OPEN ACCESS

Development of a neonatal adverse event severity scale through a Delphi consensus approach

Thomas Salaets,[1] Mark A Turner,[2] Mary Short,[3] Robert M Ward,[4] Isamu Hokuto,[5] Ronald L Ariagno,[6] Agnes Klein,[7] Sandra Beauman,[8] Kelly Wade,[9] Merran Thomson,[10] Eve Roberts,[11] Judy Harrison,[12] Theresa Quinn,[13] Gerri Baer,[14] Jonathan Davis,[15,16] Karel Allegaert,[1,17] On behalf of the International Neonatal Consortium

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/archdischild-2019-317399).

For numbered affiliations see end of article.

Correspondence to
Dr Karel Allegaert, Department of Pediatrics, Division of Neonatology, Erasmus MC Sophia Kinderziekenhuis, Rotterdam 3015, Netherlands; karel.allegaert@uz.kuleuven.ac.be

Received 11 April 2019
Revised 24 August 2019
Accepted 3 September 2019

ABSTRACT

Background Assessment of the seriousness, expectedness and causality are necessary for any adverse event (AE) in a clinical trial. In addition, assessing AE severity helps determine the importance of the AE in the clinical setting. Standardisation of AE severity criteria could make safety information more reliable and comparable across trials. Although standardised AE severity scales have been developed in other research fields, they are not suitable for use in neonates. The development of an AE severity scale to facilitate the conduct and interpretation of neonatal clinical trials is therefore urgently needed.

Methods A stepwise consensus process was undertaken within the International Neonatal Consortium (INC) with input from all relevant stakeholders. The consensus process included several rounds of surveys (based on a Delphi approach), face-to-face meetings and a pilot validation.

Results Neonatal AE severity was classified by five grades (mild, moderate, severe, life threatening or death). AE severity in neonates was defined by the effect of the AE on age appropriate behaviour, basal physiological functions and care changes in response to the AE. Pilot validation of the generic criteria revealed κ=0.23 and guided further refinement. This generic scale was applied to 35 typical and common neonatal AEs resulting in the INC neonatal AE severity scale (NAESS) V.1.0, which is now publicly available.

Discussion The INC NAESS is an ongoing effort that will be continuously updated. Future perspectives include further validation and the development of a training module for users.

What is already known on this topic?

► Communication of safety data between study investigators, sponsors and regulators remains suboptimal because of diverse ways of collection, reporting and assessment of adverse event information. In several research fields, severity scales have been developed to standardise adverse event severity reporting; however, the existing scales are not applicable to neonates.

What this study adds?

► This study describes a consensus process that led to the development of standard severity criteria for neonatal adverse events. The use of this tool could improve the quality of drug and device safety evaluations and facilitate the conduct of neonatal clinical trials.

INTRODUCTION

An adverse event (AE) is defined as 'any untoward medical occurrence associated with the use of a drug in humans, whether or not considered drug related'.[1–3] In vulnerable populations such as critically ill neonates, background rates of mortality and morbidity are high.[4] This leads to a higher incidence of reported AEs, not necessarily attributed to the investigational medicinal product (IMP) or device. However, when recognised and reported in a standardised manner, AEs can be important safety signals.

Regulatory guidelines require investigators to assess whether an AE is serious and whether there is a reasonable possibility that it is related to IMP administration (causality) (figure 1).[1 5] A strict regulatory definition exists for 'seriousness'.[5] Although causality involves clinical judgement, algorithms have been developed to make this assessment in a neonatal population more objective and homogeneous.[6 7] Data Safety Monitoring Boards and sponsors review the investigator report of an AE and adjudicate whether the AE is expected based on known side effects (Reference Safety Information) and background complication rates.[5] Finally, assessing the severity of an AE enhances the reporting process by capturing medical intensity.

Communication of safety data between study investigators, sponsors and regulators remains suboptimal because of diverse ways of collection, reporting and assessment. A common clinical research language, using standard terms and definitions, could facilitate responsible data sharing.[8 9] Within neonatology, efforts have been directed towards developing standard terminology and definitions for AEs[10] that integrate in to larger dictionaries such as the Thesaurus of the National Cancer Institute (NCI) or the Medical Dictionary for Regulatory Activities (MedDRA).[11 12]

Standardising criteria to report AE severity could make safety information more comparable across centres and trials and is a reasonable next step.[13]

Check for updates

To cite: Salaets T, Turner MA, Short M, *et al*. *Arch Dis Child* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/archdischild-2019-317399

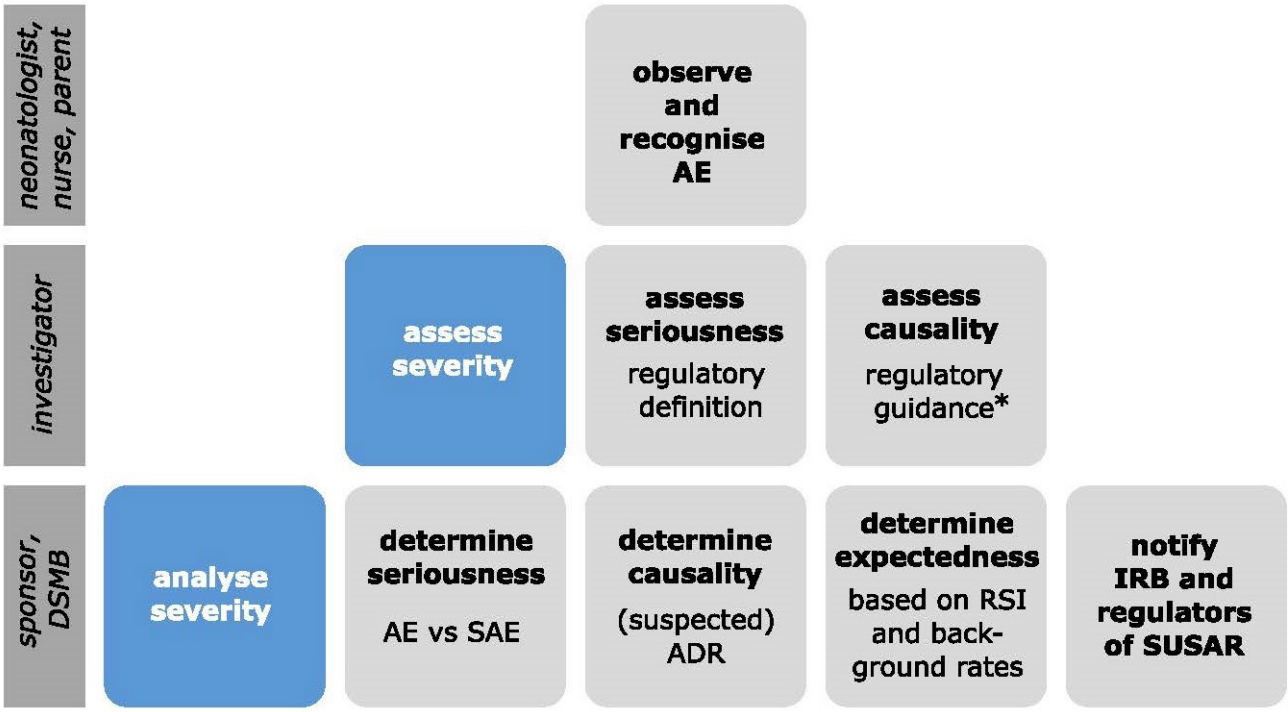

Figure 1 This figure summarises aspects of AEs that should be considered to account for safety reporting. It visualises responsibilities of the different actors and the currently available criteria and guidance. The figure is not intended to illustrate sequential activities. *Causality assessment relies on regulatory guidance; however, algorithms (eg, Du *et al*[6]) have been developed for a neonatal population. AE, adverse event; DMSB, Data Safety Monitoring Boards; SAE, serious adverse event; ADR, adverse drug reaction; RSI, reference safety information; IRB, institutional review board; SUSAR, suspected unexpected serious adverse drug reaction.

In other research fields, toxicity tables and AE severity scales are commonly used.[14–16] The most widely used example is the Common Terminology Criteria for Adverse Events (CTCAE), which was developed for oncology research.[17] As the generic severity criteria of these scales are not applicable to neonates (table 1), we identified the need for a neonatal AE severity scale (NAESS) with criteria for specific neonatal AEs.

METHODS

A stepwise consensus process, based on a Delphi approach,[18] was undertaken between December 2016 and September 2018. First, generic severity criteria for neonatal AEs were developed. Next, these severity criteria were applied to frequently occurring neonatal AEs. Finally, the terminology was integrated into larger terminology dictionaries (figure 2). Full reports for every step can be found in the appendices.

Stakeholder input

The NAESS was developed within the International Neonatal Consortium (INC). INC was established in 2015 in order to address regulatory and scientific challenges in the development of innovative drugs for neonates.[19] Throughout the neonatal AE scale development, input was requested from multiple key stakeholders involved in neonatal drug development. Academic and non-academic clinicians and researchers, industry representatives, regulators, nursing and funding organisations and parent representatives from Canada, Europe, Japan and USA participated in the process (figure 2). Respondent and participant groups were expanded at every round to incorporate feedback from a maximal number of stakeholders.

Development of generic severity criteria

Step 1: a first anonymous online survey was circulated through the network of INC and assessed in general terms regarding which aspects of AEs could be used as severity markers and provide input on the framework of NAESS.

Step 2: a second anonymous online survey presented the results and feedback of the first survey together with a proposal for generic severity criteria based on these results (Delphi approach).[18]

Step 3: in a face-to-face meeting, all components of the scale were discussed in the context of the feedback received in the

Table 1 Generic severity criteria of CTCAE, which are commonly used for adult and paediatric patients, but are not directly applicable to neonates

Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
Mild	Moderate	Severe	Life threatening	Death
Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated.	Moderate; minimal, local or non-invasive intervention indicated; limiting age-appropriate instrumental activities of daily living.	Severe or medically significant but not immediately life threatening; hospitalisation or prolongation of hospitalisation indicated; disabling; limiting self-care activities of daily living.	Life-threatening consequences; urgent intervention indicated.	Death related to AE.

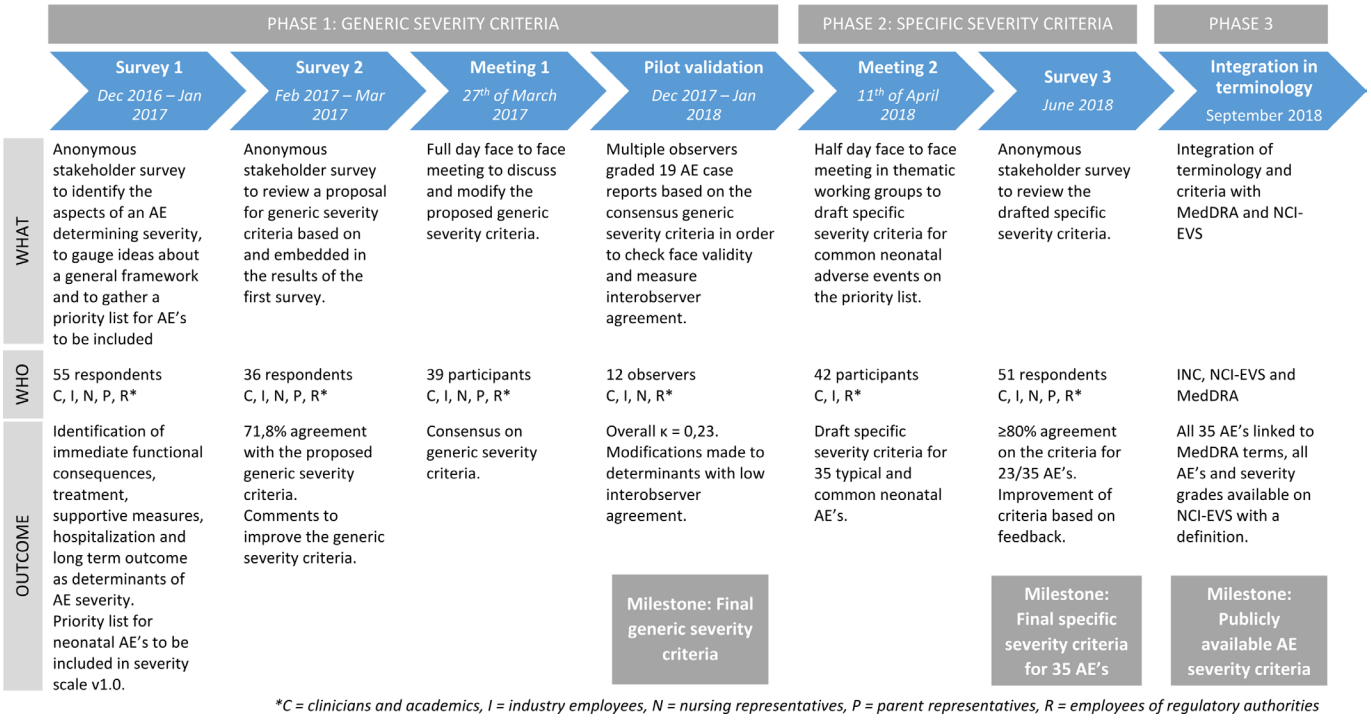

Figure 2 This figure gives an overview of the development process of the NAESS. Stakeholder involvement is indicated by C (clinicians), I (industry), N (nursing representatives), P (parent representatives) and R (regulatory authority employees). AE, adverse event; INC, International Neonatal Consortium; NCI-EVS, National Cancer Institute—Enterprise Vocabulary Services; MedDRA, Medical Dictionary for Regulatory Activities; NAESS, neonatal AE severity scale.

surveys in order to achieve consensus on the generic severity criteria.

Step 4: pilot validation was undertaken to test the validity and interobserver agreement of the consensus generic severity criteria. Nineteen written case reports of AEs were provided by the University of Liverpool.[20] All reports contained information on parameters, clinical and technical examinations, drug exposure and any resulting care changes. Twelve observers with different backgrounds from Canada, Europe, Japan and USA graded the severity of the 19 AEs using the proposed full generic severity scale and all individual markers (see online supplementary appendix 1). The results were analysed by calculating a free-marginal multirater kappa as a measure of interobserver reliability.[21 22] The generic criteria were subsequently improved based on the results of this exercise.

Development of event specific severity criteria

Step 5: during a subsequent *face-to-face meeting*, thematic subgroups (neurological, cardiovascular, respiratory, gastrointestinal and infectious/general neonatology) drafted severity criteria for a list of specific neonatal AEs provided in step 1.

Step 6: the resulting specific severity criteria were evaluated in a final anonymous online survey. For all AEs with more than 20% disagreement, modifications were made in order to align all the key stakeholders. All modified criteria were approved in a final teleconference.

Linking to existing terminology

Step 7: for each AE, a definition was used from the National Institute of Child Health and Human Development (NICHD) Pediatric AE Terminology if appropriate.[10] Every AE was linked to the corresponding MedDRA Lowest Level Terms (LLTs). MedDRA Maintenance and Support Services Organization was contacted with a proposal to adjust or add terms if no suitable LLT was available. All criteria were added to the NCI Thesaurus.

RESULTS

In total, 109 members participated in the process leading to consensus on INC NAESS V.1.0. Participant numbers and background for all steps are summarised in figure 2 and online supplementary appendix 2. All participants who were involved in at least one step are listed in the Acknowledgement section.

Development of generic severity criteria

Step 1: e received 55 responses to the first survey. Immediate functional consequences (accepted by 81% of respondents), changes in treatment (82%), prolongation of hospitalisation (75%), supportive measures (85%) and long-term outcome (73%) were accepted as indicators of AE severity. Many comments referred to the feasibility of using long-term outcome to classify AE severity (see online supplementary appendix 3).

Step 2: 36 respondents completed the second survey. Of the respondents, 72% agreed with the proposal for generic severity criteria based on immediate functional consequences, treatment (including supportive measures) and prolongation of the initial hospitalisation. The remaining 28% of respondents suggested adjustments (see online supplementary appendix 4).

Step 3: 39 experts participated in the first face-to-face meeting trying to define generic severity criteria for neonates. Consistent with other AE severity scales, severity was subdivided into five categories: mild, moderate, severe, life threatening and death. Immediate functional consequences (on age appropriate behaviour and basal physiological functions), together with resulting care changes were established as the parameters of the generic AE severity scale. Additionally, the stakeholder group

Table 2 Generic severity criteria of INC NAESS developed for use in neonates

Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
Mild	Moderate	Severe	Life threatening	Death
Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; *no change in baseline age-appropriate behaviour*; no change in baseline care or monitoring indicated.*	Moderate; *resulting in minor changes of baseline age-appropriate behaviour*; requiring minor changes in baseline care or monitoring.*†*	Severe; *resulting in major changes of baseline age-appropriate behaviour* or non-life-threatening changes in basal physiological processes†; requiring major change in baseline care or monitoring.*‡*	Life threatening; *resulting in life-threatening changes in basal physiological processes†; requiring urgent major change in baseline care.*	Death related to AE.

If the different factors of this scale result in conflicting severity grades, the highest grade should be reported. *Italics* indicate the differences with the adult generic severity criteria of CTCAE.
*Age-appropriate behaviour refers to oral feeding behaviour, voluntary movements and activity, crying pattern, social interactions and perception of pain.
†Basal physiological processes refer to oxygenation, ventilation, tissue perfusion, metabolic stability and organ functioning.
‡Minor care changes constitute: brief, local, non-invasive or symptomatic treatments.
§Major care changes constitute: surgery, addition of long-term treatment, upscaling care level.
CTCAE, Common Terminology Criteria for Adverse Events.

agreed that this AE severity scale would pertain to neonates <44 weeks postmenstrual age.[23 24]

Step 4: the proposed consensus scale resulted in a multirater free-marginal κ of 0.23 (ranging from −0.03 to 0.59 for individual AEs). Severity evaluations based on 'care changes' alone resulted in suboptimal agreement. Furthermore, marked intraobserver and interobserver variability was noted in how the 12 observers weighted the different factors involved in the determination for the final severity grade (see online supplementary appendix 1).

Examples of minor and major care changes were added for clarification. It was also specified that if individual determinants of severity resulted in a discrepant severity grade, the highest grade would be reported. These changes resulted in the final generic severity criteria as shown in table 2.

Development of event specific severity criteria

Step 5: 42 professionals, in five thematic subgroups, drafted the specific severity criteria for the 35 common neonatal AEs (table 3) during a second face-to-face meeting. Specific disease markers were used as severity criteria if it was determined that they reflected the factors incorporated in the generic severity scale. AEs based on laboratory values were not included in this first version because reliable reference values for normality were not available (for postmenstrual and postnatal age).

Step 6: among the 51 respondents of the final survey, there was ≥80% agreement for 23 out of 35 of the draft criteria (indicated with an asterisk*). For the remaining 12 AEs, criteria were then adjusted (see online supplementary appendix 5) and approved in a final teleconference. The resulting specific severity criteria for all selected 35 AEs can be found in the INC AE severity scale V.1.0 (online supplementary appendix 6); as an example, the specific severity criteria for neonatal convulsion are shown in table 4.

Linking to existing terminology

For 23 (66%) of the AEs in V.1.0 definitions were taken directly from the NICHD Paediatric AE Terminology. Nineteen (54%) of the AEs could be linked to an existing LLT in MedDRA V.21.0. New LLTs were added in MedDRA V.22.0 to match the remaining AEs. For all AEs and severity grades, specific codes for the NCI Thesaurus were generated. The terminology was made publicly available on the NCI Thesaurus (https://evs.nci.nih.gov/ftp1/INC/Adverse_Events_Terminology/).

DISCUSSION

This paper describes the development of the INC NAESS. It provides criteria that guide investigators and clinicians in assessing severity of AEs. It was developed to increase the quality of safety information.

Generic severity criteria

The common framework of the AE severity scale is contained in generic criteria shown in table 2. It defines suitable severity markers for neonates, in contrast to criteria used in other populations (eg, CTCAE; table 1). First, impact on age-appropriate activities is included, describing the signs exhibited by the neonate (feeding behaviour, voluntary movements, activity, crying, social interactions and signs of pain). The impact on basal physiological processes (changes in oxygenation, ventilation, circulation, metabolic stability and organ function) is a second severity marker. A third severity marker is the change in care in response to the AE, as it reflects the severity of the underlying

Table 3 AEs included in the current version of INC NAESS

AEs in INC neonatal AE severity scale	
Neurological	**Respiratory**
Neonatal convulsion	Infantile apnoea
Neonatal epileptic seizure	Neonatal respiratory insufficiency
Neonatal intraventricular haemorrhage*	Neonatal respiratory distress syndrome
Retinopathy of prematurity*	Neonatal pulmonary haemorrhage*
Hypoxic ischaemic encephalopathy	Persistent pulmonary hypertension of the newborn*
Periventricular leukomalacia*	Neonatal pneumothorax*
Infant irritability	Bronchopulmonary dysplasia
Infant sedation*	**Gastrointestinal**
Cardiovascular	Necrotising enterocolitis
Neonatal hypotension	Neonatal diarrhoea*
Neonatal hypertension*	Infantile vomiting*
Neonatal sinus tachycardia*	Feeding intolerance
Neonatal sinus bradycardia	Neonatal gastrointestinal bleeding*
Neonatal tachyarrhythmia*	Neonatal spontaneous intestinal perforation*
Neonatal bradycardia	Neonatal constipation*
Neonatal oedema*	**General**
Neonatal coagulation disorder*	Neonatal rash*
Infectious	Neonatal administration site complication*
Neonatal culture positive sepsis*	Neonatal fever*
Neonatal culture negative sepsis*	

For these 35 AEs, specific severity criteria were defined.
AE, adverse event; INC, International Neonatal Consortium; NAESS, neonatal AE severity scale.

Table 4 Severity criteria for neonatal convulsions, as an example of the specific severity criteria per AE given in INC NAESS (online supplementary appendix 6)

Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
Mild	Moderate	Severe	Life threatening	Death
Neonatal convulsion Definition C154952 | 10028932: *Sudden, involuntary, rapid, rhythmic or stereotyped skeletal muscular contractions in a newborn.*				
Single, self-limited suspected seizure, no treatment.	Suspected seizures controlled with one anti-seizure drug (no recurrence within 3 days after treatment).	Suspected seizures uncontrolled with one antiseizure drug (recurrence within 3 days after treatment or requiring two or more antiseizure drugs).	Suspected seizures with life-threatening consequences (eg, need for ventilation); suspected status epilepticus* despite multiple anti-seizure drugs.	Death related to suspected seizures.

*>30 min duration of convulsions within any 60 min period.
AE, adverse event; INC, International Neonatal Consortium; NAESS, neonatal AE severity scale.

event and is indicative of additional stress. All of this information is readily available after the AE occurs and permits immediate severity grading.

It is important to note that for all determinants, only changes from the baseline condition due to the AE should be considered. Also, even though these criteria provide guidance, some form of clinical judgement (and thus subjectivity) remains inherent in the severity assessment.

The final decision was not to include long-term outcome as a marker of AE severity, as it might be difficult to establish a direct causal link. Furthermore, the overall goal of this scale is to create reliable and immediate safety signals prompting awareness, which is not compatible with assessing the severity of an event when the final outcome is only known years later. Despite this decision, examining the long-term outcome associated with a neonatal drug exposure remains a crucial effort that should be encouraged.[24]

Validation

The goal of a standardised severity scale is to reduce subjectivity in severity assessments and thus reduce interobserver variability. Our pilot validation exercise on the consensus generic severity criteria revealed only fair agreement (κ=0.23) among observers of different backgrounds. This seems less rigorous than what is published for other severity scales (eg, CTCAE[25] and SAVES-2 (Spinal Adverse Event Severity System, version 2)).[26] It should be noted that our results were obtained by applying a generic scale, which is purposely broad and thus provides less direct guidance. Furthermore, our pilot validation was performed as a part of an optimisation process, resulting in improved final generic severity criteria. Future plans include measurement of interobserver agreement with the final scale on prospectively collected data in order to show the benefit of using standardised criteria for the severity of neonatal AEs.

INC NAESS V.1.0

The current version (V.1.0) contains specific severity criteria for 35 routinely encountered neonatal AEs and more will be added. For instance, laboratory-based AEs will be included when age-appropriate reference values become available. The severity of AEs currently not included in INC NAESS V.1.0 can be estimated by applying the generic criteria.

This INC NAESS V.1.0 is now publicly available in the NCI Thesaurus. This platform allows end users to embed the criteria relevant to a particular project within the concept information. Its presence in the Thesaurus improves dissemination to a relevant audience. Finally, NCI Thesaurus users can request the addition of new AEs or modifications of existing criteria that permits the terminology to be revised as needed. This ensures the sustainability of the INC NAESS.

The INC NAESS is a continuous work in progress. We are preparing to expand the number of AEs included and encourage readers to suggest new AEs for addition or comment on existing criteria if deemed not appropriate through the NCI Thesaurus website.

Anticipated benefits

Severity assessments provide a nuanced clinical appreciation of an AE. Standardising this information leads to more reliable and comparable information that can facilitate regulatory safety evaluations of drugs. Furthermore, it can improve scientific communication on AEs in publications. The availability of a standardised severity scale can also facilitate the conduct of clinical trials in neonates. For example, dose reductions can be recommended in response to severe AEs. In the neonatal population where the background rates of clinical AEs are high, protocols could stratify reporting obligations for different severity grades and background rates of AEs. This could enhance consistent AE reporting and reduce the administrative burden associated with neonatal research.

Finally, the scale could also be used in routine clinical care and for postmarketing pharmacovigilance. Neonates are vulnerable to adverse drug reactions (ADR) in clinical care, but these events are under-reported.[27] NAESS was developed to classify AEs, which would complement neonatal algorithms used to identify ADRs.[6 7] A standardised evaluation of severity would add information that is useful for pharmacovigilance, quality improvement projects, observational studies or registries and can help establish risk/benefit ratios of common therapies.[28]

CONCLUSION

In conclusion, a NAESS containing both generic and specific criteria for 35 common neonatal AEs was developed in collaboration with key stakeholders involved in neonatal drug development. The use of this tool can improve the quality of drug and device safety evaluations and can facilitate the conduct of neonatal clinical trials. Future perspectives include validation research to assess interobserver agreement and the addition of more AEs.

Author affiliations
[1]Department of Development and Regeneration, KU Leuven, Leuven, Belgium
[2]Institute of Translational Medicine, University of Liverpool, Liverpool, UK
[3]Eli Lilly and Co, Indianapolis, Indiana, USA
[4]4Department of Pediatrics, Divisions of Neonatology and Clinical Pharmacology, University of Utah, Salt Lake City, Utah, USA
[5]Department of Pediatrics, St. Marianna University, Kawasaki, Japan

[6]Department Pediatrics-Neonatology, Stanford University School of Medicine, Palo Alto, California, USA
[7]Health Canada, Ottawa, Ontario, Canada
[8]Department of Pediatrics, University of New Mexico Health Sciences Center, Albuquerque, New Mexico, USA
[9]Division of Neonatology, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA
[10]Hillingdon Hospitals NHS Foundation Trust, Uxbridge, UK
[11]Institute of Translational Medicine, University of Liverpool, Liverpool, UK
[12]Maintenance and Support Services Organization, MedDRA, McLean, Virginia, USA
[13]Enterprise Vocabulary Services, National Cancer Institute, Bethesda, Maryland, USA
[14]Office of Pediatric Therapeutics, US Food and Drug Administration, Silver Spring, Maryland, USA
[15]Floating Hospital for Children at Tufts Medical Center, Boston, Massachusetts, USA
[16]Tufts Clinical and Translational Science Institute, Boston, Massachusetts, USA
[17]Department of Pediatrics, Division of Neonatology, Erasmus MC Sophia Kinderziekenhuis, Rotterdam, The Netherlands

Acknowledgements We would like to acknowledge Anna Davis and Rebecca Spencer from the University College London for their help in initiating this process.

Collaborators The International Neonatal Consortium would like to acknowledge the following people who participated in the process: Marilee Allen, Johns Hopkins; Albert Allen, Eli Lilly; Dina Apele Freimane, PDCO, P. Stradins, Clinical University Hospital; Judy Aschner, The Joseph M Sanzari Children's Hospital at Hackensack Meridian Health School of Medicine and Albert Einstein College of Medicine; Roberta Ballard, UCSF School of Medicine; Yodit Belew, FDA; Ralph Bax, EMA; Bobby Bellflower, The University of Tennessee; Varsha Bhatt-Mehta, University of Michigan; Mike Blum, FDA; Chiara Bonardi, Chiesi Pharmaceuticals; Patti Bondurant, Greenwood Solutions; Geraldine Boylan, INFANT Research Centre & Department of Paediatrics and Child Health; Teresa Buracchio, FDA; Gilbert Burckart, FDA; Amber Burnett, Baxter International; Janelle Burnham, FDA; Cheryl Carlson, UNC; Alice Chen, NIH; Bob Clay, Highbury Regulatory Science Ltd; Alan Cohen, Bayer; Eamonn Connolly, Infant Bacterial Therapeutics; Edward Connor, Institute for Advanced Clinical Trials for Children; Edress Darsey, Pfizer Inc; Roberto De Lisa, EMA; Jennifer Degl, Speaking for Moms and Babies, Inc.; Eugene Dempsey, INFANT Research Centre & Department of Paediatrics and Child Health; Tom Diacovo, Children's Hospital of Pittsburgh; Janis Dionne, BC Children's Hospital; Jennifer Duchon, Tufts Medical Center; Wakako Eklund, National Association of Neonatal Nurses; Laura Fabbri, Chiesi Pharmaceuticals; Alistair Fielder, City University; Emily Freilich, FDA; Sabine Fürst-Recktenwald, Hoffmann-La Roche LTD; Tamorah Graham, Quorum Health Corporation; Dionna Green, FDA; Cristal Grogan, Preemie Parent Alliance; Darell Heiselman, Eli Lilly; Ann Hellström, University of Gothenburg; Ralf Herold, Bayer; Anna Maria Hibbs, Case Western Reserve University School of Medicine; Steven Hirschfeld, NIH; Mary Hise Brown, Baxter International; Jim Holberton, Mercy Health Foundation; Pablo Iveli, Hospital Quirónsalud Barcelona; Alan Jobe, Cincinnati Children's Hospital; Allyson Kayton, National Association of Neonatal Nurses; Carole Kenner, Council of International Neonatal Nurses; Walter Kraft, Thomas Jefferson University; Satoshi Kusuda, Tokyo Women's Medical University; Thierry Lacaze, The Pediatric Network in Canada; Tamorah Lewis, Children's Mercy Kansas City; Kei Lui, Australian & New Zealand Neonatal Network; Irja Lutsar, University of Tartu, Estonia; Alexandra Mangili, Takeda; Barry Mangum, Paidion Research, Inc.; Susan McCune, FDA; Karen McDonald, Barton and Associates, Inc.; Courtney McGuire, FDA; Christopher McPherson, St. Louis Children's Hospital; Washington University; Thomas Miller, Bayer; Jeffrey Ming, Sanofi Pharmaceuticals; Alison Mitzner, Pfizer; Neena Modi, Imperial College London; Maya Morrison, Baxter International; Eva Morsing, Lund University; Lily Mulugeta, FDA; Hide Nakamura, Nagano Children's Hospital; Skip Nelson, J and J Services, Inc.; Gary Noel, J and J Services, Inc.; Miki Olshansky, ELGAN; Alex Oschman, Children's Mercy Hospital and Clinics; Hulya Ozsahin, Centre Médical Qorpus; Mike Padula, Children's Hospital of Philadelphia; Vasum Peiris, FDA; Luana Pesco Koplowitz, DUCK FLATS Pharma; Josh Pham, Baxter International; Aprile Pilon, Trove Therapeutics; Ron Portman, Novartis pharmaceuticals; Ronit Pressler, Great Ormond Street Hospital; Heike Rabe, Brighton & Sussex Medical School; Tonse Raju, NIH; Douglas Andrew Roepke, Eli Lilly; Rashmin Savani, University of Texas Southwestern Medical Center; Patrick Schnell, Pfizer; Robert Segal, Windtree Therapeutics; Thibault Senterre, Baxter International; Prakesh Shah, Canadian Neonatal Network/University of Toronto; Philip Sheridan, FDA; Catherine Sherwin, University of Utah; Rajendra Singh, GlaxoSmithkline; Roger Soll, Vermont Oxford Network; Janet Soul, Harvard University; Kaye Spence, Children's Hospital at Westmead; Linda Storari, Chiesi Pharmaceuticals; Jan Taminiau, Emma Children's Hospital; Brian Tseng, Novartis Pharmaceuticals; John Van Den Anker, Sophia Children's Hospital (the Netherlands) and Children's National Medical Center, Washington, D.C.; Judy Varga, Windtree Therapeutics; Preeti Venkataraman, FDA; Norma Vivas, Sanofi; Karen Walker, Australian College of Neonatal Nurses; Janice Wilson, University of Maryland; Anne Zajicek, NICHD; Gary Zaloga, Baxter International; and Stav Zeldis, ELGAN.

Contributors TS and KA designed and guided the development process and drafted the manuscript. MAT, JD and RMW helped supervise and guide the entire process. SB, GB, IH, RLA, AK, MS, KW and MT participated in every step of the development process. JH linked the terminology to MedDRA; TQ created the concepts and links in the National Cancer Institute Thesaurus. ER designed and provided the case descriptions for the validation exercise. All authors were involved in reviewing and rewriting the manuscript.

Funding The International Neonatal Consortium (INC) is supported in part by grant number U18FD005320-01 from the US Food and Drug Administration (FDA) to the Critical Path Institute (http://c-path.org) and through annual dues of member companies. The first and last author are supported by the SafePedrug project (IWT SBO 130033).

Disclaimer The views expressed in this article are the personal views of the authors and may not be understood or quoted as being made on behalf of or reflecting the position of the US Food and Drug Administration, the position of Health Canada or the position of the National Institutes of Health.

Competing interests One coauthor is an employee of a pharmaceutical company, as noted in the list of affiliations. MT has performed paid consultancy work for Chiesi Farmaceutici S.p.A, Italy. No products are discussed in this article. The consortium aims to improve methods that can be applied to evaluating the safety and effectiveness of any medical product for neonates and is consequently drug agnostic.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

REFERENCES

1. Abraham J. International Conference On Harmonisation Of Technical Requirements For Registration Of Pharmaceuticals For Human Use. In: *Handbook of transnational economic governance regimes*. Brill, 2009: 1041–54.
2. Center for Drug Evaluation and Research. Investigational New Drug (IND) Application - Final Rule: Investigational New Drug Safety Reporting Requirements for Human Drug and Biological Products and Safety Reporting Requirements for Bioavailability and Bioequivalence Studies in Humans. Available: https://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/InvestigationalNewDrugINDApplication/ucm226358.htm [Accessed 1 Dec 2018].
3. Center for Devices and Radiological Health. Medical Device Databases - Code of Federal Regulations - Title 21 - Food and Drugs. Available: https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Databases/ucm135680.htm [Accessed 1 Dec 2018].
4. Stoll BJ, Hansen NI, Bell EF, *et al*. Trends in care practices, morbidity, and mortality of extremely preterm neonates, 1993-2012. *JAMA* 2015;314.
5. FDA. Guidance to industry and Investigaors: safety reporting requirements for IND and BA/BE. Available: https://www.fda.gov/downloads/Drugs/Guidances/UCM227351.pdf
6. Du W, Lehr VT, Lieh-Lai M, *et al*. An algorithm to detect adverse drug reactions in the neonatal intensive care unit. *J Clin Pharmacol* 2013;53:87–95.
7. Gallagher RM, Kirkham JJ, Mason JR, *et al*. Development and inter-rater reliability of the Liverpool adverse drug reaction causality assessment tool. *PLoS One* 2011;6:e28096.
8. Costeloe K, Turner MA, Padula MA, *et al*. Sharing data to accelerate medicine development and improve neonatal care: data standards and harmonized definitions. *J Pediatr* 2018;203:437–41.
9. Kush R, Goldman M. Fostering responsible data sharing through standards. *N Engl J Med* 2014;370:2163–5.
10. Gipson DS, Kirkendall ES, Gumbs-Petty B, *et al*. Development of a pediatric adverse events terminology. *Pediatrics* 2017;139:e20160985.
11. Brown EG, Wood L, Wood S. The medical dictionary for regulatory activities (MedDRA). *Drug Saf* 1999;20:109–17.
12. MedDRA. Available: https://www.meddra.org/ [Accessed 28 Feb 2019].
13. Woodcock J, LaVange LM. Master protocols to study multiple therapies, multiple diseases, or both. *N Engl J Med* 2017;377:62–70.
14. Division of AIDS table for grading the severity of adult and pediatric adverse events. Available: http://ucdmc.ucdavis.edu/clinicaltrials/StudyTools/Documents/DAIDS_AE_GradingTable_FinalDec2004.pdf
15. Division of microbiology and infectious diseases (DMID) pediatric toxicity tables. Available: http://www.fda.gov/downloads/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/Vaccines/ucm091977.pdf

16. FDA, CBER. Guidance for industry: toxicity grading scale for healthy adult and adolescent volunteers enrolled in preventive vaccine clinical trials; 2007.
17. CTCAE. Common terminology criteria for adverse events V5. Available: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_5x7.pdf
18. McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. *Int J Clin Pharm* 2016;38:655–62.
19. Turner MA, Davis JM, McCune S, *et al*. The International neonatal Consortium: collaborating to advance regulatory science for neonates. *Pediatr Res* 2016;80:462–4.
20 Roberts E, Hawcutt D, Turner M. Adrin 1 methodology study: adverse drug reactions in neonates: what are the best ways to evaluate suspected adverse drug reactions in neonates? *Arch Dis Child* 2017;102:A11.2–A11.
21 Randolph JJ. Author note: Free-Marginal Multirater kappa (multirater κfree): an alternative to Fleiss ' Fixed- marginal Multirater kappa. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.59.8776
22. Randolph J. Online kappa calculator. Available: http://justusrandolph.net/kappa/#dInfo
23. European Medicines Agency and Heads of Medicines Agencies. Guideline on good pharmacovigilance practices (GVP) - Product- or Population-Specific Considerations IV: Paediatric population; 2018.
24. Ward RM, Benjamin D, Barrett JS, *et al*. Safety, dosing, and pharmaceutical quality for studies that evaluate medicinal products (including biological products) in neonates. *Pediatr Res* 2017;81:692–711.
25. Atkinson TM, Li Y, Coffey CW, *et al*. Reliability of adverse symptom event reporting by clinicians. *Qual Life Res* 2012;21:1159–64.
26. Rampersaud YR, Anderson PA, Dimar JR, *et al*. Spinal adverse events severity system, version 2 (SAVES-V2): inter- and intraobserver reliability assessment. *J Neurosurg* 2016;25:256–63.
27. Hawcutt DB, O'Connor O, Turner MA. Adverse drug reactions in neonates: could we be documenting more? *Expert Rev Clin Pharmacol* 2014;7:807–20.
28. Abman SH, Hansmann G, Archer SL, *et al*. Pediatric pulmonary hypertension: guidelines from the American heart association and American thoracic Society. *Circulation* 2015;132:2037–99.