

Commentary: Cancer research quality and tumour classification

Tumor Biology

February 2020: 1–4

© The Author(s) 2020

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1010428320907544

journals.sagepub.com/home/tub**Ian A Cree** and **B Iciar Indave**

Abstract

Cancer researchers require accurate diagnoses for the samples, cell lines, patients or populations that they study. These diagnoses are underpinned by an internationally accepted taxonomy – the World Health Organization Classification of Tumours. This is still largely based on the histopathological examination of biopsy specimens, but increasingly also molecular methods and radiological examination of patients. Classifications evolve as new evidence arises, and for tumours that evidence is available in a quantity that is both remarkable and daunting. Evaluating this deluge of new information and incorporating it into the World Health Organization Classification of Tumours is now the responsibility of an editorial board, and up to 200 editors and authors work on each system to update it within the new 5th edition. Just as cancer researchers depend on the classification for diagnoses, so too the classification depends on the generation of high-quality, trustworthy data by cancer researchers. It is not just a case of quantity but quality too. Scientific fraud is thankfully rare, but high-profile cases are damaging and standards need to improve, not least to ensure that accurate information enters the classification.

Keywords

Cancer, tumours, classification, research, quality

Date received: 25 November 2019; accepted: 27 December 2019

Introduction

All cancer researchers depend on the diagnosis of the tumours they study. This includes the epidemiologist studying the incidence of a given cancer type worldwide or the user of a cell line from an individual patient's tumour. Many take this for granted, but underpinning every diagnosis is a global system which has until recently been based almost exclusively on histopathological examination of biopsy material. Advances in radiology, computational pathology and molecular pathology are challenging this status quo, but the need for patient diagnosis based on an internationally accepted set of standards has never been greater. There is need to know the diagnosis to put samples into studies, or patients into clinical trials, so changes in the classification has immediate relevance to research. In epidemiology, there is also a need to understand the effect that changes in classification may have on tumour registration and therefore on incidence. There is also a need for this to be international: one must be

able to compare results between studies or trials to benefit patients worldwide.

Since 1957, the World Health Organization (WHO) has had a mandate from the World Health Assembly to produce a classification of disease. The International Classification of Disease, now in its 11th iteration (<https://icd.who.int/en>), but for tumours, coding meant nothing without the standards included in what Dr Leslie Sobin and subsequently Dr Paul Kliehues developed as the WHO Classification of Tumours, the de facto gold standard for tumour diagnosis. The WHO Classification of Tumours, published by IARC from

International Agency for Research on Cancer (IARC), World Health Organization, Lyon, France

Corresponding author:

Ian A Cree, International Agency for Research on Cancer (IARC), World Health Organization, 150 Cours Albert Thomas, 69372 Lyon CEDEX 08, France.

Email: creei@iarc.fr



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial

use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

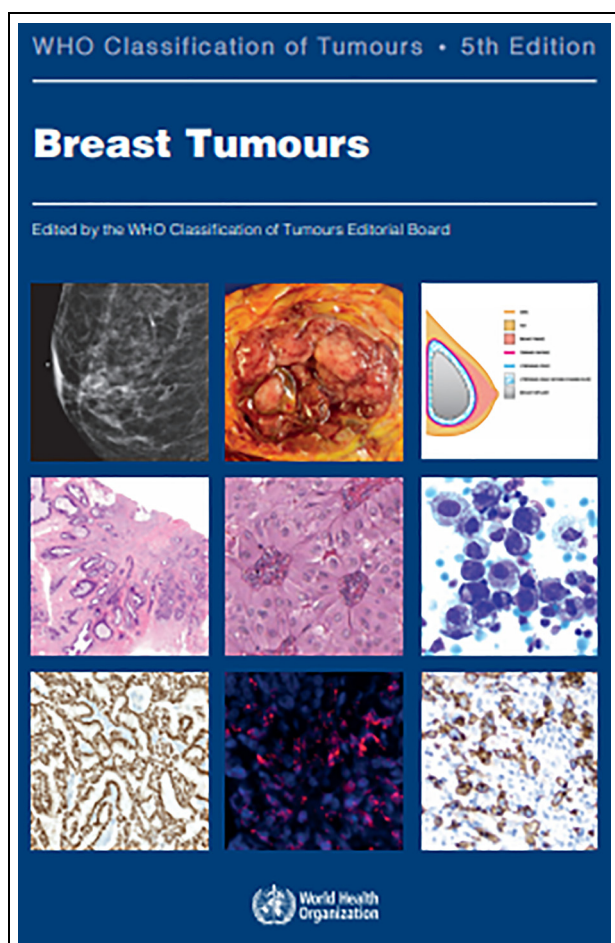


Figure 1. Front cover of the 5th edition of WHO Classification of Tumours: Breast.

2006 to 2018, in the 4th edition as a series of 12 books, and now in its 5th edition (Figure 1), classifies tumours by site and histology (<http://whobluebooks.iarc.fr>).

Classifications evolve as new evidence arises. This explains the success of Linnaeus' classification of the natural world, and even today this changes as new species are identified and their relationship to one another becomes clearer. Linnaeus classified organisms on the basis of shared characteristics,¹ and the WHO Classification of Tumours is no different. It needs to change to meet the challenge of new information, particularly from genomics, but also from other 'omic' technologies and increasingly from computational pathology and radiology.²

The challenge for the fifth series, which started in 2018 and saw its first publication^{3,4} in July 2019, is to meet the acceleration in the acquisition of knowledge and the resulting information overload, while improving the quality of the classification. Another challenge is to do this faster than ever before, to meet the clinical need for up-to-date diagnosis to benefit patients directly.

Accelerating translation in cancer diagnosis

The classification is therefore now in the hands of an editorial board, consisting of standing members, mainly pathologists, and expert members who join the board for specific books. They increasingly include radiologists, surgeons, physicians and epidemiologists. This increases the pool of knowledge available for the books, and that is further enhanced by reviews of recent literature conducted by the 150–200 authors responsible for each book.

IARC has started the IARC Collaboration for Cancer Classification and Research (IC3R) to coordinate efforts between major research organisations to increase their ability to produce high-quality data, improve the harmonisation of databases, and ensure translation using systematic review technology and careful validation of new methods.

Improving research quality

The quantity of research is not in doubt, but improving quality has become a major issue.⁵ This is not just an issue of the fraudulent production of research papers driven by the rewards available to those successful in publishing scientific papers, which can be beyond the scope of even the most expert editor to detect. There is a proliferation of new journals publishing low-quality research, including predatory journals which often approach potential authors with unsolicited requests for papers. Publication bias distorts the evidence base, and yet it can be difficult to publish confirmatory results or to dispute published findings. Even high-quality journals will publish work with inadequate controls, and universities reward those who obtain large grants or papers in high-impact factor journals without thought to the consequences.

Circulating free DNA for early cancer detection is a case in point. The literature on this potentially exciting development is rife with inadequately powered, poor-quality papers, and the problems have been identified in a recent systematic review.⁶ We concluded,

Preanalytical, analytical, and post-analytical considerations were identified which need to be addressed before such biomarkers enter clinical practice. The value of small studies with no comparison between methods, or even the inclusion of controls is highly questionable, and larger validation studies will be required before such methods can be considered for early cancer detection.

Most were single centre studies, few studies had been reproduced and the median study size was just 65 cases, and even in 2017, studies in serum were still being published despite well-publicised problems with this

Table 1. Framing a clinical research question using the PICO(T) method.

P	Patients or population	Describe patients or population to be studied. Consider demographics or other relevant characteristics
I	Investigation, intervention or indicator	This may be a new patient management or treatment strategy, a drug or a diagnostic device, investigation or determination
C	Comparator or control	What is the alternative? Usually the control, existing practice, treatment or diagnostic method
O	Outcome	Patient-relevant measurable consequences of the intervention or investigation. Mortality, adverse events, recurrence or a validated surrogate marker; are examples of common outcomes
T	Time or type of study	Time spans to consider or what study types are most likely to provide outcomes that answer your research question

approach which mean that plasma is always used. cfDNA is making its way into the clinic slowly, particularly for patient monitoring and mutation detection, but it has taken much longer and a lot more financial support than was arguably necessary.

Scientific fraud and culpably weak science are thankfully rare,⁵ but high-profile cases⁷ might make one think otherwise. Such problems detract from the overwhelmingly honest, hard-working and sometimes brilliant majority of biomedical researchers who work within imperfect systems, but have been responsible for the amazing advances in our understanding of cancer and the clinical advances from which patients now benefit.

Translation of diagnostic research into practice

We would like to believe that changes in clinical practice results from careful consideration of collected scientific clinical evidence, without bias, but the fact is that it is difficult to achieve. Experts who take part in the WHO Classification of Tumours are now selected for each volume on the basis of their expertise and publication record.²⁻⁴ Co-authors are purposely selected from different institutions and may not know each other. We believe that this reduces bias, but we ask authors to consider as much of the literature as possible between them. Statements of fact are ideally justified by two suitably sized and well-conducted studies or a review (ideally systematic, but rarely so). There are undoubtedly still problems to solve, but we are determined to fix as many as possible over the next few years.

Many of the problems start at the beginning. Clinical laboratories have to meet very strict accreditation standards (e.g. ISO15189) to ensure that samples are not mixed up – most will have this happen in less than 1 in 100,000 samples, due to rigorous identification using at least four independent identifiers and the use of accession systems including bar codes, which are then read at every stage in the process. Automation certainly helps too. Diagnostic methods within clinical

laboratories are validated: analytical validation ensures that measurements are accurate, while clinical validation ensures that they are relevant to the cases in which they are used. Documentation training, equipment maintenance, internal and external quality assurance are all pre-requisites for accreditation and improved results.⁸ Good Laboratory Practice (GLP) is an attempt to achieve similar levels of accuracy within research laboratories, but many would argue that the changing nature of the research from day to day or experiment to experiment makes it impossible to meet the standards expected of clinical laboratories.

To improve the quality of the evidence provided in the field of pathology, research projects need to define research questions carefully and ensure that study designs are methodologically sound. A number of tools can help researchers to define the core concepts of a research question. PICO is a widely used mnemonic which stands for ‘Population’, ‘Intervention’ (e.g. a new diagnostic), ‘Comparator’ (usually the gold standard) and ‘Outcome’ (the diagnosis, prognosis or other) used to frame study questions. It has its limitations, but includes the features listed in Table 1 and helps define a clinical question. In addition, the selection of an appropriate research design helps to minimise bias, addresses validity issues, ensures representative participant/case selection, defines variables and develops reliable outcome measurements, as well as addressing statistical and clinical relevance. The goal of research design is always to minimise systematic errors and descriptive work and hence risk of bias. Evidence levels observe a hierarchy, placing randomised controlled trials (RCT) or systematic reviews at the top of the evidence pyramid, considering cohort studies as a good level of evidence for prognostic and etiologic questions, whereas case reports are regarded as low-level evidence with a high risk of bias. Higher levels of evidence can be produced by applying adequate study design to assess pathology-relevant research questions, and all studies should include a sample size calculation as a matter of course. These can be applied to case-control studies

and the process of thinking this through often leads to improvements.

Many pathology studies are descriptive – this is almost inevitable when the tumour type is rare and there are less than 20 cases in the literature. The literature in question then tends to be in the form of case reports or small series, often with painstaking attention to a panel of immunohistochemical stains and molecular investigations. These can be informative, but it is rare that anyone else has used the same methods, which makes the job of assessing their merit for changes in classification difficult. It is however possible to review the information and sometimes tumour material to produce a coherent picture of the results. For larger series, central pathology review is essential and external quality assurance can be used to ensure that results obtained from different centres can be combined. Addition of a control group in a case-control design can be helpful.

Finally, the quality of publications can be improved by adherence to the guidance published by the EQUATOR network (www.equator.net) which includes the STARD guidance for diagnostic studies and the PRISMA guidance for assessing multiple diagnostic studies. All journals should insist that these are followed.

Conclusion

There is no quick fix for the problems of research quality or indeed for the incorporation of research output into the WHO Classification of Tumours. The translation of research to practice is not easy but can be improved. Lessons learned from clinical laboratories may help, and there is a chance that research laboratory accreditation will come into being at some point. The use of systematic review methods has much to recommend it to evaluate research and these being incorporated into the WHO Classification of Tumours.

Authorship

This article is based on many discussions with a large number of researchers, pathologists and other clinical colleagues as well as patients over the last 2 years. I.A.C. conceived, wrote and reviewed the manuscript at 33,000 feet over the Atlantic. He is happy to claim any errors as the product of a jet-lagged mind.

Disclaimer

The content of this article represents the personal views of the authors and does not represent the views of the authors'

employers and associated institutions. Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.


Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Ian A Cree  <https://orcid.org/0000-0001-5007-9456>

References

1. *Caroli Linnaei... Systema naturae per regna tria naturae: secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis.* Holmiae: Laurentii Salvii, 2007, pp. 1758–1759.
2. Salto-Tellez M and Cree IA. Cancer taxonomy: pathology beyond pathology. *Eur J Cancer* 2019; 115: 57–60.
3. WHO Classification of Tumours Editorial Board. *Digestive system tumours* (WHO classification of tumours series). 5th ed, vol. 1. Lyon: International Agency for Research on Cancer, 2019, <http://publications.iarc.fr/579>
4. Nagtegaal ID, Odze RD, Klimstra D, et al. The 2019 WHO classification of tumours of the digestive system. *Histopathology* 2020; 76: 182–188.
5. Campos-Varela I, Villaverde-Castañeda R and Ruano-Raviña A. Retraction of publications: a study of biomedical journals retracting publications based on impact factor and journal category. *Gac Sanit.* Epub ahead of print 14 September 2019. DOI: 10.1016/j.gaceta.2019.05.008.
6. Cree IA, Uttley L, Buckley Woods H, et al. The evidence base for circulating tumour DNA blood-based biomarkers for the early detection of cancer: a systematic mapping review. *BMC Cancer* 2017; 17(1): 697.
7. Arif N, Al-Jefri M, Bizzi IH, et al. Fake news or weak science? Visibility and characterization of antivaccine webpages returned by Google in different languages and countries. *Front Immunol* 2018; 9: 1215.
8. Gutowska-Ding MW, Deans ZC, Roos C, et al. One byte at a time: evidencing the quality of clinical service next-generation sequencing for Germline and somatic variants. *Eur J Hum Genet* 2019; 28: 202–212.