BJA

British Journal of Anaesthesia, 123 (1): e29–e37 (2019)

doi: 10.1016/j.bja.2019.02.029
Advance Access Publication Date: 24 April 2019
Special Article

Consensus clinical scoring for suspected perioperative immediate hypersensitivity reactions

P. M. Hopkins[1,2,*,†,#], P. J. Cooke[3,#,†], R. C. Clarke[4,5,#,†], A. B. Guttormsen[6,7,#,†], P. R. Platt[4,5,#,†], P. Dewachter[8,9,†], D. G. Ebo[10,†], T. Garcez[11,†], L. H. Garvey[12,13,†], D. L. Hepner[14,†], D. A. Khan[15,†], H. Kolawole[16,17,†], P. Kopac[18,†], M. Krøigaard[12,†], J. J. Laguna[19,†], S. D. Marshall[16,17,†], P. M. Mertes[20,†], M. A. Rose[21,†], V. Sabato[10,†], L. C. Savic[2,†], S. Savic[22,†], T. Takazawa[23,†], G. W. Volcheck[24,†], S. Voltolini[25,†] and P. H. M. Sadleir[4,5,26,#,†]

[1]Leeds Institute of Medical Research at St James's, University of Leeds, Leeds, UK, [2]Anaesthetic Department, Leeds Teaching Hospitals NHS Trust, Leeds, UK, [3]Department of Anaesthesia and Perioperative Medicine, Auckland City Hospital, Auckland, New Zealand, [4]Department of Anaesthesia, Sir Charles Gairdner Hospital, Nedlands, Western Australia, Australia, [5]Anaesthetic Allergy Referral Centre of Western Australia, Nedlands, Western Australia, Australia, [6]Department of Anaesthesia and Intensive Care, Haukeland University Hospital, Bergen, Norway, [7]Department of Clinical Medicine, University of Bergen, Bergen, Norway, [8]Service d'Anesthésie-Réanimation, Groupe Hospitalier de Paris-Seine-Saint-Denis, Assistance Publique-Hôpitaux de Paris, Paris, France, [9]Université Paris 13, Sorbonne-Paris-Cité, Paris, France, [10]Department of Immunology, Allergology and Rheumatology, University of Antwerp, Antwerp University Hospital, Belgium, [11]Department of Immunology, Manchester University NHS Foundation Trust, Manchester, United Kingdom, [12]Danish Anaesthesia Allergy Centre, Allergy Clinic, Department of Dermatology and Allergy, Copenhagen University Hospital, Gentofte, Denmark, [13]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark, [14]Department of Anesthesiology, Brigham and Women's Hospital, Boston, MA, USA, [15]Department of Internal Medicine, Division of Allergy and Immunology, University of Texas Southwestern Medical Center, Dallas, TX, USA, [16]Department of Anaesthesia and Perioperative Medicine, Monash University, Melbourne, Australia, [17]Department of Anaesthesia, Peninsula Health, Melbourne, Australia, [18]University Clinic of Respiratory and Allergic Diseases Golnik, Golnik, Slovenia, [19]Allergy Unit, Allergo-Anaesthesia Unit, Hospital Central de la Cruz Roja, Faculty of Medicine, Alfonso X. El Sabio University, ARADyAL, Madrid, Spain, [20]Department of Anesthesia and Intensive Care, Hôpitaux Universitaires de Strasbourg, Nouvel Hôpital Civil, Strasbourg, France, [21]Department of Anaesthesia, Royal North Shore Hospital, University of Sydney, Sydney, NSW, Australia, [22]Department of Clinical Immunology and Allergy, Leeds Teaching Hospitals NHS Trust, Leeds, UK, [23]Intensive Care Unit, Gunma University Hospital, Maebashi, Gunma, Japan, [24]Division of Allergic Diseases, Mayo Clinic College of Medicine, Rochester, MN, USA, [25]Allergy Unit, Policlinic Hospital San

Accepted: 28 February 2019

© 2019 British Journal of Anaesthesia. Published by Elsevier Ltd. All rights reserved.
For Permissions, please email: permissions@elsevier.com

Martino, Genoa, Italy and [26]Department of Pharmacology, University of Western Australia, Crawley, Western Australia, Australia

*Corresponding author. E-mail: p.m.hopkins@leeds.ac.uk

#These authors contributed equally.
†The authors are members of the ISPAR (International Suspected Perioperative Allergic Reaction) Group.

Abstract

Background: Grading schemes for severity of suspected allergic reactions have been applied to the perioperative setting, but there is no scoring system that estimates the likelihood that the reaction is an immediate hypersensitivity reaction. Such a score would be useful in evaluating current and proposed tests for the diagnosis of suspected perioperative immediate hypersensitivity reactions and culprit agents.
Methods: We conducted a Delphi consensus process involving a panel of 25 international multidisciplinary experts in suspected perioperative allergy. Items were ranked according to appropriateness (on a scale of 1–9) and consensus, which informed development of a clinical scoring system. The scoring system was assessed by comparing scores generated for a series of clinical scenarios against ratings of panel members. Supplementary scores for mast cell tryptase were generated.
Results: Two rounds of the Delphi process achieved stopping criteria for all statements. From an initial 60 statements, 43 were rated appropriate (median score 7 or more) and met agreement criteria (disagreement index <0.5); these were used in the clinical scoring system. The rating of clinical scenarios supported the validity of the scoring system. Although there was variability in the interpretation of changes in mast cell tryptase by the panel, we were able to include supplementary scores for mast cell tryptase.
Conclusion: We used a robust consensus development process to devise a clinical scoring system for suspected perioperative immediate hypersensitivity reactions. This will enable objectivity and uniformity in the assessment of the sensitivity of diagnostic tests.

Keywords: allergy; anaesthesia; anaphylaxis; drug hypersensitivity; perioperative period; surgery

Editor's key points

- A panel of 25 international multidisciplinary experts in suspected perioperative allergy used a Delphi consensus process to develop a clinical scoring system for suspected perioperative immediate hypersensitivity reactions.
- Of 60 initial statements, 43 were rated appropriate and met agreement criteria for inclusion in the clinical scoring system, which included supplementary scores for mast cell tryptase levels.
- Rating of clinical scenarios supported the validity of the scoring system.
- This new clinical scoring system should be useful for diagnosis of suspected perioperative immediate hypersensitivity reactions and for assessing the sensitivity of diagnostic tests.

Adverse perioperative events that meet published criteria for suspected immediate hypersensitivity reactions (IHRs) have been reported in up to one per353 general anaesthetics.[1,2] The clinical diagnosis of an IHR (allergic or non-allergic) is difficult in the perioperative patient because many of the clinical features occur frequently at various grades of severity through non-immune mechanisms. In addition, patients under general anaesthesia are unable to report symptoms.[3] If an IHR is diagnosed, identifying the culprit agent can be difficult because of the routine almost simultaneous exposure of multiple potential culprits.[4] The diagnosis of an IHR in the perioperative period is important because it has implications for the provision of safe anaesthesia for the patient in the future. Furthermore, having identified that a patient has had an IHR, identification of the mechanism and culprit agent along with safe alternative drugs within the same class of the culprit is required to enable the goal of safe future anaesthesia.

Guidelines for the investigation of suspected perioperative IHRs emphasise the need to combine clinical information, measurement of biomarkers of acute allergic responses, and skin testing.[5–10] *In vitro* tests to improve diagnosis[11–13] are reviewed in detail elsewhere in this issue of the *British Journal of Anaesthesia*.[14,15] A key requirement for the interpretation of any test is an understanding of its accuracy.[16] The accuracy of a test is described most simply in terms of its sensitivity (the proportion of truly positive patients or samples that have a positive test) and specificity (the proportion of truly negative patients or samples that have a negative test). Calculation of sensitivity and specificity with different cut-off values can be used to determine the optimum cut-off values for diagnosis. In combination with an estimate of *a priori* likelihood of a condition, sensitivity and specificity can be used to calculate the positive and negative predictive values of a test.

To estimate the sensitivity of any test to confirm an IHR or identify a culprit agent, it is necessary to evaluate the test in patients known to have had an IHR, that is true positives. To be an unbiased evaluation, identification of true positive cases should be independent of the results of the test or related tests

to avoid circular arguments.[17–20] This requires an objective approach to identifying true perioperative IHRs with a high degree of likelihood based on clinical information alone. Some researchers have used classification systems of allergic reactions for this purpose,[21–25] mostly based on the Ring and Messmer[26] classification. There are newer systems proposed by Niggemann and Beyer[27] (primarily for food allergy) and, specifically for perioperative reactions, by Rose and colleagues[28] and Cook and colleagues[29] to classify severity of reactions. However, none of these classification systems describe how likely a reaction is to be an IHR. Indeed, the assumption underlying such classification systems is that the patient is having an allergic reaction because there is no likely alternative explanation. This is a reasonable assumption in the absence of all of the potential confounding factors present in the perioperative period. For example, no account is taken for alternative causes of bronchospasm or hypotension[30] with classification systems derived from Ring and Messmer.[26] Therefore, we aimed to generate a clinical scoring system to assess the likelihood of an adverse event in the perioperative period being an IHR.

Methods

Although this paper does not represent development of a guideline *per se*, the methodology shares several aspects of guideline development. We therefore used the AGREE (Appraisal of Guidelines, Research, and Evaluation) checklist,[31] where relevant, to advise our approach.

Panel selection

An international multidisciplinary panel of allergists, anaesthetists, and immunologists with a track record of publication in the field of perioperative anaphylaxis was formed. From within the panel, a 'writing group' was created from those members of the panel expressing a specific interest in taking an additional role with this project.

Literature search

We used the PICO (population, intervention, comparators, outcomes) framework to formulate our literature search strategy as follows:

Population/problem: patients undergoing an operative procedure for diagnosis or treatment involving care from an anaesthetist.

Intervention: diagnosis of suspected IHR in the perioperative period.

Comparators: confounding factors for diagnosis of suspected IHR.

Outcomes: clinical diagnosis, classification, or grading of suspected perioperative IHR.

We searched PubMed and Embase databases and included publications from 1997 to [actual date] and key publications (first reports of paradigms that remain central to the PICO criteria) before 1997.

Modified Delphi process

We adopted the approach of Fitch and colleagues[32] in which statements are rated for appropriateness on a scale of 1 (completely inappropriate) to 9 (completely appropriate). Disagreement was determined using the disagreement index (DI), where the lower the value below 1, the greater is the consensus, and values >1 are considered to represent lack of consensus.[33] The median appropriateness score was used to rate each statement as inappropriate (median score 1–3.4), uncertain (median score 3.5–6.9), or appropriate (median score 7–9). We planned at least two rounds to generate a series of statements rated as appropriate with a clear consensus (DI <0.5). The process was to continue until a clear consensus was reached for each statement (DI <0.5) or the DI failed to improve by more than 15% in successive rounds.[34] The Delphi process was managed by the convener of the writing group (PMH); all other members of the panel were invited to participate in each round and were given at least 2 weeks to respond.

Round 1

A series of statements describing clinical manifestations of suspected IHRs was generated by the writing group based on relevant publications identified from the literature search and their clinical experience. The statements were sent to panel members using an online questionnaire tool (Google forms), in which panel members were asked to rate each statement on the appropriateness scale (1–9). Panel members had the option of responding N/A (not applicable) to statements that they felt to be outside their expertise. Panel members were also invited to provide freehand comments on the wording of existing statements or to propose new statements.

Round 2 and subsequent rounds

Before Round 2, panel members received their scores from Round 1 alongside de-identified scores of the other panel members (as raw data and as summary bar charts), and the calculated median appropriateness and DI values. Information on interpretation of median appropriateness and DI values was provided. Median appropriateness values were also calculated separately for panel members who were anaesthetists and those who were either allergists or immunologists and these values were also circulated to panel members.

In generating the statements for Round 2, the writing group reviewed the responses to the statements in Round 1, including freehand comments, and agreed whether each statement should be included unchanged, included in amended form, or not included in Round 2. The revised statements were formatted as an online questionnaire as for Round 1, which the panel members were invited to complete. If the stopping criteria were not met after Round 2, the process for subsequent rounds would follow that of Round 2.

Generation of the clinical scoring system

The results of the final round of the Delphi process were used to rank clinical features as to their contribution to predicting the likelihood of an IHR based first on the median appropriateness rating and then on the DI. These rankings were used to assign points within the scoring system, such that clinical features increasing the likelihood of an IHR were assigned positive values and those decreasing the likelihood (confounding features) were assigned negative values. The relative points allocation within positive and negative categories was made on the basis of the Delphi rankings supplemented by the clinical experience of the writing group that agreed on the initial scoring scheme. The content validity of the scoring scheme was initially assessed subjectively by the writing group before

testing for criterion and discriminant validity using the whole panel. For this exercise, a series of hypothetical case scenarios of suspected perioperative IHRs was developed, and panel members were asked to independently rate the likelihood of the case as 'almost certain', 'very likely', 'likely', or 'unlikely' to be an IHR. The case scenarios were compiled by the writing group convener (PMH) and were designed to evaluate how experts assessed the relative discriminant ability of items within and between scoring system categories and their combination. Minor adjustments of the points allocation within the scoring system were made in order to maximise its discriminant validity before the median likelihood ratings of the panel were used to calibrate the scoring system.

In addition to asking panel members to rate the case scenarios on clinical features alone, they were also asked to rate the scenarios when accompanied by mast cell tryptase results. This intended to assess how experts assessed: (a) 'borderline' tryptase increase; (b) the impact of no or minimal tryptase change on their evaluation of a clinical scenario with relatively high likelihood of being an IHR; and (c) the impact of a large tryptase increase on their evaluation of a clinical scenario with relatively low likelihood of being an IHR. These responses were used to produce and calibrate a scheme for supplementing the clinical scoring system when tryptase results are available (and assuming that the purpose of generating the score is not to evaluate the sensitivity of tryptase changes themselves).

Results

We approached by e-mail 33 international experts in suspected perioperative allergic reactions of which 18 were anaesthetists, 14 allergists or immunologists, and one dually accredited in anaesthesia and allergy. Of these, 15 anaesthetists, nine allergists/immunologists, and the dually accredited colleague agreed to participate. The affiliations of panel members are provided in the list of authors. The six members of the writing group are all anaesthetists.

Delphi process

From the review of the literature (literature search terms and results are provided in Supplementary Appendix 1) and their clinical experience, the writing group generated a list of 60 statements to be used in Round 1. Twenty-three of 24 members (96%) of the panel responded (the final panel member, PMH, managed the Delphi process): 39 of the statements were rated as appropriate, 20 as of uncertain appropriateness, and one inappropriate. The DI was <0.5 for 41 statements, 0.5–1 for 18 statements, and >1 for one statement ('Patients with a history of allergy are at increased risk of developing an IHR in the perioperative period'). This latter statement was one of only eight statements where the median appropriateness scores for anaesthetists differed by more than 2 from that of non-anaesthetists (Supplementary Appendix 2). Panel members contributed a total of 17 freehand comments in Round 1 although no completely new statements were proposed.

In Round 2, 32 of the statements were unchanged from Round 1, 17 statements were amended, and 11 statements were excluded with 24 members (100%) of the panel responding. Supplementary Appendix 3 shows the Round 2 statements ranked in order of highest median appropriateness score and then by the lowest DI. All statements met one or other stopping criteria for the iterative Delphi process. All but six of the statements had a median appropriateness score of 7 or more and a DI <0.5. The remaining statements were considered for use in construction of the clinical scoring system.

From Supplementary Appendix 3 it can be seen that clinical features associated with the cardiovascular system, respiratory system, and skin or mucous membranes were perceived to have value in predicting the likelihood of a perioperative IHR. Within each of these systems several confounding factors were identified that reduced the likelihood of a perioperative IHR (Supplementary Appendix 3). Supplementary Appendix 3 also highlights the high ratings for appropriateness and consensus for co-occurrence of features from more than one system. The other aspect that the writing group reflected in the initial clinical scoring system was the timing of the onset of clinical features in relation to administration of a potential culprit agent.

In transforming the consensus statements into the clinical scoring system, we realised that clinical terms needed to be defined so that the scoring system had construct validity and could be applied reproducibly. The writing group developed a series of definitions of clinical features and tested these for appropriateness with a single round Delphi process involving all panel members. Table 1 shows the definitions agreed and the high level of appropriateness and consensus of the panel for these definitions in this context.

The writing group structured the scoring system based on key areas of consensus from the Delphi process. These were: positive and confounding features within each of cardiovascular, respiratory, and dermal/mucosal categories; the added weight of combinations of features from more than one of these categories; the importance of timing of onset of features in relation to exposure to potential triggers, except for dermal or mucosal features. The writing group agreed on a provisional scoring system before conducting a validity-testing exercise involving the whole panel. The clinical scenarios used in this exercise are presented in Supplementary Appendix 4 along with the ratings of the panel members presented for the whole group and also for anaesthetists separately.

The writing group used the feedback from the clinical scenario ratings of panel members to make minor adjustments to the clinical scoring system while maintaining the principles derived from the initial consensus exercise. The final clinical scoring system is shown in Table 2. The median clinical scenario ratings were used to calibrate the clinical scoring system by converting scoring ranges to indicate almost certain, very likely, likely, or unlikely IHRs. During writing of the manuscript it was agreed to subdivide the 'likely' category into 'likely' and 'possible', as we think this will aid clinical utility. The likelihood categories are shown in Table 3.

In order to incorporate changes in mast cell tryptase concentration into the clinical scoring system, we evaluated the impact of various tryptase changes on the clinical likelihood rating by panel members. Ratings are shown in Supplementary Appendix 4. If the peak tryptase after a suspected IHR showed no change from the baseline value, most panel members considered this to have a negative impact on their assessment of the likelihood of an IHR. A change in tryptase of (1.2×baseline)+2 ng ml^{-1} with the peak tryptase remaining within the reference range was considered a better indicator of a likely IHR than a smaller relative change even if the peak tryptase was outside the reference range (>upper 95% confidence limit of the reference range). If a relative change of (1.2×baseline)+2 ng ml^{-1} was combined with a peak value greater than the upper limit of the reference range, tryptase level was considered to have a greater impact on likelihood of an IHR. An even greater relative change combined with the

Table 1 Definitions for clinical terms used in the clinical scoring scheme. DI, disagreement index; Median, median appropriateness score.

Clinical term	Definition	Median	DI
Hypotension	A fall in systolic blood pressure to <70 mm Hg (at induction or during maintenance of anaesthesia) or by >20% from a previously stable value (during maintenance of anaesthesia)	8	0.140
Severe hypotension	A fall in systolic blood pressure to <60 mm Hg (at induction or during maintenance of anaesthesia) or by >40% from a previously stable value (during maintenance of anaesthesia)	8	0.132
Cardiac arrest	The requirement for cardiopulmonary resuscitation not explained by the surgical pathology, complications of the surgical procedure, co-existing medical problems or drugs, malignant hyperthermia or technical anaesthetic problems	8	0.292
Tachycardia	An otherwise unexplained increase in heart rate of 50% or more from a previously stable value	8	0.074
Bronchospasm	The onset of wheeze on auscultation, any manifestation of otherwise unexplained increased airway resistance, or both	8	0.074
Severe bronchospasm	Bronchospasm associated with SpO_2 <85%	7.5	0.164
Urticaria	A skin rash characterised by raised pink or white raised areas of skin (wheals)	9	0.132
Angioedema	Dermal or mucosal swelling	8.5	0.132

peak being outside the normal range had the greatest impact. These rankings were used to produce an algorithm for increasing points allocation to tryptase changes to supplement the clinical scoring system, when appropriate (Table 4).

Discussion

We used an established methodological approach to generate consensus from an international multidisciplinary panel of experts in suspected perioperative allergic reactions for clinical criteria that have predictive value for estimating the likelihood that an adverse perioperative event was the result of an IHR. We used the ranking of appropriateness and consensus of the criteria to construct a clinical scoring system and went on to ensure its content, construct, criterion, and discriminant validity.

One of the key differences between previously published classification systems and our clinical scoring system is that we have enabled the impact of potential confounding factors and the time interval between potential culprit exposure and onset of signs to be assimilated. Although this increases the complexity of the final scoring system, it reflects the complexity that can be involved in forming an expert clinical judgement of the potential cause of an adverse perioperative event. The need to exclude other causes of suspected adverse drug reactions is an accepted and integral part of causality assessment used in pharmacovigilance.[35] Our validity assessments suggest that the scoring system will be able to identify with high likelihood IHRs that present with relatively subtle features involving two or more systems and IHRs with more severe features confined to a single system. The scoring system also implicitly reflects the expert consensus that timing of skin manifestations is a poor discriminator as these may be obscured by surgical drapes or delayed in appearance until a shocked patient has been resuscitated.

The value of the availability of a clinical scoring system for rare perioperative adverse reactions has been demonstrated by the enduring use of the Larach clinical grading scale for malignant hyperthermia which was developed using a Delphi consensus approach.[36] This has been used to great effect to evaluate the sensitivity of the two principally applied protocols for the laboratory diagnosis of malignant hyperthermia susceptibility[37,38] and in studies of the epidemiology of malignant hyperthermia.[39,40] As with our scoring system for IHRs, the Larach clinical grading scale was not intended for use in real-time clinical diagnosis, which for both IHR and malignant hyperthermia should be based on early pattern recognition of clinical features and rapid evaluation of differential diagnoses with a relatively low threshold for initiating treatment.

Implementation of the IHR clinical scoring system requires experience of interpretation of perioperative records, including anaesthetic charts, in order to accurately extract the data needed. Our recommendation is that this is done by an individual with the necessary expertise who was not involved directly with the case in order to minimise subconscious bias. The relevant and sufficient information to apply the scoring system to cases of suspected perioperative allergic reactions should be routinely available when patients are assessed in a specialist anaesthetic allergy clinic setting. However, the scoring should be done blinded to the results of subsequent investigations to avoid hindsight bias.

The definitions of various clinical terms, such as hypotension, bronchospasm, and tachycardia, that we have adopted for use in the clinical scoring system (Table 1) are intended to maximise the utility of the scoring system. Using hypotension as an example, our definitions differ from the physiological definition, definitions used in the context of allergy in general[41] and even definitions used elsewhere in the context of perioperative allergy.[42,43] It is inevitable that our definitions will exclude clinical features that occur in some true IHRs from contributing to the score for that reaction. It is our collective view that such subtle changes in the perioperative context have too low a predictive value for our purpose. Similarly, although a low end-tidal CO_2 has been shown to be a superior predictor of the severity of an IHR for hypotension,[44] our expert consensus was that this sign did not add to the discriminant ability of hypotension and bronchospasm to distinguish between hypersensitivity and non-hypersensitivity reactions, while potentially introducing additional confounders such as iatrogenic hyperventilation, hypothermia, pulmonary embolus, or right-to-left shunt.

A potential advantage of using a scoring system generated by expert consensus is that it is likely to reduce the potential

Table 2 The clinical scoring system. Items contributing to the clinical score for suspected perioperative immediate hypersensitivity reactions (IHRs). Points are awarded within five categories, with features suggestive of an IHR (pink cells) having positive points values and features against an IHR (green cells) having negative points values. How points may be allocated to items is indicated for each category. The overall clinical score is the sum of the net scores of all categories. *For a score from one of the three organ systems, cardiovascular (CVS), respiratory (RS), dermal/mucosal (D/M) to contribute to a combination score, the net score for that system must be >2. The net score is the sum of scores for positive features minus the sum of scores for confounders within scores for that system. For definitions, see Table 1.

1. Cardiovascular (choose hypotension, severe hypotension, or cardiac arrest if appropriate, then any other items that apply)	**Points**
Hypotension	4
Severe hypotension	6
Cardiac arrest	9
Tachycardia	2
A poor or unsustained response of hypotension to standard doses of sympathomimetics used to treat pharmacological hypotension during anaesthesia (e.g. ephedrine, phenylephrine, metaraminol)	2
A point-of-care echocardiogram showing a hyperdynamic and poorlyfilled heart	2
Recurrence or worsening of hypotension after a further dose of a drug given before the initial event	1
Cardiovascular confounders (in the presence of hypotension or cardiac arrest choose any that apply)	
Excessive dose of anaesthetic drug or drugs	–2
Surgically induced hypovolaemia or relative hypovolaemia from prolonged fasting/dehydration	–1
Acute illness predisposing to hypotension	–1
Medications affecting cardiovascular responses during anaesthesia	–2
Neuraxial regional anaesthesia (epidural/spinal)	–1
Onset of hypotension after development of increased peak airway pressure during mechanical ventilation of the lungs	–2
2. Respiratory (choose bronchospasm or severe bronchospasm if appropriate, then any other items that apply)	
Bronchospasm	2
Severe bronchospasm	4
Recurrence or worsening of bronchospasm after a further dose of a drug given before the initial event	1
Bronchospasm occurring before airway instrumentation (having excluded airway obstruction)	2
Respiratory confounders (in the presence of bronchospasm choose any that apply)	
Respiratory disease associated with reactive airways	–1
Prolonged or multiple attempts at tracheal intubation	–1
Inadequate dose of drugs to obtund airway responses before airway instrumentation	–1
3. Dermal/mucosal (choose any items that apply)	
A generalised rash is itchy in the awake patient who has not received epidural/spinal opioids	1
Angioedema	3
Generalised erythema	3
Generalised urticaria	4
Dermal/mucosal confounder	
Angioedema in a patient taking an ACE inhibitor	–3
4. Combinations (choose a maximum of one item)*	
CVS>2 and RS > 2	5
CVS>2 and D/M >2	5
RS>2 and D/M >2	5
CVS>2 and RS>2 and D/M >2	8
5. Timings (choose a maximum of one item)	
Onset of cardiovascular or respiratory features within 5 min of possible IV trigger	7
Onset of cardiovascular or respiratory features within 15 min of possible IV trigger	3
Onset of cardiovascular or respiratory features within 60 min of possible non-IV trigger	2
Onset of cardiovascular or respiratory features more than 60 min after possible non-IV trigger	–1

inter-rater variability inherent in forming an assessment of causality from an unstructured review of clinical information. The 6th National Audit Project (NAP6) of the Royal College of Anaesthetists addressed this issue by using a large multidisciplinary panel to assess each potential case of anaphylaxis.[29,45,46] Although we have not formally assessed inter-rater variability for application of the clinical scoring system, our validity exercise demonstrated the variability of an opinion-based assessment of some relatively straightforward clinical scenarios. We had anticipated that this variability would be greatest when comparing anaesthetists and non-anaesthetists. However, on the whole this was not the case with within-specialty variability being similar to between-specialty variability; this is likely to reflect the common factor of expertise in perioperative allergy.

Our evaluation of expert opinion of the interpretation of changes in mast cell tryptase indicates that uncertainty persists in how such changes impact on the clinical evaluation of suspected perioperative IHRs. The majority of laboratories use the same supplier for mast cell tryptase testing kits and

Table 3 Clinical grading scale for interpretation of clinical score for suspected perioperative immediate hypersensitivity reactions (IHRs).

Interpretation	Total (net) score
Almost certain to be an IHR	>21
Very likely to be an IHR	15–21
Likely to be an IHR	11–14
Possible IHR	8–10
Unlikely to be an IHR	<8

Table 4 Algorithm for allocating points for mast cell tryptase changes to supplement the clinical scoring system. Points should be subtracted from or added to the net score from the clinical scoring system (Table 2) with the resulting score interpreted as defined in Table 3. Criteria for mast cell tryptase changes: (a) Formula +ve: Peak tryptase is >[(1.2×baseline tryptase)+2 ng ml^{-1}]; (b) Formula −ve: Peak tryptase is < [(1.2×baseline tryptase)+2 ng ml^{-1}]; (c) ULN: upper 95% confidence limit of the reference range (11.4 ng ml^{-1}); and (d) >2×BL: peak tryptase is >2×baseline tryptase.

Mast cell tryptase change	Points
No criteria	−4
Formula −ve but > ULN	−2
Formula +ve and < ULN	0
Formula +ve and > ULN	4
>2×BL and >ULN	12

reagents. The test has a low coefficient of variation with a high level of reproducibility between laboratories.[47] This makes it even more surprising perhaps that there is not better agreement on the interpretation of acute changes in the perioperative period. One of the issues may be the lack of robust estimates for the sensitivity and specificity of mast cell tryptase changes in suspected perioperative allergic reactions. For many years, it was assumed that if the peak tryptase in the 1–2 h after a suspected perioperative allergic reaction was within the normal reference range, then the tryptase result was 'negative'. In the meantime, Brown and colleagues[48] investigated tryptase changes in volunteers in whom allergic reactions were provoked in a controlled experimental setting with venom. Such studies showed that relative change from baseline was perhaps more important in detecting mast cell activation than the absolute value of the peak tryptase concentration. Garvey and colleagues[49] found that the upper 95% confidence interval for relative change in tryptase during elective orthopaedic surgery was 39%. A consensus process was used to develop a criterion for mast cell activation based on the principle of relative change.[47] It is clear from the responses of our expert panel to the hypothetical tryptase changes presented alongside clinical scenarios that not all expert opinion is confident that the use of this formula in the perioperative setting is discriminatory. Egner and colleagues[50] conducted perhaps the largest evaluation of mast cell tryptase in suspected perioperative allergic reactions. Their data, although having to rely on the Ring and Messmer[26] classification, suggest that smaller changes in tryptase in the perioperative setting may indeed be relevant if the sensitivity of tryptase changes is to be optimised. Baretto and colleagues[51] produced similar findings but used the World Allergy Organisation criteria[52] for identifying their 'true positive' cases, which again do not account for confounding factors. We propose that evaluation of tryptase changes in a large cohort of patients categorised as 'almost certain' by our clinical scoring system would provide the best estimate to date of the sensitivity of tryptase changes in identifying perioperative IHRs. We should emphasise that the time of sampling for peak tryptase (ideally 1–2 h after onset of the reaction) is extremely important, especially when considering discrete increases.

Limitations

Although we have demonstrated several aspects of the validity of the scoring system, independent external validation was not possible within the constraints of this project. The main purpose of external validation of such a tool is to ensure that it is generalisable, but we expect that inclusion of global representation on our expert panel makes generalisability of the scoring system likely. One possible means of independent validation of the scoring system would be to utilise the NAP6 cases and compare their scores with the ratings of the NAP6 panel.[29,46] A further potential limitation is that we do not expect the clinical scoring system to be reliable when relevant clinical information is missing, emphasising the necessity to include copies of all perioperative records when referring a patient with a suspected IHR for investigation.[7,8,45]

When applying the clinical scoring system to evaluate the sensitivity of mast cell tryptase changes or skin test results, the score makes no presumption about the mechanism of the suspected IHR. This means that one can evaluate a test for its sensitivity to detect an IHR but not IHRs with a defined mechanism (allergic or non-allergic). Therefore, any test that can identify only IHRs with an allergic mechanism, for example, may not achieve 100% sensitivity to detect IHRs even though it has 100% sensitivity to detect allergic reactions. We can only guess what proportion of IHRs are allergic because mast cell tryptase changes and skin test results have been used to define a reaction as allergic, even in the absence of a clear clinical history of an IHR. We now know that both mast cell tryptase and skin tests can be 'positive' through non-allergic and even non-immune mechanisms.[53–56] From a pragmatic clinical perspective we need to know the sensitivity of tests to detect an IHR of any mechanism, because non-allergic and allergic IHRs can occur with re-exposure to the culprit agent.

Conclusions

Our clinical scoring system, with or without the incorporation of tryptase results as appropriate, has the potential to better assess the sensitivity of currently used tests that are intended to confirm that an IHR has occurred and the agent responsible. It can also provide a consistent framework for the evaluation in research settings of proposed new tests. A robust estimate of sensitivity of skin tests, for example, will also aid interpretation of investigations of cross-reactivity of chemically and pharmacologically related agents.

Authors' contributions

Study design: PMH, PC, ABG, PS, RC, PP.
Writing paper: PMH, PC, ABG, PS, RC, PP.
Revising paper and approval of final version: all authors.

Declarations of interest

PD has received lecture and travel fees from MSD France (Courbevoie, France); lecture and travel fees from Bracco Imaging France (Courcouronnes, France); Agence Nationale de Sécurité du Médicament et des Produits de Santé (Saint-Denis, France); expert for a task force group dedicated to 'neuromuscular blocking agents and anaphylactic reactions' (until 2016); and is a member of MSD Expert Board on 'neuromuscular blocking agents and fast-tracking anesthesia' (until October 2019). LHG is a consultant and adjudication committee member for Merck, NJ, USA and Novo Nordisk Denmark. PMH is an editorial board member of *British Journal of Anaesthesia*. PMM is a scientific advisor for the ALPHO study (NCT02250729), funded by a consortium of pharmaceutical companies: Zambon, Urgo, Pierre Fabre, Boots, Hepatoum, Biocodex, Sanofi, LBR, GSK, APL, Bells Healthcare, Pinewood, T & R, and Ernest Jackson. PK has received lectures fees from Novartis Pharma Services Inc. and Shire Pharmaceuticals Group Plc.

Funding

SM is in receipt of an Australian National Health and Medical Research Council (NHMRC) ECR Fellowship for the investigation of cognitive aids in emergencies.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bja.2019.02.029.

References

1. Savic LC, Kaura V, Yusaf M, et al. Incidence of suspected perioperative anaphylaxis: a multicenter snapshot study. *J Allergy Clin Immunol Pract* 2015; **3**: 454–455 e1
2. Berroa F, Lafuente A, Javaloyes G, et al. The incidence of perioperative hypersensitivity reactions: a single-center, prospective, cohort study. *Anesth Analg* 2015; **121**: 117–23
3. Ebo DG, Fisher MM, Hagendorens MM, Bridts CH, Stevens WJ. Anaphylaxis during anaesthesia: diagnostic approach. *Allergy* 2007; **62**: 471–87
4. Hunting AS, Nopp A, Johansson SG, Andersen F, Wilhelmsen V, Guttormsen AB. Anaphylaxis to patent blue V: I. Clinical aspects. *Allergy* 2010; **65**: 117–23
5. Kolawole H, Marshall SD, Crilly H, Kerridge R, Roessler P. Australian and New Zealand anaesthetic allergy group/ Australian and New Zealand College of anaesthetists perioperative anaphylaxis management guidelines. *Anaesth Intensive Care* 2017; **45**: 151–8
6. Scolaro RJ, Crilly HM, Maycock EJ, et al. Australian and New Zealand anaesthetic allergy group perioperative anaphylaxis investigation guidelines. *Anaesth Intensive Care* 2017; **45**: 543–55
7. Harper NJ, Dixon T, Dugue P, et al. Suspected anaphylactic reactions associated with anaesthesia. *Anaesthesia* 2009; **64**: 199–211
8. Ewan PW, Dugue P, Mirakian R, Dixon TA, Harper JN, Nasser SM. BSACI guidelines for the investigation of suspected anaphylaxis during general anaesthesia. *Clin Exp Allergy* 2010; **40**: 15–31
9. Kroigaard M, Garvey LH, Gillberg L, et al. Scandinavian Clinical Practice Guidelines on the diagnosis, management and follow-up of anaphylaxis during anaesthesia. *Acta Anaesthesiol Scand* 2007; **51**: 655–70
10. Mertes PM, Malinovsky JM, Jouffroy L, et al. Reducing the risk of anaphylaxis during anesthesia: 2011 updated guidelines for clinical practice. *J Investig Allergol Clin Immunol* 2011; **21**: 442–53
11. Sudheer PS, Hall JE, Read GF, Rowbottom AW, Williams PE. Flow cytometric investigation of peri-anaesthetic anaphylaxis using CD63 and CD203c. *Anaesthesia* 2005; **60**: 251–6
12. Uyttebroek AP, Sabato V, Leysen J, Bridts CH, De Clerck LS, Ebo DG. Flowcytometric diagnosis of atracurium-induced anaphylaxis. *Allergy* 2014; **69**: 1324–32
13. Ebo DG, Faber M, Elst J, et al. In vitro diagnosis of immediate drug hypersensitivity during anesthesia: a review of the literature. *J Allergy Clin Immunol Pract* 2018; **6**: 1176–84
14. Takazawa T, Sabato V, Ebo DG. In vitro diagnostic tests for perioperative hypersensitivity: potential, limitations and perspectives. *Br J Anaesth* 2019; **123**: e117–25
15. Opstrup MS, Garvey LH. Optimizing diagnostics in perioperative allergy. *Br J Anaesth* 2019; **123**: e11–3
16. Brockow K. Dilemmas of allergy diagnosis in perioperative anaphylaxis. *Allergy* 2014; **69**: 1265–6
17. Krishna MT, York M, Chin T, et al. Multi-centre retrospective analysis of anaphylaxis during general anaesthesia in the United Kingdom: aetiology and diagnostic performance of acute serum tryptase. *Clin Exp Immunol* 2014; **178**: 399–404
18. Leysen J, Uyttebroek A, Sabato V, Bridts CH, De Clerck LS, Ebo DG. Predictive value of allergy tests for neuromuscular blocking agents: tackling an unmet need. *Clin Exp Allergy* 2014; **44**: 1069–75
19. Ramirez LF, Pereira A, Chiriac AM, Bonnet-Boyer MC, Demoly P. Negative predictive value of skin tests to neuromuscular blocking agents. *Allergy* 2012; **67**: 439–41
20. Aalberse RC, Kleine Budde I, Mulder M, et al. Differentiating the cellular and humoral components of neuromuscular blocking agent-induced anaphylactic reactions in patients undergoing anaesthesia. *Br J Anaesth* 2011; **106**: 665–74
21. Fisher MM, Bowey CJ. Intradermal compared with prick testing in the diagnosis of anaesthetic allergy. *Br J Anaesth* 1997; **79**: 59–63
22. Karila C, Brunet-Langot D, Labbez F, et al. Anaphylaxis during anesthesia: results of a 12-year survey at a French pediatric center. *Allergy* 2005; **60**: 828–34
23. Laroche D, Gomis P, Gallimidi E, Malinovsky JM, Mertes PM. Diagnostic value of histamine and tryptase concentrations in severe anaphylaxis with shock or cardiac arrest during anesthesia. *Anesthesiology* 2014; **121**: 272–9
24. Mertes PM, Laxenaire MC, Alla F. Anaphylactic and anaphylactoid reactions occurring during anesthesia in France in 1999–2000. *Anesthesiology* 2003; **99**: 536–45
25. Tacquard C, Collange O, Gomis P, et al. Anaesthetic hypersensitivity reactions in France between 2011 and 2012: the 10th GERAP epidemiologic survey. *Acta Anaesthesiol Scand* 2017; **61**: 290–9
26. Ring J, Messmer K. Incidence and severity of anaphylactoid reactions to colloid volume substitutes. *Lancet* 1977; **1**: 466–9
27. Niggemann B, Beyer K. Time for a new grading system for allergic reactions? *Allergy* 2016; **71**: 135–6

28. Rose MA, Green SL, Crilly HM, Kolawole H. Perioperative anaphylaxis grading system: 'making the grade. *Br J Anaesth* 2016; **117**: 551–3
29. Cook TM, Harper NJN, Farmer L, et al. Anaesthesia, surgery, and life-threatening allergic reactions: protocol and methods of the 6th National Audit Project (NAP6) of the Royal College of Anaesthetists. *Br J Anaesth* 2018; **121**: 124–33
30. Garvey LH. Perioperative hypersensitivity reactions: diagnosis, treatment and evaluation. *Curr Treat Options Allergy* 2016; **3**: 113–28
31. Brouwers MC, Kerkvliet K, Spithoff K. The AGREE Reporting Checklist: a tool to improve reporting of clinical practice guidelines. *BMJ* 2016; **352**: i1152
32. Fitch K, Bernstein SJ, Aguilar MD, et al. *The RAND/UCLA appropriateness method user's manual*. Santa Monica, CA: RAND; 2001
33. Maverakis E, Wang EA, Shinkai K, et al. Stevens–Johnson syndrome and toxic epidermal necrolysis standard reporting and evaluation guidelines: results of a national institutes of Health working group. *JAMA Dermatol* 2017; **153**: 587–92
34. Diamond IR, Grant RC, Feldman BM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *J Clin Epidemiol* 2014; **67**: 401–9
35. Edwards IR, Aronson JK. Adverse drug reactions: definitions, diagnosis, and management. *Lancet* 2000; **356**: 1255–9
36. Larach MG, Localio AR, Allen GC, et al. A clinical grading scale to predict malignant hyperthermia susceptibility. *Anesthesiology* 1994; **80**: 771–9
37. Ording H, Brancadoro V, Cozzolino S, et al. In vitro contracture test for diagnosis of malignant hyperthermia following the protocol of the European MH Group: results of testing patients surviving fulminant MH and unrelated low-risk subjects. The European Malignant Hyperthermia Group. *Acta Anaesthesiol Scand* 1997; **41**: 955–66
38. Allen GC, Larach MG, Kunselman AR. The sensitivity and specificity of the caffeine–halothane contracture test: a report from the North American malignant hyperthermia registry. The North American malignant hyperthermia registry of MHAUS. *Anesthesiology* 1998; **88**: 579–88
39. Larach MG, Brandom BW, Allen GC, Gronert GA, Lehman EB. Cardiac arrests and deaths associated with malignant hyperthermia in North America from 1987 to 2006: a report from the north american malignant hyperthermia registry of the malignant hyperthermia association of the United States. *Anesthesiology* 2008; **108**: 603–11
40. Brandom BW, Larach MG, Chen MS, Young MC. Complications associated with the administration of dantrolene 1987 to 2006: a report from the North American malignant hyperthermia registry of the malignant hyperthermia association of the United States. *Anesth Analg* 2011; **112**: 1115–23
41. Sampson HA, Munoz-Furlong A, Campbell RL, et al. Second symposium on the definition and management of anaphylaxis: summary report—second national institute of allergy and infectious disease/food allergy and anaphylaxis network symposium. *J Allergy Clin Immunol* 2006; **117**: 391–7
42. Harboe T, Guttormsen AB, Irgens A, Dybendal T, Florvaag E. Anaphylaxis during anesthesia in Norway: a 6-year single-center follow-up study. *Anesthesiology* 2005; **102**: 897–903
43. Gurrieri C, Weingarten TN, Martin DP, et al. Allergic reactions during anesthesia at a large United States referral center. *Anesth Analg* 2011; **113**: 1202–12
44. Gouel-Chéron A, de Chaisemartin L, Jönsson F, et al. Low end-tidal CO_2 as a real-time severity marker of intra-anaesthetic acute hypersensitivity reactions. *Br J Anaesth* 2017; **119**: 908–17
45. Harper NJN, Cook TM, Garcez T, et al. Anaesthesia, surgery, and life-threatening allergic reactions: management and outcomes in the 6th National Audit Project (NAP6). *Br J Anaesth* 2018; **121**: 172–88
46. Harper NJN, Cook TM, Garcez T, et al. Anaesthesia, surgery, and life-threatening allergic reactions: epidemiology and clinical features of perioperative anaphylaxis in the 6th National Audit Project (NAP6). *Br J Anaesth* 2018; **121**: 159–71
47. Valent P, Akin C, Arock M, et al. Definitions, criteria and global classification of mast cell disorders with special reference to mast cell activation syndromes: a consensus proposal. *Int Arch Allergy Immunol* 2012; **157**: 215–25
48. Brown SG, Blackman KE, Heddle RJ. Can serum mast cell tryptase help diagnose anaphylaxis? *Emerg Med Australas* 2004; **16**: 120–4
49. Garvey LH, Bech B, Mosbech H, et al. Effect of general anesthesia and orthopedic surgery on serum tryptase. *Anesthesiology* 2010; **112**: 1184–9
50. Egner W, Sargur R, Shrimpton A, York M, Green K. A 17-year experience in perioperative anaphylaxis 1998–2015: harmonizing optimal detection of mast cell mediator release. *Clin Exp Allergy* 2016; **46**: 1465–73
51. Baretto RL, Beck S, Heslegrave J, et al. Validation of international consensus equation for acute serum total tryptase in mast cell activation: a perioperative perspective. *Allergy* 2017; **72**: 2031–4
52. Simons FE, Ardusso LR, Bilo MB, et al. World allergy organization guidelines for the assessment and management of anaphylaxis. *World Allergy Organ J* 2011; **4**: 13–37
53. Berg CM, Heier T, Wilhelmsen V, Florvaag E. Rocuronium and cisatracurium-positive skin tests in non-allergic volunteers: determination of drug concentration thresholds using a dilution titration technique. *Acta Anaesthesiol Scand* 2003; **47**: 576–82
54. Spoerl D, Nigolian H, Czarnetzki C, Harr T. Reclassifying anaphylaxis to neuromuscular blocking agents based on the presumed patho-mechanism: IgE-mediated, pharmacological adverse reaction or "innate hypersensitivity"? *Int J Mol Sci* 2017; **18**: 1223
55. Sprung J, Weingarten TN, Schwartz LB. Presence or absence of elevated acute total serum tryptase by itself is not a definitive marker for an allergic reaction. *Anesthesiology* 2015; **122**: 713–4
56. Sprung J, Larson KJ, Divekar RD, Butterfield JH, Schwartz LB, Weingarten TN. Refractory intraoperative hypotension with elevated serum tryptase. *Asia Pac Allergy* 2015; **5**: 47–50

Handling editor: H.C. Hemmings Jr