Vaccine 26 (2008) 6349–6360

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

Applicability, reliability, sensitivity, and specificity of six Brighton Collaboration standardized case definitions for adverse events following immunization☆

Katrin S. Kohl[a,*,1], Manya Magnus[b], Robert Ball[c], Neal Halsey[d,e], Sean Shadomy[a,2], Thomas A. Farley[f]

[a] Immunization Safety Office, Office of the Chief Science Officer, Centers for Disease Control and Prevention, 1600 Clifton Road, Mailstop E-03, Atlanta, GA, USA
[b] Department of Epidemiology and Biostatistics, George Washington University School of Public Health, Washington, DC, USA
[c] Office of Biostatistics and Epidemiology, Center for Biologics Evaluation and Research, Food and Drug Administration, Rockville, MD, USA
[d] Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
[e] Clinical Immunization Safety Assessment network, USA
[f] Department of Community Health Sciences, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, USA

ARTICLE INFO

Article history:
Received 26 August 2008
Accepted 1 September 2008
Available online 19 September 2008

Keywords:
Adverse events following immunization
Case definitions
Evaluation methodology
Fever
Persistent crying
Hypotonic–hyporesponsive episode
Intussusception
Nodule at injection site
Generalized convulsive seizure

ABSTRACT

We evaluated the applicability, reliability, sensitivity, and specificity of six standardized case definitions for adverse events following immunization (AEFI) (for fever, generalized convulsive seizure, hypotonic–hyporesponsive episode, intussusception, nodule, and persistent crying) developed by the Brighton Collaboration using the U.S. Vaccine Adverse Event Reporting System (VAERS). The evaluation included: (a) the development of codified search strings using standardized coding terminology, and (b) for sensitivity and specificity analyses, the development of a "gold standard" for case determination by clinical expert reviews, and its comparison against the application of the definitions to VAERS reports by nonclinicians. Application of the case definitions in an automated approach proved to be valid, feasible, and unlikely to miss confirmed cases of the reported clinical event. The definitions had variable but generally high sensitivity and specificity compared to clinician review, which in itself yielded inconsistent case determination. The study demonstrated the need for the developed standardized definitions for AEFI and their usefulness in passive surveillance.

1. Background

The Brighton Collaboration (http://www.brightoncollaboration.org) is an international voluntary collaboration with the primary goal of facilitating the development, evaluation, and dissemination of information about adverse events following immunization (AEFI) about vaccines used in human populations through the development of standardized case definitions of AEFI with corresponding guidelines for standardized data collection, analysis, and presentation [1–6]. The Brighton Collaboration case definitions (BCCDs) are intended to increase comparability of AEFI data across surveillance systems or research studies for retrospective analyses, and through standardized prospective data collection in a variety of geographic settings. Standardized terminology and case definitions can improve our ability to monitor AEFI and vaccine safety, as well as compare results of studies of different populations, vaccines, and geographic regions [1,2,4,7].

The first six BCCDs, which were published in 2004, were fever, generalized convulsive seizure, hypotonic–hyporesponsive episode (HHE), intussusception, nodule at the injection site, and persistent crying [8–13]. The Brighton Collaboration has since published 14 additional papers of standardized case definitions of AEFI, as well as a revised version of the HHE case definition [14] based on further review by an international vaccine safety working group

☆ *Support*: The study was funded in part through the Unmet Needs Awards by the National Vaccine Program Office (http://www.dhhs.gov/nvpo/research/awards_2003.htm).

* Corresponding author. Tel.: +1 404 639 8073; fax: +1 404 639 4441.
E-mail address: kfk0@cdc.gov (K.S. Kohl).
[1] Present address: Division of Global Migration and Quarantine, National Center for Preparedness, Detection, and Control of Infectious Diseases, Mailstop E-61, Centers for Disease Control and Prevention, USA.
[2] Present address: Division of Foodborne, Bacterial and Mycotic Diseases, National Center for Zoonotic, Vector-Borne, and Enteric Diseases, Centers for Disease Control and Prevention, USA.

0264-410X/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.vaccine.2008.09.002

(http://www.cioms.ch/jan2008_current_programme_and_planned_activities.pdf). Minor revisions to the all published definitions can be found on the Brighton Collaboration website as attachments to each published version.

Several studies have investigated the use of study-specific case definitions including BCCD in identifying and standardizing AEFI for HHE, intussusception, encephalitis, fever, anaphylaxis, and acute encephalopathy, as well as evaluating AEFI in association with specific vaccines, (e.g., the Japanese encephalitis, pertussis, and rotavirus vaccines) [15–24]. However, these studies tend to use a case-series approach, comparing cases with BCCD that were felt to be highly probable by another screening method (e.g., ICD-9, clinician diagnosis) or evaluating responses to vaccines. These studies can improve our understanding of the performance of the definitions in respective study settings and provide suggestions for revising the definitions, but cannot provide sensitivity/specificity parameter estimates or compare estimates for multiple BCCD within the same population of reports.

The primary objective of our study was to evaluate the first six published case definitions (Box 1) with respect to their applicability, reliability, sensitivity, and specificity within the U.S. Vaccine Adverse Event Reporting System (VAERS), one of the largest passive AEFI surveillance systems globally. The secondary objective was to propose methodology for doing so that is generalizable to other surveillance systems and research studies.

2. Methods

The Vaccine Adverse Event Reporting System, which was established in 1990, is operated as a collaboration between the Food and Drug Administration (FDA) and the Centers for Disease Control and Prevention (CDC) (http://vaers.hhs.gov/) [25]. This passive surveillance system receives over 14,000 adverse event reports each year from health care providers, vaccine recipients, parents/guardians of vaccine recipients, vaccine manufacturers, and other interested persons. Objectives of VAERS include identifying signals of previously unrecognized adverse events, improving characterization of known events, and identifying risk factors for AEFI. As with other passive surveillance systems, VAERS cannot provide unbiased estimates of population prevalences of AEFIs, is limited by the variability of detail in reports, and cannot be used to prove causality. Another limitation of VAERS is the inherent absence of case definitions to standardize measurement of AEFI and allow comparison between time periods as well as with other surveillance systems or studies. VAERS is limited by the literal translation process by which its coding terminology is applied to AEFI descriptions by coding staff.

Until January 2007, reports submitted to VAERS were coded according to Coding Symbols for a Thesaurus of Adverse Reaction Terms or COSTART [26]. COSTARTS included 1048 codes for symptoms text and diagnosis in VAERS reports. The analysis for this study was conducted by using COSTARTS. In January 2007, the VAERS coding system was converted to the Medical Dictionary for Regulatory Activities (MedDRA) (http://www.meddramsso.com/MSSOWeb/index.htm), a globally used coding system for drug and vaccine adverse events with several thousand coding terms. In both systems, each report can be coded with several terms depending on the description of the AEFI provided on the report.

All reports submitted to VAERS during 1991–2003 were eligible if at least one symptom was present and coded, a birth date was available, and the age at the time of onset of the AEFI was $\leq$18 years. Persons $\leq$ 18 years were selected because four of the six BCCDs under evaluation were most common among children and young adolescents, and use of the same age range allowed comparison between BCCD. Data on vaccines administered were not required as AEFI characteristics were evaluated independent from the vaccinations provided.

All six BCCDs were evaluated for their applicability to VAERS reports by assessing the ease with which criteria of the definitions could be applied to coded reports and text strings in the VAERS data base. BCCDs were also assessed for their reliability when being applied to VAERS reports by different reviewers; interrater reliability between clinical reviewers blinded to the definitions was equally assessed. And finally, sensitivity and specificity were assessed by using independent clinical reviews of VAERS reports as the gold standard against which to measure the application of the BCCD to the same reports by nonclinicians.

Several steps were taken to evaluate the six BCCDs. These include:

(1) Developing search strategies for reports with high and low probability of representing the AEFI under evaluation. Search strategies were first required to identify VAERS reports with the potential to be an AEFI of interest. To develop the search strategies, each BCCD was deconstructed into its component symptom categories. For example, three categories were included for the HHE BCCD: limpness, reduced responsiveness, and pallor. COSTARTS [20] encompassing each symptom within each category were identified by using a multidisciplinary process that included pediatricians and pediatric subspecialists, epidemiologists, and vaccine safety specialists. For each category, component criteria were developed, each composed of a series of COSTARTS. For example, for HHE, the COSTARTS for the criterion of *Limpness* included hypotonia and paralysis flaccid; the criterion *Reduced responsiveness* included coma, stupor, and syncope; and the criterion *Pallor* included pallor, cyanosis, and hypoxia. Combinations of symptoms and categories were mapped from the BCCD onto each COSTART combination, and eligible reports were analyzed to describe the proportion of reports potentially representing the AEFI. For example, where the Level 1 of the HHE BCCD requires "sudden onset of limpness (i.e., muscular hypotonia) and reduced responsiveness (i.e., hyporesponsiveness) or unresponsiveness and pallor or cyanosis," query logic identified reports with one of each of the selected criteria. Due to the broad nature of COSTARTS, when detailed clinical, surgical, or radiographic information was required (as for intussusception), BCCD levels requiring only clinical criteria that could be described by COSTARTS were used for search strategies. Reports having the combination of COSTARTS required for the BCCD were considered "high probability"; that is, reports that were likely to have identified reports suggestive of the AEFI or likely to match the BCCD. Search strategies for "low-probability" reports were then developed. These search strategies used COSTARTS describing symptoms similar to those in "high-probability" reports. For example, for HHE, reports with any COSTART of *Loss of consciousness/reduced responsiveness*, *Generalized motor manifestations*, *Limpness* or *Pallor* were eligible for comparison. This strategy reduced the bias in favor of the BCCD that would emerge if too-dissimilar reports were being presented (e.g., HHE vs. fever). "High-probability" reports and "low-probability" reports are conceptually similar to "cases" and "controls," respectively; however, the COSTART terms were not the criteria themselves, and were used only as a means to identify reports for further review.

(2) Assess the applicability and validity of these search strategies qualitatively and quantitatively. Applicability assessment included (a) demonstrating that COSTARTS would be an appropriate method for assessment of BCCD within VAERS compared