IJC
International Journal of Cancer

Invited commentary—WHO Classification of Tumours: How should tumors be classified? Expert consensus, systematic reviews or both?

Lesley Uttley[1], Blanca Iciar Indave[2], Chris Hyde[3], Valerie White[2], Dilani Lokuhetty[2] and Ian Cree[2]

[1]School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, United Kingdom
[2]World Health Organization, Classification of Tumours Group, International Agency for Research on Cancer, Lyon, France
[3]Exeter Test Group, College of Medicine and Health, University of Exeter, Exeter, United Kingdom

Dear Sir,

The World Health Organization (WHO) Classification of Tumors series, published as the WHO Blue Books (Fig. 1) and the accompanying website, are essential resources for pathologists across the globe, providing the standards against which tumors are classified to aid cancer diagnosis, research, treatment and prognosis. Each book tackles the classifications of up to 300 tumor types, defining for each of these the etiology, pathogenesis, epidemiology, clinical features, macroscopic appearances, histology, cytology, molecular pathology, essential and desirable diagnostic features, staging, prognostic factors and predictive biomarkers. If each tumor were to require one formal review, this would mean finding or conducting up to 3,600 such reviews for each book which would not be practicable or time-efficient. The current approach therefore relies largely on subject experts drawing on published literature searches according to their individually perceived need to inform the content of these books. However, these decisions affect the classification, and hence the diagnosis and management of cancer patients worldwide. Consequently, to minimize the risk of including biased information, the evidence is weighed and decisions made by an editorial board. This editorial board is composed of standing members and content experts (mainly practicing pathologists) who meet, propose, revise and agree on definitions and core criteria for each tumor systematically for every new edition.

Over-reliance on expert consensus at editorial board meetings for this purpose may lead to problems (Table 1). First, evidence needs to be reviewed to assess all relevant issues in a WHO Blue Book. But, even if all participating experts are required to perform literature reviews, their different expertise and inevitable time constraints produce variable results with potential for studies that are relevant to decisions to be missed. In the first book of the new 5th edition, up to 130 tumors or subjects were described as "unknown" and 200 labeled as clinically not relevant, this often being the only description of a whole section.[1] Sections such as etiology or pathogenesis are frequently summarized as "unknown," without providing more details, leaving unclear whether an exhaustive search of the literature has been performed. This does not necessarily indicate that systematic reviews should be routinely performed when such questions arise and indeed, it would not be possible for standing members and content experts to do so given the number that would be required. Often the statements requiring evidence are more related to background information than issues relevant to classification, and introducing systematic approaches to the literature searches would not improve the evidence base of the review results. In addition, this would allow authors to focus on the pressing or contentious issues of tumor classification: topics that need to be assessed by systematic reviews, due to their potential controversy. Second, there is a potential risk that unintended bias from content experts could influence decisions in the absence of a structured and controlled process of evidence synthesis. This could lead to certain studies being unduly highlighted or overlooked depending on who is reviewing the literature and what their own clinical or research knowledge may be.[2] Third, if the convened expert panel is not sufficiently representative,[3] skewed decisions may be made, which may not be fully informed by the best and most relevant evidence.[4] Fourth, due to interpersonal and cultural differences among experts during meetings, it is possible for individuals to dominate discussion.[3,5,6] Finally, previously included but incorrectly referenced evidence may be carried forward during the updating process.

It is possible that these problems could be mitigated by the addition of evidence-based practices to the editorial process. Systematic reviews are widely regarded as the cornerstone of evidence-based medicine, encouraging comprehensive literature searching, transparency in methods and rigorous study appraisal.[7] Performing systematic reviews for some tumor types could certainly improve the reliability of decisions taken by the editorial board, but they may not always be the best solution.

Systematic reviews require training in methods, as well as expertise.[8–10] They can be laborious to complete, time-consuming and difficult to interpret, particularly when there is a lot of evidence, or when the evidence is low-grade. Systematic review methods have been described as being formulaic.[11,12] Best practice guidelines such as Cochrane[13] and PRISMA[14] are closely aligned with meta-analytic reviews for medical interventions, the methods of which are not necessarily appropriate to pathology. Systematic reviews can be long and unwieldy documents, difficult to interpret and hollow in their conclusions.[15,16]

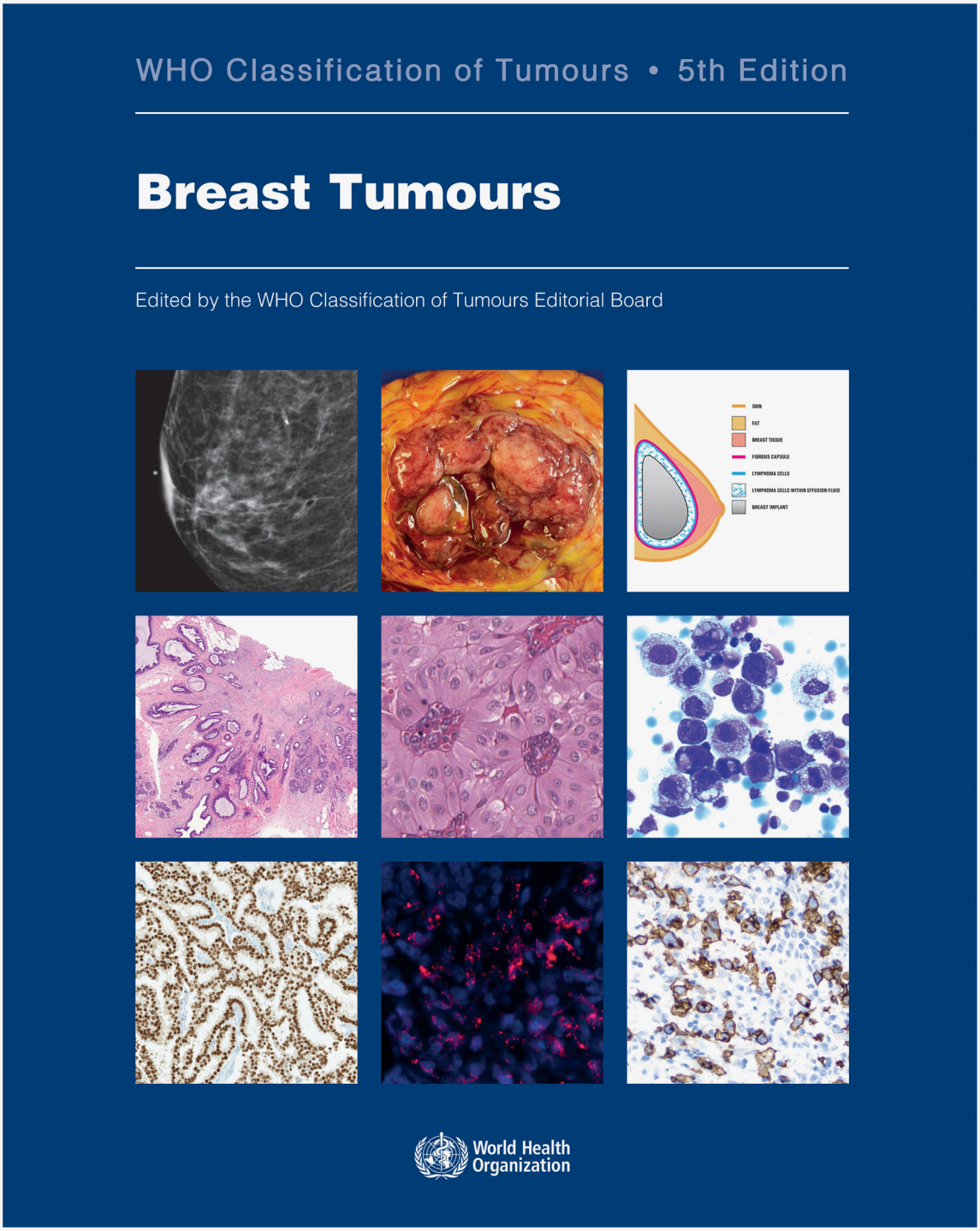

Figure 1. The World Health Organization (WHO) Classification of Tumors series. Breast tumors. Fifth Edition.

The hierarchy of evidence-based medicine with systematic reviews at the apex has been resisted by some academic groups and clinicians for oversimplifying complex issues relating to clinical care.[17] Some argue that clinical judgment based on experience is more exact and applies better to real-life clinical situations, because of its emphasis on individual cases rather than evidence derived from randomized controlled trials (RCTs) that are produced without social and political confounding effects and of other variables.[18] While the scientific rigor of systematic reviews and indeed RCTs are preserved

Table 1. Problem arising from consensus-based approach and proposed solutions by an evidence-based approach

Problem of a consensus-based approach	Solution by an evidence-based approach	Potential problems not solved by an evidence-based approach
Risk of missing relevant research	Comprehensive searching which is part of systematic reviewing may improve identification of important literature. A structured, systematic process allows summarizing and evaluating complex information such as big data or basic research information provided for molecular pathology.	Risk of missing research which does not fit into standard study design framework used for systematic reviews. Publication bias may not be addressed if only searching published evidence.
Selection of the literature may be biased	Systematic reviews require clearly stated inclusion criteria, so cherry-picking of particular studies to prove a particular point is easier to spot. In addition, the setting of acceptable evidence levels and assessment of risk of bias of studies avoids the use of inappropriate evidence.	Presentation of results may still allow a certain degree of "cherry-picking" when presenting only on selected outcomes (outcome reporting bias).
Interpretation of the literature may be biased	Systematic reviews consider each included study equally, unless there is a specific reason why less emphasis should be placed on it such as small sample or poor study quality. Risk of bias assessment of individual studies, but also the body of evidence can be undertaken to aid an appropriate interpretation of the retrieved evidence.	The use of several reviewers may not provide the desired control of bias effect and instead interesting information, may not be incorporated due to disagreement.
Panel of experts may be biased in composition or be dominated by particular individuals	A systematic review with clear eligibility criteria made available in a protocol may provide a reference point against which "extreme" views by particular panel members can be mitigated.	Panel may still be biased in developing eligibility criteria, even if an evidence-based approach helps in the discussion.
Difficulties in documentation of included evidence (especially in the updating process)	Systematic review protocol and reports document the biomedical databases searched and over what time period therefore uncertainty about whether a particular study has been included or not is much less likely to occur.	Relies on the well-designed and appropriate literature searches and implementation of reporting standards for systematic reviews.
Credibility of classification may be undermined if not evidence-based	Use of systematic review methods will improve the credibility of the classification, as well as the reliability of tumour classification.	Credibility may also be affected by other factors not addressed by a systematic review process. Experts in the field, important to the credibility of the books, may be put-off by the systematic review process.

through adhering to reporting guidelines, they have been shown to be less readable than article types which allow some spin or journalistic creativity, such as editorials and narrative reviews.[19] They are also subject to influence if not conducted objectively and not properly assessed prior to publication.[20] Managing the influence and expertise of team members in systematic reviews, like nonsystematic reviews, remains critical to producing reliable and externally valid conclusions.[10]

Although essential for the synthesis of available scientific evidence in controversial questions, it is clear that systematic review methods would need to be adapted to the specific needs of the WHO Classification of Tumours and its editorial process. Timely procurement of a valid evaluation of evidence is one of the major problems for process-driven systems such as the WHO Classification of Tumours, and this is something that traditional, comprehensive systematic reviews struggle to provide. More succinct review methods including, mapping, rapid or scoping reviews could improve the quality of scientific evidence used to make statements in the WHO Blue Books, without the need for full systematic reviews.[21–23] While alternative methods that abbreviate systematic reviews may be appealing, they do increase the risk of bias and errors in the review process and may not always provide the expected results.[21,22,24] Careful assessment of the need for evidence synthesis may show that a structured expert description of the subject is enough. We believe that full systematic reviews of the available evidence should be limited to the assessment of research questions that directly affect the classification of a tumor, or have major consequences in cancer diagnosis.

Nowhere is the need to consider consensus-based *versus* evidence-based approaches more apparent than in the field of molecular pathology, where research groups may suggest different molecular tumor classifications for the same tumor type. In these instances, evidence synthesis and evaluations to underpin incorporation of molecular advances into the classification will be greatly assisted by the use of the Standards for Reporting Diagnostic Accuracy Studies (STARD),[25] Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or

Diagnosis (TRIPOD)[26] and other standardized reporting guidance (www.equator.net).

Herein lies the challenge for the WHO Classification of Tumours, currently published as a series of 15 WHO Blue Books, and now as a website. Should it continue to rely primarily on experts without formal systematic review methods, or should it take steps to incorporate systematic reviews and/or adapted methods into the process of using expert consensus? The importance of systematic reviews and expert interpretation do not need to be viewed as mutually exclusive. Mickenautsch argues that to view one as superior to another is erroneous as each represents necessary concepts of analysis and synthesis.[27] Indeed, there is an opportunity to ensure that reviews are fit for purpose without enforcing rigid guidelines which can result in review products which are overly detailed and unwieldy to interpret.[15] We propose an approach that is adapted to the needs of WHO Blue Books which requires authors to employ a series of "non-negotiables" when performing the literature reviews that feed into key decisions for the WHO Blue Book series. They are:

1. *Transparency*: clearly stated, prespecified methods are critical and should be described in detail to assure reproducibility and reliability of the process. A publicly available protocol setting out the intended methods and eligibility criteria would help receiving external input into the review and facilitate updating.
2. *Searching rigor*: at least two relevant bibliographic sources should be searched for relevant literature
3. *Double checking*: a proportion of randomly selected study selection and data extraction decisions should be double-checked by a second reviewer to verify reliability and consistency.
4. *Risk of bias assessment*: included individual studies should be assessed using standardized or adapted tools that consider areas of bias to which studies of pathology may be prone. Assessment of the body of the collected evidence may be useful, ranking it using validated criteria.

Detail on the rationale and risk underlying each "non-negotiable" are outlined in Table 2. The issue of using

Table 2. Methodological non-negotiables for systematic reviews for the purposes of tumor classification

Rationale	Risks (if not considered in the review)
1. *Transparency*	
Methods should be clearly stated and previously defined. Inclusion and exclusion criteria stated and applied. A review protocol should be written and made publicly available as an explicit statement of intended methods where deviations to these methods can be noted (with justifications). This ensures accountability by authors and facilitates replication of the review. Conflict of interests of the review team, as well as funding information, needs to be disclosed.	• Methods may not be appropriate to ensure equitable representation of literature globally • Unjustified deviations to planned methods remain unchallenged • Undeclared conflicts of interest or researcher allegiance from authors may influence conclusions
2. *Searching rigor*	
Searching two major bibliographic databases, (e.g., PubMed and Web of Science), minimizes the chance that a highly relevant study will be missed. While there are overlaps in medical bibliographic databases, indexing varies considerably. Therefore searching only one database means that retrieval of relevant literature is highly dependent on appropriateness/accuracy of the search strategy.	• Reliance on one database capturing all relevant studies, reliance on all relevant studies being accurately indexed and reliance on a single search strategy being sufficient to capture all relevant literature • Failure to identify all relevant literature
3. *Double checking*	
Duplication of the data extraction and a proportion of the total study selection done by the primary author should be completed by a second reviewer for accuracy. Where multiple discrepancies are noted, further checking may be required for consistency.	• Reliance on the accuracy and consistency of one author for all study selection and data extraction • Bias in selection of studies • Greater chance of erroneous study selection or data extraction
4. *Risk of bias assessment*	
A methodological quality assessment tool for pathology reviews should be adapted, based on standardized risk of bias assessment tool. This helps review authors to assign more weight to findings from studies of higher quality or at lower risk of bias in interpretation.	• No objective method of appraising studies for higher risk of bias • Biases from primary studies are perpetuated in the review • Bias in interpretation of studies may be applied by review authors

validated methods for WHO Classification of Tumours is becoming increasingly important because, due to rapid advances in molecular pathology, more information is becoming evident from closer analysis of tumors than was previously available from histopathological analysis alone. Systematic review methods may need to be adapted to assess this information, as some steps of a validated systematic review process may not meet current need, or simply not add sufficient value to the editorial process.

Not every decision made in the WHO Classification of Tumours is contentious; many areas may be regarded as close to fact due to strong logic or pathophysiological basis: they can be considered as equivalent of parachutes—known to be efficacious without the need to conduct a RCT.[28] However, when issues are debatable or likely to vary depending on context or interpretation then a systematic review may be deemed necessary. An example of this occurs with Kaposi's sarcoma where there is some debate over whether the characteristic tumors can reliably be regarded as neoplasms or as a "reactive" proliferations of cells in response to infections with human herpesvirus-8 (HHV-8).[29] Cancerous tumors are commonly regarded to be clonal neoplasms which proliferate uncontrollably, even after the stimulus which evoked the change has been removed,[30] Kaposi sarcoma is contentious because there are cases where it does not persist after successful antiretroviral treatment.[31] However, reclassification of Kaposi sarcoma could have ramifications for patient diagnosis and care, as well as service delivery and funding for research. In response to this dilemma, a systematic review protocol was designed and registered prospectively by the WHO tumor classification group on the PROSPERO database (CRD42018087595). The review aims to inform and support the decision of the expert panel regarding Kaposi Sarcoma classification using available evidence without restriction by study type.

There are challenges for the classification of tumors in evidence-based pathology. Differing evidence levels and considerations are applied within the field of pathology and consensus is needed to promote a comparable assessment of studies. Review methods should allow reliable assessments of evidence in an appropriate time frame and without being systematic review methodologists. Experts who serve as authors and editors require basic training and methodological support in systematic review and literature searching methods. Furthermore, subjects in need of assessment should be registered and prioritized based on the WHO Classification of Tumours requirements and scientific considerations. The WHO Classification of Tumours Group at the International Agency for Research on Cancer (IARC) have set themselves the challenge to revise their procedures to incorporate systematic review methods into the current expert-led approach through the appointment of a systematic reviewer for the Blue Books and by planning an international Joint Action Group. The methods described here aim to start this deliberation and we invite discussion and feedback as to the potential challenges and solutions to the intended "non-negotiables" proposed to ensure the methods for reviewing literature which feed into tumor classification are robust.

In conclusion, we believe that an evidence-based approach to informing key decisions that feed into tumor classification would allow editors to mitigate any potential risks of bias and also benefit authors by providing structured, transparent and reliable methods for the synthesis of available evidence for each tumor type. Our hope over time is that this approach will increase the rigor of the decisions feeding into WHO Classification of Tumour, by addressing critical questions and identifying research gaps, as well as reaching recommendations for research to inform future editions. Such an approach will help maintain the reliability of tumor classification, and help to provide solutions to challenges like the exponential rise in number of scientific publications and the need to manage new types of information such as evidence from genetics or big data.

Yours sincerely
Lesley Uttley
Blanca Iciar Indave
Chris Hyde
Valerie White
Dilani Lokuhetty
Ian Cree

Conflict of interest

Dr Lesley Uttley declares that she undertook two paid visiting scientist trips to the International Agency for Research on Cancer to plan and develop this work in July/August 2018. The other authors declare that they have no conflict of interest.

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.

References

1. WHO Classification of Tumours. *Digestive system tumours*, 5th edn. Lyon: IARC Publications, 2019.
2. Young SN. Bias in the research literature and conflict of interest: an issue for publishers, editors, reviewers and authors, and it is not just about the money. *J Psychiatry Neurosci* 2009;34:412–7.
3. McKee M, Priest P, Ginzler M, et al. How representative are members of expert panels? *Quality Assurance Health Care* 1991;3:89–94.
4. Doust J, Del Mar C. Why do doctors use treatments that do not work? *BMJ* 2004;328:474–5.
5. Kea B, Sun BC. Consensus development for healthcare professionals. *Intern Emerg Med* 2015;10:373–83.
6. Tajeddin G, Safayeni F, Connelly CE, et al. The influence of emergent expertise on group decision processes. *Small Group Res* 2011;43:50–74.
7. Mulrow CD. Rationale for systematic reviews. *BMJ* 1994;309:597–9.
8. Gotzsche PC, Ioannidis JP. Content area experts as authors: helpful or harmful for systematic reviews and meta-analyses? *BMJ* 2012;345:e7031.
9. Piehl JH, Green S, McDonald S. Converting systematic reviews to Cochrane format: a cross-sectional survey of Australian authors of systematic reviews. *BMC Health Serv Res* 2003;3:2.

10. Uttley L, Montgomery P. The influence of the team in conducting a systematic review. *Syst Rev* 2017;6:149.
11. Collins JA, Fauser BC. Balancing the strengths of systematic and narrative reviews. *Hum Reprod Update* 2005;11:103–4.
12. MacLure K, Paudyal V, Stewart D. Reviewing the literature, how systematic is systematic? *Int J Clin Pharm* 2016;38:685–94.
13. Higgins JPT, Green S. *Cochrane handbook for systematic reviews of interventions*. West Sussex, UK: The Cochrane Collaboration and John Wiley & Sons Ltd, 2008.
14. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097.
15. Greenhalgh T, Thorne S, Malterud K. Time to challenge the spurious hierarchy of systematic over narrative reviews? *Eur J Clin Invest* 2018;48:e12931.
16. Newman TB. The power of stories over statistics. *BMJ* 2003;327:1424–7.
17. Williams DD, Garner J. The case against "the evidence": a different perspective on evidence-based medicine. *Br J Psychiatry* 2002;180:8–12.
18. Parker M. False dichotomies: EBM, clinical freedom, and the art of medicine. *Med Humanit* 2005;31:23–30.
19. Loke YK, Derry S. Does anybody read "evidence-based" articles? *BMC Med Res Methodol* 2003;3:14.
20. Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q* 2016;94:485–514.
21. Dobbins M. *Rapid review guidebook. Steps for conducting a rapid review [Guideline]*. Hamilton: The National Collaborating Centre of Methods and Tools (NCCMT), 2017 Available from: https://www.nccmt.ca/uploads/media/media/0001/01/a816af720e4d587e13da6bb307df8c907a5dff9a.pdf.
22. Khangura S, Konnyu K, Cushman R, et al. Evidence summaries: the evolution of a rapid review approach. *Syst Rev* 2012;1:10.
23. Polisena J, Garritty C, Umscheid CA, et al. Rapid review summit: an overview and initiation of a research agenda. *Syst Rev* 2015;4:111.
24. Haby MM, Chapman E, Clark R, et al. What are the best methodologies for rapid reviews of the research evidence for evidence-informed decision making in health policy and practice: a rapid review. *Health Res Policy Syst* 2016;14:83.
25. Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 2016;6:e012799.
26. Moons KG, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162:W1–73.
27. Mickenautsch S. Systematic reviews, systematic error and the acquisition of clinical knowledge. *BMC Med Res Methodol* 2010;10:53.
28. Smith GC, Pell JP. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. *BMJ* 2003;327:1459–61.
29. Brooks JJ. Kaposi's sarcoma: a reversible hyperplasia. *Lancet (London, England)* 1986;2:1309–11.
30. Willis RA. Pathology of Tumours. By R. A. Willis, D.Sc., M.D., F.R.C.P., and Sir William H. Collins Professor of Human and Comparative Pathology, Royal College of Surgeons, London. $9^{5/8} \times 6^{5/8}$ in. Pp. 1044 + xxiii, with 500 illustrations. 1948. London: Butterworth & Co. (Publishers) Ltd. 63s. net. *BJS* 1948;35:446.
31. Chang Y, Cesarman E, Pessin MS, et al. Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science (New York, NY)* 1994;266:1865–9.

DOI: 10.1002/ijc.32975

History: Received 16 Dec 2019; Accepted 14 Feb 2020; Online 13 Mar 2020

Correspondence to: Blanca Iciar Indave, E-mail: indavei@iarc.fr