rules of a Good Evaluation Practice (GEP). Implementation of GEP may help reduce the existing mess, improve transparency, promote research quality, and facilitate the exchange of information between authors and readers. This will also benefit reviewers and editors, who will find it easier to verify the correctness of the evaluation process. Although the article has been written with analytical chemistry in mind, the proposed rules are general enough to be easily extrapolated to other chemical domains.

Received 4th February 2025,
Accepted 6th May 2025

DOI: 10.1039/d5gc00615e

rsc.li/greenchem

Green foundation

1. In analytical chemistry, the use of dedicated metrics for assessing greenness and whiteness of new methods is becoming very popular. However, does this entail an increase in the overall scientific value? In this article, I explain why the correct answer is “not always”. I am trying to remedy this situation by proposing the five general rules of a Good Evaluation Practice (GEP).
2. Although the article has been written with analytical chemistry in mind, the proposed rules are general enough to be easily extrapolated to other chemical domains.
3. Implementation of GEP may help reduce the existing mess, improve transparency, promote research quality, and facilitate the exchange of information between authors and readers. This will also benefit reviewers and editors, who will find it easier to verify the correctness of the evaluation process.

1. Introduction

In analytical chemistry, assessing new methods in terms of the Green Analytical Chemistry (GAC)^{1–5} concept has recently become a mainstream topic.⁶ The interest is also seen in White Analytical Chemistry (WAC),⁷ the idea inspired by the red–green–blue colour model where “whiteness” is understood as the best possible compromise between greenness and functionality represented by red and blue colours. Assessment aimed at GAC and WAC can be performed using a wide spectrum of available metrics, *e.g.*, National Environmental Methods Index (NEMI),⁸ Analytical Eco-Scale (AES),⁹ Green

Analytical Procedure Index (GAPI),^{10–12} Analytical GREENness Metric (AGREE),^{13–16} Analytical Method Greenness Score (AMGS),¹⁷ Analytical Method Volume Intensity (AMVI),¹⁸ Sample Preparation Metric of Sustainability (SPMS),¹⁹ Environmental Assessment Tool (EAT) for chromatographic methods,²⁰ Chloroform-oriented Toxicity Estimation Scale (ChlorTox Scale),²¹ Multicriteria Decision Analysis (MCDA),²² hexagon,²³ Red–Green–Blue (RGB) model,^{7,24,25} Blue Applicability Grade Index (BAGI),²⁶ and Red Analytical Performance Index (RAPI).²⁷ These and other metrics have been well described in many review articles.^{28–31} Greenness-oriented tools are aimed at providing additional information regarding environmental friendliness and safety of use, and in the case of “other colours” – regarding a holistic potential (whiteness), analytical performance (redness), and practical criteria (blueness). This additional insight into the character-

Department of Analytical Chemistry, Faculty of Chemistry, Jagiellonian University in Kraków, Gronostajowa 2, 30-387 Kraków, Poland. E-mail: pm.nowak@uj.edu.pl



istics of the method is undoubtedly a valuable addition to the article and can help readers to form their own opinion about the potential method applicability. Thus, this trend seems positive and worth following, and one may even suppose that evaluations in terms of GAC and WAC principles will soon become a generally accepted standard. Nevertheless, is such an idealistic picture of this phenomenon fully justified?

It is worth realizing that there are many reasons why conducting an assessment of the method (in the way it is currently commonly done in analytical journals) does not always lead to expanded information about the method and, therefore, to an increase in the quality of the article itself. Furthermore, sometimes one may have the impression that the greenness/whiteness assessment is supposed to be an "extra value gained at low cost", thanks to which the manuscript will become publishable in the eyes of the reviewers and the editor, thereby preventing potential critical comments regarding the lack of innovation and low added value.

The main causes of counter-productive assessments will be discussed below, together with the general guidelines called Good Evaluation Practice (GEP) and general tips on its implementation, which, in my opinion, can help promote quality and eliminate possible abuses. The proposed rules are so general that they can be used in any chemical subdomain, beyond analytical chemistry. My considerations are based on the knowledge and experience I have gained in recent years in this matter.^{7,21,24,25,27,32-35} This is also the continuation of my last theoretical work,³⁶ in which I presented an attempt at a multilateral description of "greenness" as a central concept and the various implications arising from it.

2. Should the assessment of greenness/whiteness be mandatory?

But first, let's ask ourselves whether the greenness/whiteness assessment of analytical methods should be mandatory or at least a generally recommended standard? In my opinion, the correct answer is: "it depends". Despite the great effort I have put into promoting the GAC and WAC concepts and varied metrics, I believe that the only mandatory step at the end of the method development process should be analytical validation, followed by demonstration of the method's use on real samples. Of course, making a reliable assessment taking into account ecological and practical aspects broadens the picture of the method, but it does not make the method more innovative. If the new method allows the analysis of samples that were previously unavailable to other methods or its analytical capabilities, *e.g.*, LOD, significantly exceed known methods, then the lack of use of assessment metrics focusing on parameters beyond analytical validation should not be treated as a significant shortcoming. In this case, the analytical merits are so strong that no one should have any doubts that the method is worth publishing.

Sometimes, however, the value of a new method may not be so evident, *e.g.*, when it is supposed to be an alternative to other well-performing procedures for the same analytes and samples. Then, the recommendation to assess and compare

the new method with known methods seems fully justified. In other cases, the assessment process may deliver the main results, *e.g.*, when writing review articles analysing "how one method compares to another" may be the basis for the authors to draw key conclusions. It is also worth mentioning that performing the assessment may be strongly advised in specific cases, *e.g.*, in journals focusing on greenness or for specific special issues.

To sum up, the assessment of greenness/whiteness cannot currently be treated as a general formal standard, such as analytical validation. Nevertheless, in most cases, it should be performed to advocate the merits of new analytical methods. Thus, the authors' own initiative to perform reliable assessment and evaluation of the new method they want to publish should always be welcome but not required in every possible case. Editors and reviewers should finally decide when this is necessary.

3. Good evaluation practice

The five general GEP rules will be presented below, along with justification and tips for their implementation.

3.1. Rule 1

Use quantitative indicators based on empirical data, and to ensure a more comprehensive picture, combine them with models with varied structures.

Currently, the most popular among available assessment metrics in analytical chemistry are models, such as various versions of GAPI¹⁰⁻¹² and AGREE,¹³⁻¹⁶ AES,⁹ RGB and RGB12 models,^{7,24} BAGI,²⁶ etc. Their basic limitation is that they combine many assessment criteria according to a certain pre-defined, arbitrary scheme, and they often require making far-reaching estimates and assumptions, which may introduce inaccuracies and leave room for abuse. An example of a criterion that often appears in models is the amount of electricity needed to analyse one sample (kWh per sample), the value of which (as I assume based on my own observations and experience) is almost never measured directly with appropriate meters monitoring the operation of individual instruments, but instead approximated more or less scrupulously. But how can we be sure that errors made during estimation will not ultimately affect the assessment result and will not change the indication of the best method? Another reason for inaccuracy is that many criteria are assessed discretely rather than linearly. To illustrate, an energy demand of 0.5 and 0.9 kWh per sample may be from the same assumed range of <1.0 kWh per sample and rated for the same number of points, while a value of 1.1 can already be rated worse, although 1.1 is closer to 0.9 than 0.9 to 0.5. Therefore, again, comprehensive assessment models, based on arbitrarily adopted assumptions, cannot be treated as oracles. The results obtained can constitute valuable support and extend the picture of the method's characteristics, but should be interpreted with great caution. In order to compensate for certain deviations and obtain a more reliable picture of the method, it is worth using different models sim-



ultaneously, differing in structure, scoring scheme and adopted assumptions.

Therefore, the question arises: "why do we so eagerly use models when assessing analytical methods, but so rarely simple, specific indicators based on directly measurable, empirical, quantitative data?" Indeed, the spectrum of potential indicators that are more direct and objective than models is very wide, let me propose a few:

(a) the amount of electricity required to carry out a certain large number of analyses, *e.g.*, 100 samples, including instrument preparation and calibration, measurement and ending procedure, obtained by measuring the amount of energy consumed by specific instruments using a wattmeter;

(b) carbon footprint resulting from carrying out a certain number of analyses, resulting from energy consumption (point a) and the emissivity of the energy used to power the method in a given place ($\text{gCO}_2 \text{ kW}^{-1} \text{h}^{-1}$);

(c) the total mass/volume of waste generated during a certain amount of analyses, including preparation of reagents, calibration of instruments, measurements and final procedures;

(d) the total mass/volume of reagents used that may be considered particularly hazardous;

(e) the total mass of solid waste subject to special disposal regulations, such as plastic, packaging and laboratory glassware that came in contact with chemicals;

(f) total volume of tap, distilled and ultrapure water used by the method;

(g) the total time needed to implement and apply a given procedure to a specific type and number of samples, taking into account prior optimization, calibration and validation of the method, expressed in units of time or man-hours (taking into account the number of analysts).

The fact is that we very rarely use these or similar indicators,^{18,20,34} however, if used properly, they could say a lot about the method's greenness and practicality. These parameters are directly measurable, quantitative (and therefore easily comparable) and transparent. Perhaps the reason why we do not prefer using them in analytics is that it would require changing habits and more effort from us? Or perhaps it lacks clear recommendations of using them in the literature? Regardless of the reason for this state of affairs, it is worth and even necessary to use quantitative indicators whenever it can be done in a reasonable and relatively easy way.

An example of good practice would be to permanently connect all electric devices to energy meters and monitor energy consumption as a standard daily procedure in the laboratory. This would allow us to calculate the amount of energy consumed by a method as a function of the energy demand of a given device and its operating time. Assessment performed in this way could be presented in a simple format, *e.g.*, a table, providing a deep and valuable insight into one of the key greenness elements – carbon footprint.³⁶ One should however note that electricity consumption is not the only source of greenhouse gas emissions. The perspective of including a carbon footprint related to the manufacturing of reagents, sol-

vents and disposal of waste from cradle to grave is for now difficult to implement but still worth developing as a future goal we should strive for.

An example of another noteworthy greenness indicator based on hard empirical data, of wide potential applicability, is the aforementioned ChlorTox scale,²¹ developed by me in cooperation with the authors of GAPI and AGREE (Justyna Płotka-Wasylka and Marek Tobiszewski). This indicator expresses the total risk associated with the use of chemical reagents, taking into account their mass and the hazards posed, calculated in relation to chloroform as a reference substance (eqn (1)).

$$\text{ChlorTox} = \frac{\text{CH}_{\text{sub}}}{\text{CH}_{\text{CHCl}_3}} \cdot m_{\text{sub}} \quad (1)$$

Where the ChlorTox value, expressed in the mass of chloroform (g), reflects a degree of chemical risk associated with the substance-of-interest, considering its properties (hazards) and the amount used. $\text{CH}_{\text{sub}}/\text{CH}_{\text{CHCl}_3}$ represents a relative chemical hazard (CH) of using the assessed substance in relation to chloroform, and m_{sub} is a mass of the substance-of-interest needed to apply the method.

Although it is based on some arbitrary assumption – the method of comparing chemical hazards is based on the classification used in safety data sheets (SDS), it is still fully quantitative and transparent. The only data needed for calculating ChlorTox values are quantities of specific reagents and data from SDS cards, which are easily accessible/measurable.³³

To sum up, whenever possible, the fundamental for assessment should be measurable, quantitative parameters based on empirical data, although (I realize) we are not yet used to them. In parallel, we should not abandon the use of overall greenness and whiteness assessment models, which can be a valuable support and help in carrying out the final evaluation. Many different models should be used simultaneously and selected in a way that they complement each other. The potential lack of consistency between models seems expectable given their diversity; thus it should not be decisive for final verdicts.

A certain example regarding the selection of specific assessment tools from the currently available spectrum is shown in Fig. 1. Due to different focuses and the level of complexity and operation scheme, they nicely complement each other and thus help perform high quality evaluation (this is just a proposition; my goal is not to impose the choice of specific metrics). Noticeably, Click Analytical Chemistry Index (CACI)³⁷ and RAPI²⁷ are recently introduced metrics that complement numerous models dedicated to greenness. RAPI is based on formal recommendations for the validation of analytical methods,^{38,39} allowing assessment of the analytical capabilities of the method depending on the target concentration range.

3.2. Rule 2

When assessing a new method, use reference methods addressed to the same analytical purpose for comparison. Start with assessment, then interpret the results and make an



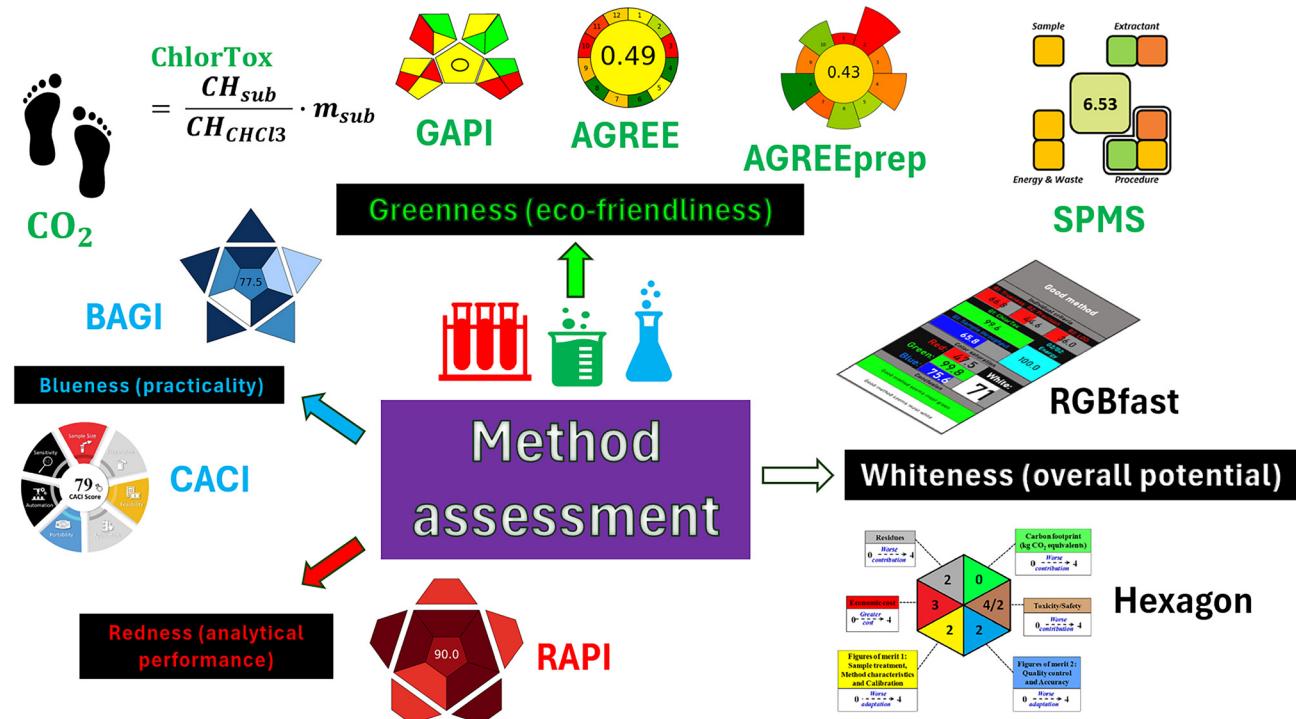


Fig. 1 An example of mutually complementary and currently available tools that can be used to assess methods in terms of different “colours” including both simple indicators/models based on empirical data (carbon footprint, ChlorTox, and RGBfast) and models based on more complex assumptions and guidelines (GAPI, AGREE, AGREEprep, SPMS, hexagon, BAGI, CACI, and RAPI).

evaluation taking into account the specificity of the intended application.

Indeed, assessment and evaluation are not synonymous and should not be used as interchangeable terms in analytical chemistry. Assessment is the process of obtaining certain qualitative or quantitative information, *e.g.*, in the form of a pictogram or number, which characterizes a given method from a certain narrow or wide angle. Thus, the assessment will include both the calculation of the total weight of waste produced and the use of a model that results in obtaining a coloured pictogram presenting the overall method's greenness. Evaluation is made on the basis of a previous assessment; it is a process of judging the suitability of a method in relation to specific expectations (stemming from assumed application), which requires an objective look at the assessment results and awareness of the limitations of the metrics used. An assessment not supported by appropriate interpretation referring to the specificity of the planned use of the method is of little value because the information for the reader is too superficial to be useful in a specific case. The ultimate goal should therefore be evaluation, carried out in an objective and critical manner, based on the assessment results, the evaluator's experience and knowledge of the specificity of the method's intended application.

An example of bad but quite frequent practice is the assessment described and interpreted using laconic statements like for example: “(...) after validation, the greenness of the method was assessed using the X model, (...) the resulting pictogram is

mostly green, thus confirming that the method is green (and thus deserves publication)”. In this example, it is a mistake to treat the assessment model as an oracle without any reflection. For instance, knowing what the target application of the method is and the resulting requirements, we could state that the model's guidelines are in this case appropriate, too lenient or too strict. Even if the green pictogram indicates high greenness according to the model, as a result of the evaluation we can come to the conclusion that it is insufficient (*e.g.*, due to the simple sample and high analyte concentration there is a large potential for reagent and energy saving), and *vice versa*.

Nothing makes the evaluation of a method easier than comparing the assessment results with other alternative methods used as generally accepted standards or with new, highly innovative methods that are just gaining popularity and which can soon constitute real competition. However, it is crucial that all compared methods have the same analytical purpose. Ideally, they would show full compliance regarding the analyte(s) and sample matrix. It is worth remembering that the same analyte, but present in different matrices, *e.g.*, water and urine, may pose different requirements to the analytical method. The matrix plays a major role in determining the sample preparation process, *e.g.*, the need for extraction, and therefore influences the techniques used and the final assessment of the entire procedure (method) – see the principles of green sample preparation.⁴⁰ In the absence of an appropriate method for comparison showing full compliance of analytes and matrix, one can consider using another method, with a lower degree



of compliance, but emphasizing precisely at the evaluation stage that the purpose of the compared methods is slightly different.

The comparison will also gain credibility if it is not limited to a single reference method. Looking idealistically, it would be best to combine as many different methods as possible (showing compatibility of analytes and matrix), both the older ones that are probably most frequently used and the newer ones that may constitute real competition for “our new method” in the near future. However, in practice this can be extremely difficult. The main obstacle is that the description of methods available in the analytical literature is in fact cursory and devoid of much important information to parameterize them comprehensively and sincerely assess. While the validation criteria (determining the redness of the method according to the WAC idea) are well described, the problem occurs in the case of the green and blue criteria. For example, it is difficult to imagine how the energy consumption of a method can be reliably estimated and expressed as kWh per sample, relying only on a “routine” article in which all instruments are listed, but without providing their average electrical power and normalized operation time. As a result, we are forced to rely on estimates and approximations. Therefore, although the use of reference methods in assessment is quite common today, the reliability of such comparisons is often questionable. Here are some suggestions on how to improve this state of affairs:

(a) ask for an assessment of the reference method directly from the authors, indicating the purpose for which you want to use the data and offering in return a personal acknowledgement or, if cooperation develops, co-authorship of the article;

(b) implement the reference method in your own laboratory (depending on the availability of equipment and infrastructure), re-validate it, and then assess;

(c) ask your friends to evaluate all the methods (increase the number of evaluators), let everyone rate the methods according to their knowledge, and give the average values as the main result;

(d) develop and use an artificial intelligence (AI)-based model as a dedicated support.

The growing capabilities of AI, in particular Large Language Models (LLMs), which can effectively analyse, collect and process data contained in scientific texts, seem to provide real hope for the future. Perhaps a properly trained LLM would be able to reliably estimate missing parameters in the indicated method description, with acceptable accuracy, if trained on a sufficiently large and representative body of literature and empirical data. However, creating a training set for AI able to simulate lacking data would require many coordinated experiments, taking into account the variety of currently available methodologies. Nevertheless, this is a new interesting research path worth following.

To finish this fragment, I feel it is also necessary to strongly condemn a practice that may seem tempting, but is in fact unethical. It is unacceptable to compare a new method with another, called the “reference” method, in such a way that, not having access to appropriate data about the reference method,

we consciously (or subconsciously) make certain unjustified assumptions that result in its worse assessment. In other words, assessment and evaluation of our own methods should be done with the same rigor and meticulousness as those of others, even if the final result puts our methods in a bad light.

3.3. Rule 3

Report all data used in the evaluation, describe and justify all estimates and assumptions. The assessment results should be reproducible based on the data provided.

This rule is an extension of the previous discussion. Regardless of whether we are evaluating a new or old method, ours or someone else's, we are obliged to provide all data used in the evaluation and justify the assumptions made. Of course, the space available in the main text of the manuscript may not be sufficient to include and describe all the data, but we can always place it in a supplement or a publicly available database to which we will provide a link. An example would be the data and assumptions based on which we concluded that: “the energy consumption of our method is X kWh per sample, and the production of hazardous waste is Y mL per sample, etc.”.

It is worth emphasizing that article reviewers should always have the possibility to verify whether the assessment using specific metrics was carried out “according to the art” and whether the assumptions made are rational. From my experience as an editor of greenness-related journal (Green Analytical Chemistry, Elsevier),³⁵ I can say that reviewers very rarely, perhaps even too rarely, question the assessment process, the quality of the input data and the correctness of the application of specific metrics. Considering the currently observed situation, meeting this rule seems to be an urgent need.

There is a way to easily verify whether the submitted description contains all the necessary information. We need to honestly ask ourselves a simple question: “will a reviewer who wants to verify the correctness of our assessment based on the data and information we provide receive identical results?” In other words: “are the assessment results fully reproducible?” If we answer “yes” or “very likely yes”, it means that the rule is met.

At the same time, it should be remembered that while the assessment should be fundamentally reproducible, evaluation is by definition something more subjective and discretionary, just like the conclusions we draw based on the obtained research results. For some, the results of the greenness/whiteness assessment in relation to the planned purpose of the method will seem “good”, for others “very good” for yet others “only acceptable”, such diversity of opinions is normal. However, in some situations, *e.g.*, in the decision-making process regarding the selection of a method, evaluation results must be subject to some kind of compromise. The key may be, again, to conduct a survey among a sufficiently large group of specialists with appropriate knowledge and experience, without conflicts of interest, or elaborate some additional framework referring to the previous assessment process. However, this topic is out of the scope of this work.



3.4. Rule 4

Use concise, clear and critical language.

The idea of this rule can be expressed through a question that often arises after reading articles describing the greenness assessment of a new method: “so, you claim that your method is green, but what does it mean?” Unfortunately, the increased interest in GAC and WAC often results in the overuse or accidental use of terms such as a “green method”, a “white method”, and a “sustainable method”. In a recently published article, I faced the problem of the theoretical description of greenness as a central concept of green chemistry,³⁶ among other things, suggesting what greenness actually should mean. For the purpose of this work, I will only present the most important conclusions from this theoretical study.

Greenness may mean the degree of colour saturation, *i.e.*, an attribute of the method indicating its potential adverse impact on the environment and the user, which can be quantified and compared between methods (this is exactly what the greenness assessment is being made for). Secondly, greenness may also mean “the state of being green”, *i.e.*, a certain zero-one property. When the method achieves this state, it can be referred to as “green” in the general sense. The situation becomes more complicated when we realize that the state of being green can be defined in various ways. I proposed three interpretations: purist, pragmatic and formal.³⁶ At the same time, I did not indicate which of them is the best, because in my opinion, the most important thing is to be consistent and not to mix them.

In short, a purist greenness is one that is unattainable in practice, *i.e.*, no method can be considered green since each method has some, even negligible, unfavourable impact, expressed *e.g.*, in carbon footprint. Pragmatic greenness is one that is, by definition, relative, it is stated in relation to another object, so method A may be and may not be green at the same time in relation to two different methods (B and C). Formal greenness is one that can be determined in relation to more or less formal top-adopted standards, which, in fact, are currently lacking (however, they can be developed for the needs of a specific comparison).

In order not to leave readers without further practical instructions, I propose that, for maximum purity and clarity of language, the purist interpretation should be assumed by default. In addition, I suggest not using statements such as “green method” or “green procedure” at all, because without knowledge of the adopted interpretation, such terms do not provide any valuable information and are redundant. In particular, I suggest avoiding such terms in article titles, because they may even discourage people from reading (especially those sceptical about green chemistry as a distinct discipline). However, one can still use “greenness” as a parameter, *e.g.*, “according to model X, method A seems greener than B” or “model X rated the greenness of method A (colour saturation) at 80% and method B at 60%”, *etc.* Similar conclusions also apply to whiteness and other colours. We can also state something more specific, *e.g.*, in the final conclusions, that “it was found that the greenness/whiteness of method X is (or is not) adequate to the planned application”, or that “method X

seems to meet the GAC/WAC assumptions sufficiently/insufficiently regarding its planned use”, *etc.*

It is also worth mentioning that another problematic term is “sustainable method/procedure”. The point here is that there is no complete agreement among chemists (and even people in general) on what “sustainable” actually means. It is certainly incorrect to use it as a synonym for green, because its meaning should be broader. A “sustainable method” should, therefore, mean either one that meets the 17 Sustainable Development Goals – here the problem is how to assess the method with reference to specific goals, especially since many of them are not related to chemistry at all,^{41,42} or one that combines care for the environment with social (accessibility for people) and economic aspects – this is a quite common interpretation in analytical chemistry.^{43,44} Overall, this term seems actually vaguer than the previously mentioned greenness, and it is difficult to indicate exactly what these additional aspects are expressed in and whether they include the analytical effectiveness tested at the validation stage or not. My personal suggestion, which I made already when developing the WAC concept,⁷ was to use whiteness as a term that not only combines greenness and analytical criteria (redness), but also socio-economic criteria (blueness). However, sometimes it is not necessary to consider all three colours in decision-making. For instance, improving analytical criteria much beyond the threshold of acceptability does not always bring real benefits to the user. In such cases decision key could be restricted to green and blue criteria. That is why I have recently been preparing a new concept that defines sustainability as a junction of greenness and blueness, quantified with a simple indicator (S-factor) - whose determination will require the implementation of GEP. However, this will be the subject of another article.

Finally, let me say a few words about criticism of the description of the evaluation process. In general, it gives a bad impression to use strong statements that suggest our certainty when we are describing something questionable, subjective and relative. An example would be the interpretation of results obtained using models such as GAPI, AGREE, AES, RGB, BAGI, or RAPI. Statements such as: “the results obtained using model X CONFIRM that method A is better than B” or “the obtained pictograms PROVE that the new method is more environmentally friendly than B” are exaggerated. In such situations, one should accept the imperfection and inherent subjectivity of the models used and stop at weaker statements, like “the evaluation results SUGGEST/INDICATE that method A MAY BE less harmful/better than B” or “the greenness/whiteness of method A SEEMS higher than method B ACCORDING TO model X”, *etc.* It is always better for “analytical trueness” to underestimate than to overstate the expected merit.

3.5. Rule 5

Do not limit evaluation to the retrospective approach (*ex-post*). Perform *ex-ante* evaluation to justify the need for a new method and better plan the development phase.

The last of the proposed rules no longer applies to errors or bad practices that may occur in the evaluation process itself.



The purpose of this rule is to point out that there are actually two appropriate moments to evaluate a new method, one of which is virtually unnoticed. We are used to the evaluation being carried out retrospectively, *i.e.*, *ex-post*, in relation to a previously developed, optimized and validated procedure. This is what the previously mentioned metrics are designed for. However, we must be aware that if our goal is to obtain the best possible greenness/whiteness score at the end, by limiting ourselves only to a retrospective approach, we reduce the chance of getting a good result. In fact, greenness/whiteness assessment should both begin and finish the method development process (see Fig. 2). As soon as we come up with the idea of developing a new method based on specific techniques and methodologies, we should conduct a prospective *ex-ante* evaluation. Its aim is to simulate the potential advantages and disadvantages of the method we want to develop (expected greenness and whiteness), based on our knowledge, experience and available information in the literature. Notably, when we do it properly, additional benefits will emerge.

First, *ex-ante* evaluation can provide a solid justification for decisions regarding method development, which often entails efforts and investment of valuable resources. When our goal is to develop an alternative to a standard procedure that is both greener and whiter (indeed a laudable motivation), and the result of the *ex-ante* evaluation shows that we should expect rather a worse characteristic in the key criteria, it suggests a waste of time and resources. When the simulation shows that there are a few relatively small obstacles, it may be a green light for our idea.

Moreover, reporting the *ex-ante* evaluation in the text of a scientific article describing the development of a new method will be an expression of our experience and knowledge, but also rationality, which, I assume, will be appreciated by the vast majority of reviewers. Indeed, starting method development solely because: "such a method has not existed yet" implies the authors care more about paper publication than scientific quality.

In addition to making decisions, outcomes of *ex-ante* assessments may imply an experimental plan, the range of variables that will be optimized, and bottlenecks that should

be given the most attention; in other words, they help to better plan the next steps. The experimental design found in publications often seems too narrow or random, which indicates that the authors have closed themselves off from focusing on significant shortcomings in advance or enhancing the advantages offered by the techniques and instruments used.

So the question arises: "how to assess *ex-ante*?" In general, previous rules regarding *ex-post* evaluation, such as the need to use objective and diverse metrics or accurate reporting and description of the evaluation process, do not apply strictly in this case. The *ex-ante* assessment may be based on estimates, assumptions and hypotheses, as long as they are reasonable and supported by available knowledge. By definition, the guidelines should be less restrictive than for the final assessment, and the format of the tools used should be more flexible. Such a type of tool is the RGB_ *ex-ante* model,⁴⁵ which I recently used to evaluate a new potential tandem technique, microscale thermophoresis (MST) combined in a stop-flow format with mass spectrometry (MST-MS). The *ex-ante* evaluation showed that the potential applicability of this technique is quite limited, and the expected benefits, in particular regarding greenness, seem to compensate to a small extent for the significant limitations in the red and blue areas (see Fig. 3).

However, it is not necessary at all to use a specific model for *ex-ante* evaluation, and it can be done in a simple descriptive way. It is sufficient to describe our assumptions and predictions precisely and clearly and refer to the literature whenever possible. Assessing greenness and whiteness in a quantitative way should be optional, depending on how much emphasis we want to put on the evaluation process.

Combining and confronting the results of *ex-ante* and *ex-post* evaluations may be a source of valuable knowledge regarding the correctness of the adopted assumptions and even refute stereotypes regarding techniques and their alleged intrinsic advantages and disadvantages. The publication of the assessment results obtained at both stages, along with a commentary explaining the observed correctness of the predictions or its lack, may make the subsequent authors use this knowledge, and their *ex-ante* assessment will be more accurate. By propagating such collaborative efforts in the research community, the chances for developing the best possible methods increase. Noteworthy, *ex-ante* evaluation has already been combined with *ex-post* evaluation by other research groups; for this purpose a modified version of the older RGB model was used.^{46,47}

Moreover, it is worth mentioning again the benefits that the use of AI, especially LLMs, can bring in the future. It is easy to imagine the use of a specially developed Chatbot that would have extensive and constantly updated knowledge in the field of analytical chemistry journals, which would be trained to perform *ex-ante* analysis of the analytical procedures we plan to develop. It is worth mentioning that even now, the use of AI for the purpose of collecting and processing data and knowledge contained in articles is fully possible and, notably, does not violate general ethical principles. Indeed, how can we



Fig. 2 Subsequent stages of the process of developing a new analytical method, indicating the moments when prospective (*ex-ante*) and retrospective (*ex-post*) evaluations should be carried out according to the GEP rules.



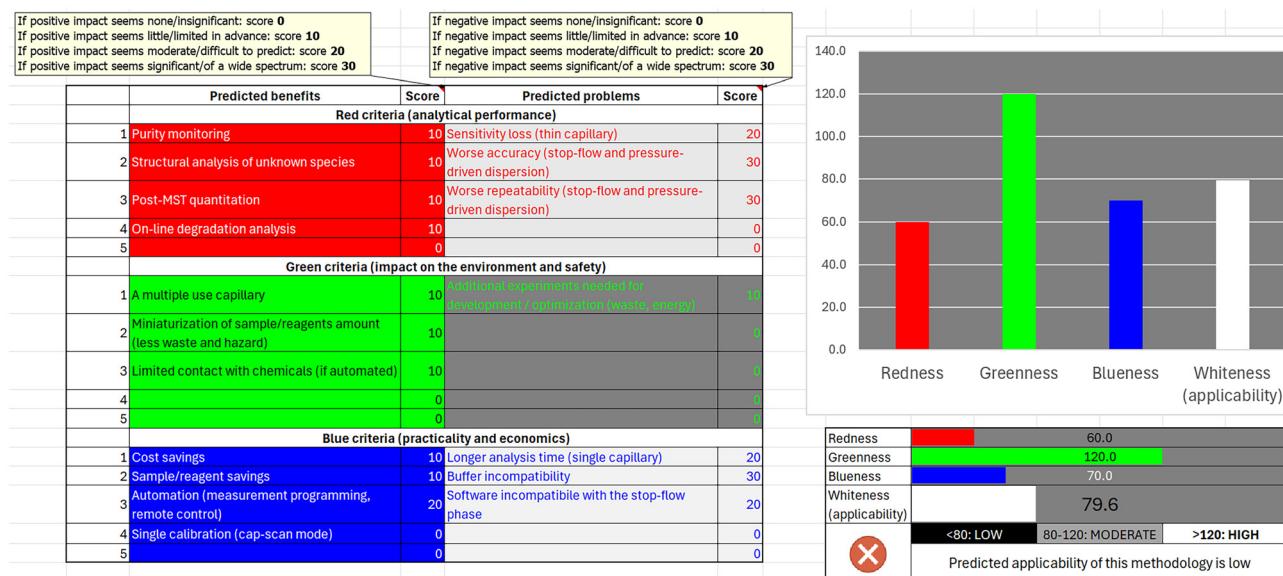


Fig. 3 Results of the prospective evaluation of the new analytical technique MST-MS, obtained using the RGB_ex-ante model, the details of which are available in the original paper. Reproduced from ref. 45 published under a CC-BY licence.

accuse someone of making the decision to develop a new method based on the information he received from AI? What was the impulse to make the decision is irrelevant from the point of view of the ready-to-use method and its confirmed capabilities. So, let's make our lives easier and explore opportunities offered by the currently available computer technologies in this aspect.

4. GEP protocol

To facilitate the implementation of the above rules in a routine practice, I have developed a simple questionnaire called the “GEP protocol”, which is presented in Table 1. It consists of three columns: the first contains the GEP rules, the second contains suggested questions to ask oneself – they can be optimized by keeping the correlation with the particular rules in mind, whereas the third column is supposed to be completed each time by authors of a new method subject to evaluation and present concise answers to these questions. Affirmative answers, if they agree with the facts and are justified appropriately, will confirm the implementation of GEP. The GEP protocol may be included in publications directly, or in a supplement, and provide evidence of GEP for reviewers and readers.

5. Perspective for general chemistry

Beyond analytical chemistry, there is also a pursuit of greenness and sustainable development, but in my humble opinion, it is generally more structured, and the quality of information presented in articles is higher. The advantage of synthetic chemistry is the popular use of simple empirical indicators,

such as E-factor, atom economy, mass intensity, carbon efficiency or effective mass yield,⁴⁸ which are in fact very simple to compare and interpret. Their accuracy in describing new methods and processes is however limited, and here, simple models such as Eco-scale⁴⁹ or Green Star (GC)⁵⁰ come to the rescue. The deepest insight into the ecological effects of a chemical process can be provided by a reliably conducted Life Cycle Assessment (LCA).⁵¹ Nevertheless, if we are assessing a synthesis method that is just a certain recipe and not a specific process carried out in a specific place and time (analogously to analytical chemistry methods), LCA may be difficult to perform due to the lack of data specifying, for example, the origin of reagents or waste disposal. Laboratories can differ in these aspects, although they use the same method, *i.e.*, a recipe given in the literature. The choice between LCA and simpler models should depend on what we actually want to assess (see more discussion in my last theoretical paper³⁶).

In general chemistry, evaluation of “other colours” has not yet become the mainstream. There are not many models that cover parameters defining the reaction yield, product purity, time consumption, cost, and greenness. That is why I recently started cooperation with Prof. Zajdel's group specializing in mechanochemistry. Together, we developed a version of the RGB model dedicated to the evaluation of synthesis methods – RGBsynt,³² which allows for comparing and selecting optimal synthesis procedures (the whitest ones) based on easily accessible quantitative data representing three different colours. However, there is still room for creating new models dedicated to colours other than green.

As for specific GEP rules, each of them can be successfully applied in synthetic chemistry. The selection of metrics, the use of reference methods, a full description of the evaluation process and the data used, and the use of a more critical and



Table 1 GEP protocol confirming the implementation of individual rules in a specific case

GEP rules	Questions to ask oneself	Example of answer ^a
1. Use quantitative indicators based on empirical data, and to ensure a more comprehensive picture, combine them with models with varied structures.	Have I made measurements/objective estimates of the amount of energy and reagents used by the method? Have I used (for the assessment) any other quantitative indicators based on empirical data? Have I used greenness/whiteness assessment models based on differentiated structures that complement each other?	I have calculated the exact amounts of reagents and used them to apply the ChlorTox scale. ²¹ I have estimated the amount of energy based on the average power of the devices provided by the manufacturer and their operating time. I have used these data to estimate the carbon footprint of the method related to energy consumption. To further assess greenness, I have used complementary models: ComplexGAPI, ¹¹ AGREE, ¹³ and AGREEprep. ¹⁴ To assess whiteness, I have used RGBfast, ²⁵ complemented by BAGI (focus on blueness), ²⁶ and RAPI (focus on redness). ²⁷
2. When assessing a new method, use reference methods addressed to the same analytical purpose for comparison. Start with assessment, then interpret the results and make evaluation taking into account the specificity of the intended application.	Have I compared the new method with reference methods aimed at the same analytes and sample matrix? Have I selected enough reference methods that constitute viable alternatives to the proposed one? Have I made a critical evaluation by interpreting the assessment results taking into account the specificity of the planned application?	I have made comparisons with two reference methods of the same intended use, the older, routinely used method A and the newer method B with high analytical value. I have made an evaluation indicating high expectations regarding trueness and precision (forensic analysis). The proposed method provides analytical values comparable to method A and inferior to method B, but still acceptable. The advantages of the new method are green and blue criteria, which are confirmed by every metric used. The new method seems fit-for-purpose and highly competitive.
3. Report all data used in the evaluation, describe and justify all estimates and assumptions. The assessment results should be reproducible based on the data provided.	Have the results of the assessment using different metrics turned out consistent? What is the main evaluation outcome? Where have I provided all the data and explanation of assumptions used in the evaluation? Have I described them clearly? How have I estimated the data needed to evaluate the reference methods? Have I asked someone to reproduce the assessment with the same data available? (If yes) Have the consistent results been obtained and has a similar evaluation been made?	I have provided all information in the ESI.† I have shown all numerical data in tables and I described the assumptions in detail in the text. Data on reference methods came from literature descriptions; some of the data on greenness were unavailable, so I have estimated it to the best of my knowledge. I asked another specialist (who did not participate in the method development) to reproduce the assessment based only on the data provided in the ESI.† Consistent results have been obtained and similar conclusions have been drawn.
4. Use concise, clear and critical language.	Have I limited the use of empty buzzwords in the evaluation description (e.g., "green method", "sustainable method")? Is the meaning of colour concepts clear to the reader (terms such as "greenness", "whiteness", etc.)? Have I avoided overinterpreting assessment results and demonstrated an appropriate level of self-criticism?	With regard to colours, I have adopted a purist interpretation, ³⁶ eliminating redundant terms. I consider greenness and whiteness as relative parameters, which are expressed by the metrics used. I treat the assessment results as a premise, not evidence. I have evaluated the new and reference methods with the same level of criticism.
5. Do not limit evaluation to the retrospective approach (<i>ex-post</i>). Perform <i>ex-ante</i> evaluation to justify the need for a new method and better plan the development phase.	Did I start the process of developing a new method with a prospective <i>ex-ante</i> evaluation? How did I perform it? Have the obtained results indicated a high added value of the new method, and if so, what exactly? Have I used these results when planning experiments? Have I analysed the consistency of the <i>ex-ante</i> and <i>ex-post</i> evaluations? Can the data presented be helpful to others in future evaluations of similar methods?	I started with making the <i>ex-ante</i> evaluation using the RGB- <i>ex-ante</i> template. ⁴⁵ The obtained outcomes pointed out that one can expect little reagent and energy consumption and, thereby, high potential greenness, maintaining comparable practicality and slightly worse analytical parameters to reference methods. I have used these data to focus on optimizing analytical performance during method development. <i>Ex-post</i> and <i>ex-ante</i> evaluations have occurred to differ to small extent. Analytical performance improved slightly as a result of optimization. Green criteria were correctly predicted. The practical criteria have turned out even better than assumed. These outcomes can be further used by others.

^a These are exemplary answers that should be given anew each time by the method evaluators.

precise language are universal issues that require refinement by both analysts and synthetic engineers. Notably, I see the possibility of implementing the last rule, *i.e.*, prospective evaluation, as particularly interesting from the point of view of synthetic chemistry. Its implementation can take place by integrating known and new assessment metrics with computer synthesis planning algorithms,^{52–54} allowing for an *ex-ante* evaluation of new potential routes leading to the target. Computer synthesis planning creates enormous opportunities from the point of view of green chemistry,⁵⁵ so it is worth ensuring that the ultimately selected paths are efficient, cheap and environmentally friendly. A separate issue is that developing computer programs that require the use of energy-intensive technologies, like deep learning algorithms, in itself generates a significant carbon footprint. Therefore, GEP should go beyond synthetic chemistry and include computational chemistry as well.

6. Conclusions and future outlook

The evaluation of new methods in terms of GAC and WAC is becoming more and more common, which is undoubtedly a positive phenomenon, but, as has been shown, it can be carried out incorrectly and therefore become counter-productive. Therefore, I have presented a proposal of several general rules, the observance of which may increase the quality of published articles in the field of analytical chemistry and beyond. I am aware that full implementation of GEP on a large scale will require both the willingness of authors and their time. For now, any change towards GEP should be appreciated, even if not all rules can be met. The potential advantage is huge: authors will benefit – they will be able to assess and evaluate new methods according to transparent guidelines and thus introduce additional value to their publications; reviewers and editors – they will find it easier to verify the correctness and reliability of the evaluation process; and readers – they will find it easier to understand the article and verify the true potential of the presented method. However, the most important thing is that science itself will benefit because the implementation of GEP will promote quality, transparency, and the ideas of GAC and WAC in a way that may attract current sceptics. The GEP protocol can be used as a simple source of information on whether GEP rules are met comprehensively. Raising awareness of good practice at an early stage of education is also important and should go hand in hand with increased requirements placed on article authors and greater involvement of reviewers in verifying the reliability of the assessment. An example of a simple activity expected from article reviewers should be to figure out whether the assessment of reference methods (which are often evaluated worse than the new method advertised by the authors) actually results directly from the attached data.

In the near future, in addition to the implementation of GEP, it seems crucial to reliably recognize the possibilities that AI technology can bring in the context of evaluating new chemical methods. Properly trained algorithms under the

supervision of experts can become extremely helpful at various stages of evaluation. Therefore, combining traditional method evaluation according to GEP with various forms of using AI should be welcome and thus can allow for gaining the missing “know-how”. Reporting such Artificial Intelligence-assisted Method Assessment Protocols (AIMAPs) can be a valuable addition to publications presenting new methods of analysis or synthesis. Sharing these types of data is crucial from the point of view of quick scientific progress.

Data availability

No data other than those included in the text were used to create the manuscript.

Conflicts of interest

The author declares that he has no conflicts of interest.

Acknowledgements

This research was funded by the National Science Centre, Poland (P. M. N.; OPUS, 2020–2024, grant no. 2019/35/B/ST4/01022).

References

- 1 P. T. Anastas, *Crit. Rev. Anal. Chem.*, 1999, **29**, 167–175, DOI: [10.1080/10408349891199356](https://doi.org/10.1080/10408349891199356).
- 2 M. Koel and M. Kaljurand, *Pure Appl. Chem.*, 2006, **78**, 1993–2002, DOI: [10.1351/pac200678111993](https://doi.org/10.1351/pac200678111993).
- 3 S. Armenta, S. Garrigues and M. de la Guardia, *TrAC, Trends Anal. Chem.*, 2008, **27**, 497–511, DOI: [10.1016/j.trac.2008.05.003](https://doi.org/10.1016/j.trac.2008.05.003).
- 4 A. Gałuszka, Z. Migaszewski and J. Namieśnik, *TrAC, Trends Anal. Chem.*, 2013, **50**, 78–84, DOI: [10.1016/j.trac.2013.04.010](https://doi.org/10.1016/j.trac.2013.04.010).
- 5 L. H. Keith, L. U. Gron and J. L. Young, *Chem. Rev.*, 2007, **107**, 2695–2708, DOI: [10.1021/cr068359e](https://doi.org/10.1021/cr068359e).
- 6 M. Koel and M. Kaljurand, *Curr. Opin. Green Sustainable Chem.*, 2021, **31**, 100541, DOI: [10.1016/j.cogsc.2021.100541](https://doi.org/10.1016/j.cogsc.2021.100541).
- 7 P. M. Nowak, R. Wietecha-Posłuszny and J. Pawliszyn, *TrAC, Trends Anal. Chem.*, 2021, **138**, 116223, DOI: [10.1016/j.trac.2021.116223](https://doi.org/10.1016/j.trac.2021.116223).
- 8 L. H. Keith, H. J. Brass, D. J. Sullivan, J. A. Boiani and K. T. Alben, *Environ. Sci. Technol.*, 2005, **39**, 173A–176A, DOI: [10.1021/es0532411](https://doi.org/10.1021/es0532411).
- 9 A. Gałuszka, Z. M. Migaszewski, P. Konieczka and J. Namieśnik, *TrAC, Trends Anal. Chem.*, 2012, **37**, 61–72, DOI: [10.1016/j.trac.2012.03.013](https://doi.org/10.1016/j.trac.2012.03.013).
- 10 J. Płotka-Wasylka, *Talanta*, 2018, **181**, 204–209, DOI: [10.1016/j.talanta.2018.01.013](https://doi.org/10.1016/j.talanta.2018.01.013).

11 J. Płotka-Wasylka and W. Wojnowski, *Green Chem.*, 2021, **23**, 8657, DOI: [10.1039/d1gc02318g](https://doi.org/10.1039/d1gc02318g).

12 F. R. Mansour, K. M. Omer and J. Płotka-Wasylka, *Green Anal. Chem.*, 2024, **10**, 100126, DOI: [10.1016/j.greeac.2024.100126](https://doi.org/10.1016/j.greeac.2024.100126).

13 F. Pena-Pereira, W. Wojnowski and M. Tobiszewski, *Anal. Chem.*, 2020, **92**, 10076–10082, DOI: [10.1021/acs.analchem.0c01887](https://doi.org/10.1021/acs.analchem.0c01887).

14 W. Wojnowski, M. Tobiszewski, F. Pena-Pereira and E. Psillakis, *TrAC, Trends Anal. Chem.*, 2022, **149**, 116553, DOI: [10.1016/j.trac.2022.116553](https://doi.org/10.1016/j.trac.2022.116553).

15 F. Pena-Pereira, M. Tobiszewski, W. Wojnowski and E. Psillakis, *Adv. Sample Prep.*, 2022, **3**, 100025, DOI: [10.1016/j.sampre.2022.100025](https://doi.org/10.1016/j.sampre.2022.100025).

16 M. Marć, W. Wojnowski, F. Pena-Pereira, M. Tobiszewski and A. Martín-Esteban, *ACS Sustainable Chem. Eng.*, 2024, **12**, 12516–12524, DOI: [10.1021/acsuschemeng.4c03874](https://doi.org/10.1021/acsuschemeng.4c03874).

17 M. B. Hicks, W. Farrell, Ch. Aurigemma, L. Lehmann and L. Weisel, *Green Chem.*, 2019, **21**, 1816–1826, DOI: [10.1039/c8gc03875a](https://doi.org/10.1039/c8gc03875a).

18 R. Hartman, R. Helmy, M. Al-Sayah and C. J. Welch, *Green Chem.*, 2011, **13**, 934–939, DOI: [10.1039/c0ge00524j](https://doi.org/10.1039/c0ge00524j).

19 R. Gonzalez-Martín, A. Gutierrez-Serpa and V. Pino, *J. Chromatogr. A*, 2023, **1707**, 464291, DOI: [10.1016/j.chroma.2023.464291](https://doi.org/10.1016/j.chroma.2023.464291).

20 Y. Gaber, U. Tornvall, M. A. Kumar, M. Ali Amin and R. Hatti-Kaul, *Green Chem.*, 2011, **13**, 2021–2025, DOI: [10.1039/c0gc00667j](https://doi.org/10.1039/c0gc00667j).

21 P. M. Nowak, R. Wietecha-Posłuszny, J. Płotka-Wasylka and M. Tobiszewski, *Green Anal. Chem.*, 2023, **5**, 100056, DOI: [10.1016/j.greeac.2023.100056](https://doi.org/10.1016/j.greeac.2023.100056).

22 R. Jędrkiewicz, S. Tsakovski, A. Lavenu, J. Namieśnik and M. Tobiszewski, *Talanta*, 2018, **178**, 928–933, DOI: [10.1016/j.talanta.2017.10.044](https://doi.org/10.1016/j.talanta.2017.10.044).

23 A. Ballester-Caudet, P. Campíns-Falco, B. Perez, R. Sancho, M. Lorente, G. Sastre and C. Gonzalez, *TrAC, Trends Anal. Chem.*, 2019, **118**, 538–547, DOI: [10.1016/j.trac.2019.06.015](https://doi.org/10.1016/j.trac.2019.06.015).

24 P. M. Nowak and P. Kościelniak, *Anal. Chem.*, 2019, **91**, 10343–10352, DOI: [10.1021/acs.analchem.9b01872](https://doi.org/10.1021/acs.analchem.9b01872).

25 P. M. Nowak and F. Arduini, *Green Anal. Chem.*, 2024, **10**, 100120, DOI: [10.1016/j.greeac.2024.100120](https://doi.org/10.1016/j.greeac.2024.100120).

26 N. Manousi, W. Wojnowski, J. Płotka-Wasylka and V. Samanidou, *Green Chem.*, 2023, **25**, 7598–7604, DOI: [10.1039/D3GC02347H](https://doi.org/10.1039/D3GC02347H).

27 P. M. Nowak, W. Wojnowski, N. Manousi, V. Samanidou and J. Płotka-Wasylka, *Green Chem.*, 2025, **27**, 5546–5553, DOI: [10.1039/D4GC05298F](https://doi.org/10.1039/D4GC05298F).

28 M. Sajid and J. Płotka-Wasylka, *Talanta*, 2022, **238**, 123046, DOI: [10.1016/j.talanta.2021.123046](https://doi.org/10.1016/j.talanta.2021.123046).

29 A. Y. Shishova and O. B. Mokhodoeva, *J. Anal. Chem.*, 2024, **79**, 487–499, DOI: [10.1134/S1061934824050125](https://doi.org/10.1134/S1061934824050125).

30 F. A. Semysim, B. K. Hussain, M. A. Hussien, E. A. Azooz and D. Snigur, *Crit. Rev. Anal. Chem.*, 2024, 1–14, DOI: [10.1080/10408347.2024.2304552](https://doi.org/10.1080/10408347.2024.2304552).

31 L. Yin, L. Yu, Y. Guo, Ch. Wang, Y. Ge, X. Zheng, N. Zhang, J. You, Y. Zhang and M. Shi, *J. Pharm. Anal.*, 2024, **14**, 101013, DOI: [10.1016/j.jpha.2024.101013](https://doi.org/10.1016/j.jpha.2024.101013).

32 P. M. Nowak, M. Kamiński, W. Trybała, V. Canale and P. Zajdel, *Green Chem.*, 2025, **27**, 1102–1112, DOI: [10.1039/D4GC05097E](https://doi.org/10.1039/D4GC05097E).

33 P. M. Nowak, A. Bis and A. Zima, *Green Anal. Chem.*, 2023, **6**, 100065, DOI: [10.1016/j.greeac.2023.100065](https://doi.org/10.1016/j.greeac.2023.100065).

34 P. M. Nowak, A. Bis, M. Rusin and M. Woźniakiewicz, *Green Anal. Chem.*, 2023, **4**, 100051, DOI: [10.1016/j.greeac.2023.100051](https://doi.org/10.1016/j.greeac.2023.100051).

35 J. Pawliszyn, D. Barceló, F. Arduini, L. Mondello, Z. Ouyang, P. M. Nowak and R. Wietecha-Posłuszny, *Green Anal. Chem.*, 2022, **1**, 100001, DOI: [10.1016/j.greeac.2022.100001](https://doi.org/10.1016/j.greeac.2022.100001).

36 P. M. Nowak, *Green Chem.*, 2023, **25**, 4625–4640, DOI: [10.1039/D3GC00800B](https://doi.org/10.1039/D3GC00800B).

37 F. R. Mansour, A. Bedair and M. Locatelli, *Adv. Sample Prep.*, 2025, **14**, 100164, DOI: [10.1016/j.sampre.2025.100164](https://doi.org/10.1016/j.sampre.2025.100164).

38 *Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics*, ed. B. Magnusson and U. Örnemark, 2nd edn, 2014, ISBN 978-91-87461-59-0, available at <https://www.eurachem.org>.

39 I. Taverniers, M. De Loose and E. Van Bockstaele, *TrAC, Trends Anal. Chem.*, 2004, **23**, 535–552, DOI: [10.1016/j.trac.2004.04.001](https://doi.org/10.1016/j.trac.2004.04.001).

40 Á. I. López-Lorente, F. Pena-Pereira, S. Pedersen-Bjergaard, V. G. Zuin, S. A. Ozkan and E. Psillakis, *TrAC, Trends Anal. Chem.*, 2022, **148**, 116530, DOI: [10.1016/j.trac.2022.116530](https://doi.org/10.1016/j.trac.2022.116530).

41 P. T. Anastas and J. B. Zimmerman, *Curr. Opin. Green Sustainable Chem.*, 2018, **13**, 150–153, DOI: [10.1016/j.cogsc.2018.04.017](https://doi.org/10.1016/j.cogsc.2018.04.017).

42 H. L. Moore, *J. Int. Dev.*, 2015, **27**, 801–815, DOI: [10.1002/jid.3114](https://doi.org/10.1002/jid.3114).

43 R. Marcinkowska, J. Namieśnik and M. Tobiszewski, *Curr. Opin. Green Sustainable Chem.*, 2019, **19**, 19–23, DOI: [10.1016/j.cogsc.2019.04.003](https://doi.org/10.1016/j.cogsc.2019.04.003).

44 J. Płotka-Wasylka, H. M. Mohamed, A. Kurowska-Susdorf, R. Dewani, M. Y. Fares and V. Andruch, *Curr. Opin. Green Sustainable Chem.*, 2021, **31**, 100508, DOI: [10.1016/j.cogsc.2021.100508](https://doi.org/10.1016/j.cogsc.2021.100508).

45 P. M. Nowak, A. Zima, M. Gołąb and M. Woźniakiewicz, *Green Anal. Chem.*, 2025, **12**, 100185, DOI: [10.1016/j.greeac.2024.100185](https://doi.org/10.1016/j.greeac.2024.100185).

46 A. M. Jiménez-Carvelo, A. Arroyo-Cerezo and L. Cuadros-Rodríguez, *TrAC, Trends Anal. Chem.*, 2024, **170**, 117463, DOI: [10.1016/j.trac.2023.117463](https://doi.org/10.1016/j.trac.2023.117463).

47 P.-Y. Sacre, Ch. A. Waffo Tchounga, Ch. De Bleye, Ph. Hubert, R. D. Marini and E. Ziemons, *Green Anal. Chem.*, 2024, **11**, 100158, DOI: [10.1016/j.greeac.2024.100158](https://doi.org/10.1016/j.greeac.2024.100158).

48 J. Martínez, J. F. Cortés and R. Miranda, *Processes*, 2022, **10**, 1274, DOI: [10.3390/pr10071274](https://doi.org/10.3390/pr10071274).

49 K. Van Aken, L. Strekowski and L. Patiny, *Beilstein J. Org. Chem.*, 2006, **2**, 3, DOI: [10.1186/1860-5397-2-3](https://doi.org/10.1186/1860-5397-2-3).



50 M. G. T. C. Ribeiro, D. A. Costa and A. A. S. C. Machado, *Green Chem. Lett. Rev.*, 2010, **3**, 149–159, DOI: [10.1080/17518251003623376](https://doi.org/10.1080/17518251003623376).

51 S. M. Mercer, J. Andraos and Ph. G. Jessop, *J. Chem. Educ.*, 2012, **89**, 215–220, DOI: [10.1021/ed0200249v](https://doi.org/10.1021/ed0200249v).

52 T. Klucznik, B. Mikulak-Klucznik, M. P. McCormack, H. Lima, S. Szymkuć, M. Bhowmick, K. Molga, Y. Zhou, L. Rickershauser, E. P. Gajewska, A. Touthkine, P. Dittwald, M. P. Startek, G. J. Kirkovits, R. Roszak, A. Adamski, B. Sierdzińska, M. Mrksich, S. L. J. Trice and B. A. Grzybowski, *Chem*, 2018, **4**, 522–532, DOI: [10.1016/j.chempr.2018.02.002](https://doi.org/10.1016/j.chempr.2018.02.002).

53 B. Mikulak-Klucznik, P. Gołębiewska, A. A. Bayly, O. Popik, T. Klucznik, S. Szymkuć, E. P. Gajewska, P. Dittwald, O. Staszewska-Krajewska, W. Beker, T. Badowski, K. A. Scheidt, K. Molga, J. Mlynarski, M. Mrksich and B. A. Grzybowski, *Nature*, 2020, **588**, 83–88, DOI: [10.1038/s41586-020-2855-y](https://doi.org/10.1038/s41586-020-2855-y).

54 Y. Lin, R. Zhang, D. Wang and T. Cernak, *Science*, 2023, **379**, 453–457, DOI: [10.1126/science.ade8459](https://doi.org/10.1126/science.ade8459).

55 A. Wołos, D. Koszelewski, R. Roszak, S. Szymkuć, M. Moskal, R. Ostaszewski, B. T. Herrera, J. M. Maier, G. Brezicki, J. Samuel, J. A. M. Lummiss, D. T. McQuade, L. Rogers and B. A. Grzybowski, *Nature*, 2022, **604**, 668–676, DOI: [10.1038/s41586-022-04503-9](https://doi.org/10.1038/s41586-022-04503-9).

