Accreditation and Quality Assurance
https://doi.org/10.1007/s00769-025-01649-7

RESEARCH

Is everything wrong in analytical chemistry? A study on reproducibility

Bruna Drielen Ferreira[1] · **Igor Renato Bertoni Olivares**[1] · **Emanuel Carrilho**[1] · **Vitor Hugo Polisél Pacces**[1]

Received: 1 April 2025 / Accepted: 18 May 2025

Abstract
Reproducibility in science, particularly in fields that develop and apply analytical methods, such as analytical chemistry and related disciplines, has been increasingly questioned. A survey by Monya Baker (2016) indicated that most researchers acknowledge a reproducibility crisis. Frequently cited causes include low statistical power or poor analysis, insufficient replication in the original laboratory, unavailability of methods, poor experimental design, and absence of raw data. Chemistry was the field in which researchers most frequently reported difficulty reproducing both their own and others' experiments. This study was conducted to confirm the existence of this crisis in analytical methods, quantify its extent, and evaluate its relationship with method validation and measurement uncertainty, which are based on statistical approaches and metrological principles, using the Analytical Quality Assurance Cycle. The results suggest that the crisis is directly associated with incorrect statistical procedures, inadequate validation criteria, and deficient execution of performance characteristics, factors that directly contribute to elevated measurement uncertainty. In 28% of the evaluated methods, expanded uncertainties exceeded 100% at the first point of the linearity assessment, compromising both result reliability and metrological traceability. These observations support concerns from previous studies regarding statistical errors, insufficient replication, and limited methodological transparency. Enhancing statistical training and quality assurance in academic programs, broader adoption of open science practices, and clearer policies aligned with international guidelines, such as EURACHEM, ISO/IEC 17025, and the OECD Principles of Good Laboratory Practice, may improve the reproducibility and reliability of analytical research.

Keywords Reproducibility · Analytical methods · Measurement uncertainty · Method validation · Metrological traceability · Statistical analysis

Introduction

Scientific reproducibility has drawn increasing attention from the scientific community in recent years [1–8]. Data presented in the 2016 article "1,500 scientists lift the lid on reproducibility" [1] indicate that researchers from various fields acknowledge the existence of a reproducibility crisis. According to the survey, the main causes identified by the respondents include insufficient knowledge and inadequate application of statistical tools, as well as poor replication of experiments within the original laboratories.

This study [9] investigates the existence of this crisis, quantifies its extent within analytical chemistry and related areas, and explores its association with key contributing factors, particularly method validation [10–16] and measurement uncertainty [17–22], as represented by the Analytical Quality Assurance Cycle (AQAC) [23].

Methods

This research was developed based on the Analytical Quality Assurance Cycle (AQAC) [23], a diagram that clarifies the interrelationships among method validation, measurement uncertainty, and analytical quality controls [13, 28] and illustrates how these quality assurance [13, 14] concepts are integrated to ensure traceable and reliable

✉ Bruna Drielen Ferreira
bdrielen21@gmail.com; brunadrielen@usp.br

Igor Renato Bertoni Olivares
igorolivares@iqsc.usp.br

Emanuel Carrilho
emanuel@iqsc.usp.br

Vitor Hugo Polisél Pacces
vitorusp@iqsc.usp.br

[1] Institute of Chemistry of São Carlos (IQSC), University of São Paulo, São Carlos, SP, Brazil

Published online: 16 June 2025

results. Supported by metrological traceability [13, 14, 17, 18, 23], method validation establishes whether a method is fit for its intended purpose, while measurement uncertainty quantifies the reliability and dispersion of the result. Continuous monitoring through quality control tools is essential to maintain method reliability [13–28].

Method validation confirms whether an analytical procedure developed for a particular objective has met its specific requirements [14, 18]. Once a method is validated, results obtained under the same conditions can be considered valid. However, to ensure the persistence of analytical performance over time, appropriate quality control (QC) strategies are required, so that results remain in conformity with the validation acceptance criteria [13, 17, 23].

However, according to the International Vocabulary of Metrology (VIM) [18], the true value of a quantity is, in principle and practice, unknowable [18]. Some approaches assume that no single true value exists, but rather a set of true values, due to the inherently incomplete detail in defining a quantity [18]. Other approaches disregard the concept of a true value altogether, considering instead that the validity of a measurement result depends on its metrological compatibility and traceability [18]. In summary, the statement of a measurement result is only complete when the assigned value is accompanied by its associated measurement uncertainty.

Some performance characteristics evaluated during method validation, such as linearity, precision, and accuracy, can be used as sources of uncertainty, as presented in the AQAC diagram [23]. Therefore, in this study, the measurement uncertainty associated with method validation is used as the statistical tool to assess the quality of results from analytical methods published in scientific literature.

Based on these results, the objectives of this study are to:

a) Investigate the existence of a reproducibility crisis in analytical chemistry and related fields;
b) Determine whether method validation was appropriately applied and whether suitable validation protocols were used;
c) Evaluate, using previously validated software (ConfLab Validation and ConfLab Uncertainty) [29], whether the mathematical calculations and statistical tests employed in method validation are correct and sufficient;
d) Identify the most common errors found in method validation practices;
e) Assess whether the findings from items (a) to (d) are related to the estimation of measurement uncertainty; and
f) Investigate whether a relationship exists between item (e) and the reproducibility crisis.

The performance characteristics of method validation [11–14] and their corresponding uncertainties were examined to investigate potential sources of the scientific reproducibility crisis in analytical chemistry. The analysis encompassed academic and scientific documents that reported the development of analytical methods, including articles, thesis, and dissertations retrieved from online platforms, databases, and virtual libraries.

Verifications were conducted by comparing the method validation results reported in the documents with those recalculated using ConfLab Software [29], a validated tool commonly employed in laboratories accredited under ISO/IEC 17025 [13] and OECD Principles of Good Laboratory Practice [14].

The validation performance characteristics [11–13] were recalculated from the raw data available in the selected documents using ConfLab Validation Software [29], which was considered the reference (or correct) value. The associated expanded measurement uncertainty for each performance characteristic was calculated using ConfLab Uncertainty Software [29], evaluating three different concentration levels per method.

The validation performance characteristics [11–13] considered in the analysis included linearity, precision, intermediate precision, accuracy, limit of quantification (LOQ), limit of detection (LOD), and expanded measurement uncertainty.

Results and discussion

Most (59%) of the analytical methods evaluated were found in articles from various countries, which constituted the primary focus of the analysis. However, due to the frequent unavailability of raw data in these publications, additional documents, such as theses, dissertations, and undergraduate final papers, were also included to allow a more complete assessment. This limitation directly reflects the issues highlighted by researchers in the referenced survey [1], which identified "methods, code unavailable" and "raw data not available from original laboratory" as major contributors to irreproducibility. These deficiencies compromise metrological traceability by hindering the verification of analytical results. The inclusion of academic works in this study is further justified because methods developed and validated in such contexts are often later published in peer-reviewed journals.

The selected articles were published in journals with an average impact factor of 4, placing them among the top 25 analytical chemistry journals worldwide, out of 120 indexed in the SJR—Scimago Journal & Country Rank [30, 31]. The thesis, dissertations, and undergraduate papers evaluated were produced at the top ten-ranked Brazilian universities, which are positioned between the 400th and 500th places

among 1300 institutions in the QS World University Rankings [32, 33].

Validation

The main aspects evaluated regarding method validation included the application of protocols about conformity with the analytical conditions required for the assessment of performance characteristics [10–15], the acceptance criteria for these results, the mathematical calculations and statistical tests used for performance characteristics, the presentation of results, and the adequacy of metrological traceability and the information provided.

As an evaluation criterion, when the method developer failed to apply an appropriate protocol to define the analytical conditions and acceptance criteria for assessing performance characteristics, their understanding of the objective and scope of the analytical technique under development was deemed insufficient or unsatisfactory.

As shown in Fig. 1, the assessment of performance characteristics and the analytical conditions adopted during validation revealed deficiencies in most methods, with many cases classified as unsatisfactory, insufficient, or inadequate. This is evident in the results derived from the incorrect or incomplete application of procedures and statistical calculations used to evaluate these characteristics, findings that are consistent with the concerns raised by researchers in the article "1,500 scientists lift the lid on reproducibility" [1].

With respect to protocol application, 47% of the authors utilized the appropriate validation protocol. However, even among these, 63% failed to execute all required performance characteristic studies, and 81% did not apply the prescribed analytical conditions. Linearity was the performance characteristic with the most significant impact on validation conclusions, as all cases exhibited methodological deficiencies, either in statistical treatment (inappropriate selection of regression metrics and omission of significance testing) or in failure to follow established protocols.

These deficiencies are consistent with the reproducibility challenges identified by researchers in the referenced study [1], including "low statistical power or poor analysis," "poor experimental design," and "not replicated enough in original laboratory." Intermediate precision evaluations were frequently insufficient, and only 8% of the studies cited or reported the use of calibrated equipment, undermining metrological traceability.

Such traceability deficiencies compromised result reliability and raised concerns about the studies' methodological rigor. Figure 1 summarizes the assessment outcomes, identifying three major deficiencies in method validation: inadequate protocol application, calculation errors, and insufficient metrological traceability.

Uncertainty results

Measurement uncertainty was calculated using ConfLab Uncertainty Software [29] based on the performance characteristics obtained from method validation data. Since only one document explicitly reported uncertainty values, only evaluations could be performed rather than comparisons.

The study adopted the EURACHEM/CITAC approach [1] for uncertainty estimation, expressing results as expanded uncertainties, calculated as the product of the combined standard uncertainties and the coverage factor. As only the uncertainties associated with method validation performance characteristics were considered, the reported values may be underestimated in broader contexts. Based on the calibration curve, uncertainty was

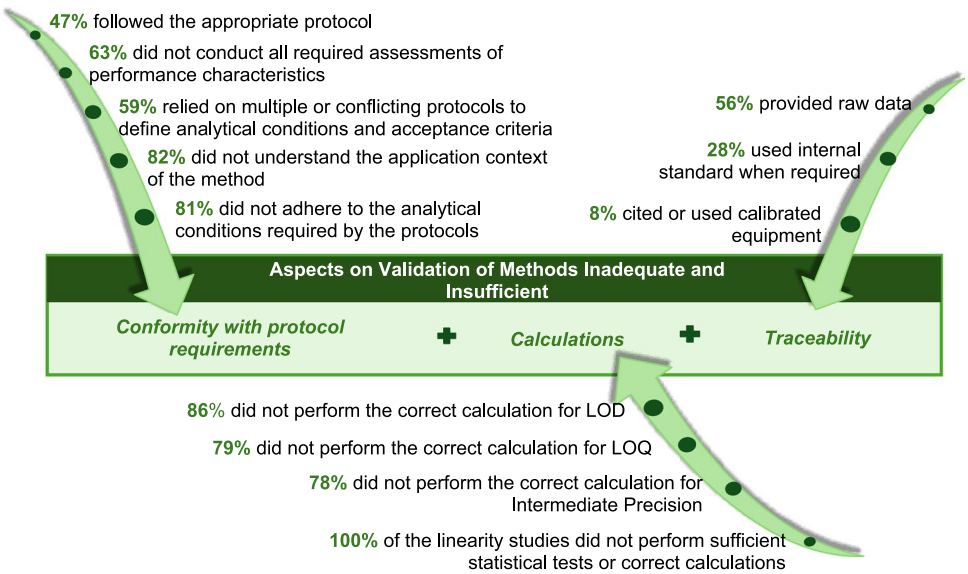

Fig. 1 Overview of method validation results. Inadequate application of validation protocols, incomplete assessment of required performance characteristics, and frequent inconsistencies in statistical treatment compromise the validity of analytical methods

assessed at three concentration levels for each method, namely low (or limit of quantification), medium, and high.

At the low concentration level, only 43% of the expanded uncertainty values were considered acceptable (< 40%), 29% were questionable (40–100%), and 28% were aberrant (> 100%) (Fig. 2). In approximately 80% of the cases, the standard uncertainty derived from linearity was the dominant contributor to the expanded uncertainty, followed by intermediate precision.

Figure 2 presents the distribution of expanded uncertainty results at the lowest concentration level assessed. Acceptable uncertainties (< 40%) were observed in 43% of the validations, while questionable values (40–100%) accounted for 29%, and aberrant uncertainties (> 100%) were found in 28% of the cases.

Expanded uncertainty results above 40% indicate a low probability of obtaining values close to the measured result upon repetition of the method. Values exceeding 100% suggest an almost negligible probability of obtaining results near the expected value. This variability arises from the high dispersion of values associated with measurements obtained using the evaluated methods, even when analysis conditions, equipment, and analysts remain unchanged.

Such variability reflects a central aspect of the reproducibility crisis previously discussed in the literature [1–8], particularly the insufficient use of and limited knowledge about statistical tools, which has been identified as one of the most critical contributing factors. These aspects are summarized in Fig. 3.

Measurement uncertainty can also be directly linked to the reproducibility crisis, considering its definition: "a parameter characterizing the dispersion of the values being attributed to a measurand, based on the information used" [17, 18]. As illustrated in Figs. 1, 2, and 3, higher uncertainty levels, especially stemming from improperly validated methods, lead to greater dispersion in measured values and lower probabilities of obtaining results close to the expected ones.

This relationship is clearly illustrated in the Analytical Quality Assurance Cycle (AQAC) diagram [23], in which method validation directly influences measurement uncertainty, and both impact the reliability and metrological traceability of analytical results. According to ISO/IEC 17025 [13], validation is defined as the demonstration, through objective evidence, that a method is fit for its intended purpose. When this process is conducted inadequately, whether due to the adoption of incorrect acceptance criteria,

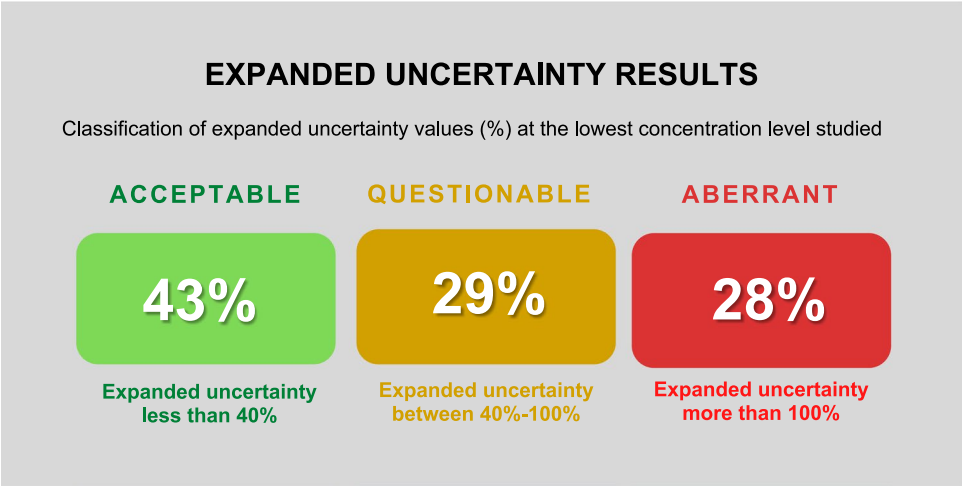

Fig. 2 Classification of expanded uncertainty results at the lowest concentration level studied (LOQ) Only 43% of the results presented acceptable expanded uncertainty values (< 40%), while 29% were questionable (40–100%), and 28% exceeded 100%, indicating critically low measurement reliability at this level

Percentage	Description
Only 19%	Authors who applied the performance characteristics according to the protocol appropriate for the developed method..
Only 17%	Validations in which all performance characteristics met the acceptance criteria defined by the correct protocol.
Only 43%	Expanded uncertainty results that showed acceptable values (<40%) at the lowest concentration levels evaluated during the validations.

Fig. 3 Highlights and key observations on method validation and uncertainty. Note the critically low percentages of adequate protocol implementation and complete assessment of performance characteristics, highlighting methodological deficiencies that compromise metrological traceability and reproducibility in analytical results

inappropriate application of statistical tools, or incomplete evaluation of performance characteristics, the method fails to meet its intended purpose, as it is unable to consistently produce results close to the expected values due to high associated uncertainty. These deficiencies propagate throughout the development and application of analytical methods, increasing variability and ultimately compromising their reproducibility. Therefore, method validation and measurement uncertainty are interdependent and form a causal chain at the core of the reproducibility crisis.

Conclusion

So, is everything wrong? No. But this study confirms that scientists and researchers' highlighted challenges regarding the difficulty of reproducing or obtaining consistent results are justified. This is supported by observing extremely high measurement uncertainty values, primarily resulting from deficiently conducted method validation. Fundamental pillars, such as method validation, measurement uncertainty, and traceability, are often applied inadequately and incompletely. These deficiencies directly intensify the reproducibility crisis and must be addressed to restore confidence in analytical results.

A collective effort across academia, laboratories, and editorial standards will be essential to rebuild trust in the scientific process. The following actions may help mitigate the reproducibility crisis:

a. Universities should include courses on chemical metrology, particularly applied aspects such as method validation and uncertainty estimation, as well as quality management systems in laboratories, including ISO/IEC 17025 [13] and the OECD Principles of Good Laboratory Practice [14].
b. If formal implementation of a quality management system [13, 14] is not feasible, research institutions can adopt relevant principles to ensure the technical competence of their laboratories. As an alternative, consistent application of AQAC principles [23] is recommended, including implementing quality control and assurance procedures to support traceability and reliable results.
c. Scientific journals should define minimum requirements for method validation and uncertainty reporting as a prerequisite for manuscript acceptance.
d. The EURACHEM/CITAC Guide [17] should be used for uncertainty estimation with the same level of importance as method validation, to promote standardization of uncertainty calculations across research studies.

These results justify and confirm the concerns raised by researchers in the survey [1], particularly those related to poor statistical practices, insufficient replication, and restricted access to essential methodological data, all of which were also identified in this study as key contributors to irreproducibility in analytical chemistry. This study contributes to a better understanding and implementation of analytical quality assurance, supporting laboratories in achieving greater reliability, scientific credibility, and regulatory compliance.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00769-025-01649-7.

Acknowledgements BDF discloses support for the research of this work from Funder CAPES (grant number 88887.483785/2020-00) and the Institute of Chemistry of São Carlos, University of São Paulo (IQSC-USP).

Author contributions BDF* was responsible for data acquisition, study conception and design, data analysis, and manuscript writing. IRBO developed the software for data analysis and participated in the critical review of the manuscript, focusing on quality management in laboratories, statistics, and metrology in analytical chemistry. EC supervised the applied analytical chemistry area of the work, acting as co-advisor of the study. He was also responsible for the critical review of the manuscript, emphasizing applied analytical chemistry. VHPP served as the research group leader, overseeing all work areas and serving as the study advisor. He also developed the software used in data analysis, critically reviewed the manuscript, and contributed significantly to the project.

Funding Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, 88887.483785/2020-00.

Data availability Data are provided within the manuscript or supplementary information files.

Declarations

Conflict of interest The authors declares that they have no conflict of interest.

References

1. Baker M (2016) 1500 scientists lift the lid on reproducibility. Nature 533:452–454. https://doi.org/10.1038/533452a
2. Nature Index (2024) Peer-replication model aims to address science's 'reproducibility crisis'. Nature Index. https://www.nature.com/peer-replication-model-aims-to-address-sciences-reproducibility-crisis. Accessed 25 Mar 2025
3. Marshall-Cook J, Farley M (2024) The hidden sustainability cost of the reproducibility crisis. Nat Rev Phys 6:4–5. https://doi.org/10.1038/s42254-023-00674-0
4. Nature (2021) Understand the real reasons reproducibility reform fails. Nature. https://www.nature.com/articles/d41586-021-03617-w. Accessed 25 Mar 2025
5. Button K (2018) Reboot undergraduate courses for reproducibility. Nature. https://www.nature.com/articles/d41586-018-06692-8. Accessed 25 Mar 2025
6. Begley C, Buchan A, Dirnagl U (2015) Robust research: institutions must do their part for reproducibility. Nature 525:25–27. https://doi.org/10.1038/525025a.Accessed25Mar2025

7. Goodman SN, Fanelli D, Ioannidis JPA (2016) What does research reproducibility mean? Sci Transl Med. https://doi.org/10.1126/scitranslmed.aaf5027
8. Berg J (2018) Progress on reproducibility. Science 359:9. https://doi.org/10.1126/science.aar8654
9. Ferreira BD (2021) Scientific reproducibility crisis in analytical chemistry: Evaluation with a focus on method validation and result uncertainty. Dissertation, University of São Paulo, São Carlos. https://doi.org/10.11606/D.75.2021.tde-11012022-110019. Accessed 25 Mar 2025
10. National Institutes of Health (2021) Final Report Advisory Committee Animal Research – ACD Working Group on Enhancing Rigor, Transparency, and Translatability in Animal Research. US Department of Health & Human Services. https://acd.od.nih.gov/working-groups/eprar.html. Accessed 25 Mar 2025
11. International Council for Harmonisation (ICH) (2022) ICH M10 on bioanalytical method validation—Scientific guideline. https://www.ema.europa.eu/en/ich-m10-bioanalytical-method-validation-scientific-guideline. Accessed 25 Mar 2025
12. International Conference on Harmonisation (ICH) (2021) Q14 Analytical procedure development and validation. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/q14-analytical-procedure-development. Accessed 25 Mar 2025
13. ISO/IEC 17025:2017. General requirements for the competence of testing and calibration laboratories. International Organization for Standardization, Geneva
14. OECD. OECD Principles of Good Laboratory Practice (as revised in 1997). OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 1, ENV/MC/CHEM(98)17, Organisation for Economic Co-operation and Development, Paris (1998)
15. Cantwell H (ed) (2025) The fitness for purpose of analytical methods—A laboratory guide to method validation and related topics, 3rd edn. EURACHEM. https://www.eurachem.org
16. Thompson M, Ellison SLR, Wood R (2002) Harmonized guidelines for single-laboratory validation of methods of analysis (IUPAC technical report). Pure Appl Chem 74:835–855. https://doi.org/10.1351/pac200274050835
17. EURACHEM/CITAC (2012) Quantifying uncertainty in analytical measurement, 3rd edn. https://www.eurachem.org/index.php/publications/guides/quam. Accessed 25 Mar 2025
18. JCGM (2012) International vocabulary of metrology – Basic and general concepts and associated terms (VIM), 3rd edn. https://www.bipm.org/en/committees/jc/jcgm/publications. Accessed 25 Mar 2025
19. Raposo FC, Ramsey MH (2024) The role of measurement uncertainty in the validation of a measurement procedure. Accred Qual Assur 29:267–274. https://doi.org/10.1007/s00769-024-01593-y
20. Belouafa S, Bouchmila R, Benhidour A et al (2017) Statistical tools and approaches to validate analytical methods: methodology and practical examples. Int J Metrol Qual Eng 8:9. https://doi.org/10.1051/ijmqe/2016030
21. Alvarez-Prieto M, Jiménez-Chacón J (2018) A fitness for purpose approach to validation and verification of analytical measurements. Accred Qual Assur 23:219–229. https://doi.org/10.1007/s00769-018-1331-y
22. Feinberg M, Laurentie M (2006) A global approach to method validation and measurement uncertainty. Accred Qual Assur 11:3–9. https://doi.org/10.1007/s00769-005-0081-9
23. Olivares IRB, Lopes FA (2012) Essential steps to providing reliable results using the analytical quality assurance cycle. TrAC Trends Anal Chem 35:109–121. https://doi.org/10.1016/j.trac.2012.01.004
24. Olivares IRB (2023) Quality management in laboratories, 5th ed., revised and expanded. Átomo, São Paulo. ISBN 978-65-87322-22-3
25. Feinberg M (2022) Analytical quality by design and quality culture: key factors in laboratory quality management systems. Accred Qual Assur 27:73–82. https://doi.org/10.1007/s00769-022-01531-5
26. Baker M (2016) How quality control could save your science. Nature 529:456–458. https://doi.org/10.1038/529456a
27. Barwick V (ed) (2016) EURACHEM/CITAC Guide: Guide to Quality in Analytical Chemistry – An Aid to Accreditation, 3rd edn. EURACHEM/CITAC. https://www.eurachem.org
28. Barwick V (ed) (2023) EURACHEM Guide: Terminology in Analytical Measurement – Introduction to VIM 3, 2nd edn. EURACHEM. https://www.eurachem.org
29. ConfLab Software (2024) Validation, uncertainty, and quality control. https://www.conflab.com.br/home. Accessed 25 Mar 2025
30. QS Quacquarelli Symonds Limited (2024) QS World University Rankings 2024. https://www.topuniversities.com. Accessed 25 Mar 2025
31. Center for World University Rankings (2024) CWUR World University Rankings—Brazil. https://cwur.org/2024.php. Accessed 25 Mar 2025
32. US News & World Report LP (2024) 2024 Best Global Universities Rankings. https://www.usnews.com/education/best-global-universities/rankings. Accessed 25 Mar 2025
33. SJR—Scimago Journal & Country Rank (2024) Journal Rankings on Analytical Chemistry. https://www.scimagojr.com/journalrank.php?category=1602&area=1600&order=cpd&ord=desc&page=1&total_size=121. Accessed 25 Mar 2025