Standardizing the Structure of Stroke Clinical and Epidemiologic Research Data

The National Institute of Neurological Disorders and Stroke (NINDS) Stroke Common Data Element (CDE) Project

Jeffrey L. Saver, MD; Steven Warach, MD, PhD; Scott Janis, PhD; Joanne Odenkirchen, MPH; Kyra Becker, MD; Oscar Benavente, MD; Joseph Broderick, MD; Alexander W. Dromerick, MD; Pamela Duncan, PhD; Mitchell S.V. Elkind, MD; Karen Johnston, MD; Chelsea S. Kidwell, MD; James F. Meschia, MD; Lee Schwamm, MD; for the National Institute of Neurological Disorders and Stroke (NINDS) Stroke Common Data Element Working Group

Background and Purpose—The National Institute of Neurological Disorders and Stroke initiated development of stroke-specific Common Data Elements (CDEs) as part of a project to develop data standards for funded clinical research in all fields of neuroscience. Standardizing data elements in translational, clinical, and population research in cerebrovascular disease could decrease study start-up time, facilitate data sharing, and promote well-informed clinical practice guidelines.

Methods—A working group of diverse experts in cerebrovascular clinical trials, epidemiology, and biostatistics met regularly to develop a set of stroke CDEs, selecting among, refining, and adding to existing, field-tested data elements from national registries and funded trials and studies. Candidate elements were revised on the basis of comments from leading national and international neurovascular research organizations and the public.

Results—The first iteration of the National Institute of Neurological Disorders and Stroke (NINDS) stroke-specific CDEs comprises 980 data elements spanning 9 content areas: (1) biospecimens and biomarkers; (2) hospital course and acute therapies; (3) imaging; (4) laboratory tests and vital signs; (5) long-term therapies; (6) medical history and prior health status; (7) outcomes and end points; (8) stroke presentation; and (9) stroke types and subtypes. A CDE website provides uniform names and structures for each element, a data dictionary, and template case report forms, using the CDEs.

Conclusions—Stroke-specific CDEs are now available as standardized, scientifically vetted, variable structures to facilitate data collection and data sharing in cerebrovascular patient-oriented research. The CDEs are an evolving resource that will be iteratively improved based on investigator use, new technologies, and emerging concepts and research findings. **(*Stroke*. 2012;43:967-973.)**

Key Words: database ■ education ■ educational campaigns

Downloaded from http://ahajournals.org by on March 4, 2022

The mission of the National Institute of Neurological Disorders and Stroke (NINDS) is "to reduce the burden of neurological disease." Supporting translational, clinical, and population research in stroke is fundamental to this mission, as stroke is the single greatest nervous system cause of death and disability, both in the United States and worldwide.[1–3]

Accordingly, the NINDS supports a diverse array of translational, clinical trial, epidemiological, and additional patient-oriented research in cerebrovasular disease, which have had a substantial beneficial effect on health policy, clinical care, and patient outcomes.[4] However, the fullest potential benefit of these research endeavors has not been realized due to the absence of uniform, widely accepted formats to characterize demographic, disease, care process, and outcome variables. Data elements are often characterized in varying manners in different studies, hampering cross-study comparisons, recognition of population differences, data sharing, and pooled analyses.

Received August 3, 2011; accepted December 9, 2011.

Bo Norrving, MD, was the Guest Editor for this paper.

From the National Institute of Neurological Disorders and Stroke (S.W., S.J., J.O.), National Institutes of Health, Bethesda, MD; the University of Washington School of Medicine (K.B.), Seattle, WA; the University of British Columbia (O.B.), Vancouver, British Columbia; the University of Cincinnati (J.B.), Cincinnati, OH; Georgetown University School of Medicine (A.D.), Washington, DC; Duke University (P.D.), Durham, NC; Columbia University (M.S.V.E.), New York, NY; the University of Virginia (K.J.), Charlottesville, VA; Georgetown University (C.S.K.), Washington, DC; Mayo Clinic College of Medicine (J.F.M.), Jacksonville, FL; and Harvard Medical School (L.S.), Boston, MA.

The online-only Data Supplement is available with this article at http://stroke.ahajournals.org/lookup/suppl/doi:10.1161/STROKEAHA.111.634352/-/DC1.

Correspondence to Jeffrey L. Saver, MD, UCLA Stroke Center, 710 Westwood Plaza, Los Angeles, CA 90095. E-mail jsaver@ucla.edu

© 2012 American Heart Association, Inc.

Stroke **is available at http://stroke.ahajournals.org** **DOI: 10.1161/STROKEAHA.111.634352**

To harmonize data collected across diverse translational, clinical, and population studies, NINDS began the Common Data Element (CDE) Project in 2006.[5] The project aims to standardize naming, definitions, data structure, and response options for all variables commonly used in NINDS-funded patient and population research. The CDE project complements study-level guidelines from the Enhancing the QUAlity and Transparency Of health Research (EQUATOR) Network.[6] However, where EQUATOR guidelines focus on the reporting of study level results, the CDE Project focuses on the collection and reporting of individual variable level results, providing guidance for data collection, and facilitating data sharing, at a more granular patient level.

The overall CDE Project has 4 primary goals: (1) disseminate standards for the collection of data from participants enrolled in studies of neurological diseases; (2) create easily accessible data-collection tools for investigators that are ready to use "off the shelf"; (3) encourage focused and simplified data collection to reduce burden on investigators and practice-based clinicians to increase clinical research participation; and (4) improve study quality and reduce cost of data entry, cleaning and analysis by providing uniform data descriptions and tools across NINDS-funded clinical studies of treatment for neurological diseases.[5,7] The anticipated benefits of the CDE Project are multiple, and include (1) rapid and efficient study start-up by allowing investigators access to appropriate data elements, definitions, and case report form templates; (2) improved patient safety by facilitating development of common report templates that can be submitted to oversight committees such as Data and Safety Monitoring Boards (DSMBs); (3) enriched data sharing and data aggregation by employing standard definitions and common forms; and (4) wide adoption of common outcome measures (eg, functional, cognitive) that may be relevant across the neurological diseases.[5,7]

The CDE Project first developed a set of general CDEs commonly collected in all neuroscience clinical studies, including demographic information, medical history data, medication use, and data needed for safety reporting.[8] Next, development of disease-specific CDEs was undertaken. This report describes the process and outcome for the development of Stroke CDEs.

Methods

NINDS convened a Stroke CDE Working Group consisting of 52 experts with experience in National Institutes of Health (NIH)-funded cerebrovascular disease clinical and population research or international initiatives in stroke data integration. The Working Group comprised extramural clinical scientists, intramural NINDS researchers, and the NINDS CDE team, supported by a contracted clinical research organization, KAI Research Inc (Rockville, MD). Participants represented a broad spectrum of specialties and research domains, including adult and pediatric stroke; acute care and prevention; outpatient, prehospital, emergency department, inpatient, and neurointensive care; neurology, neurosurgery, emergency medicine, radiology, and hematology; nursing, physicians, biostatisticians, public health, and epidemiologists. The input of experienced research coordinators and scientists involved with the creation of data and specimen repositories and of an industry scientist was also solicited. Members from diverse institutions were selected to ensure the inclusion of different perspectives and experience. Specific federal agencies, such as the Food and Drug Administration (FDA) and the Centers for Disease Control (CDC), were requested to appoint representatives. The CRO was responsible for scheduling the Working Group conference calls and working on administrative action items such as development of the CDE data dictionaries based on the Working Group's content recommendations. Wherever possible efforts were made to harmonize the Stroke CDE work product with existing data structure recommendations, including from the Specialized Program of Translational Research in Acute Ischemic Stroke (SPOTRIAS) Common Clinical Database, the Centers for Disease Control and Prevention's Paul Coverdell National Acute Stroke Registry (Coverdell), and the American Heart Association/American Stroke Association Get With the Guidelines (GWTG)–Stroke national quality improvement registry.

The final CDE products were to include:

(1) Listing of standardized data elements that can be incorporated into any data system

(2) Data Dictionary with data specifications for each element, including definition, format of response (numeric versus multiple choice versus free text), recommended response options (for multiple choice), explanations, and references

(3) Template study forms in Microsoft WORD and PDF format

(4) Manuals of Procedures (MOPs), as appropriate

(5) Website that facilitates access to the CDE data elements, data dictionary, forms, and MOPS (http://www.commondataelements.ninds.nih.gov/Stroke.aspx).

When a measure with copyright protection is included in the CDE recommendations, the NINDS CDE Team contacts the copyright holder of the proprietary instrument seeking permission to (1) post the actual instrument to the NINDS CDE Web site or (2) post a link that will direct people to a Web site or paper where the instrument may be found. If permission is granted, the instrument is made available on the CDE Web site, and if not, CDE users are at a minimum provided with a Web site link and reference that they can use to access the instrument themselves.

The Working Group first met at the 2009 Annual American Academy of Neurology meeting in Seattle, WA. Early on, the group decided that the Stroke CDEs would not specify a set of mandatory data elements to be included in all studies whatever their purpose and character. The selection of elements most appropriate for a study would be performed by the study investigators. However, the CDEs would specify a mandatory structure for each data element, so that if investigators chose to collect a specific variable in their study, they would need to use the standard definition and format of that variable or to specify a reason for departing from the standard version. The goal was to be descriptive, not prescriptive, with regard to variable selection, but fully prescriptive with regard to variable structure. The Working Group classified data variables into 3 tiers: Core, Supplemental, and Exploratory. Core variables are data elements commonly used in stroke clinical and population research that many studies should consider employing. Supplemental variables are more specialized data elements that are appropriate for particular patient populations or study subtypes. Exploratory variables show promise but were judged not fully ready for widespread use in clinical research studies at this time (eg, additional validation required first).

Nine stroke-specific CDE-domains were identified: (1) biospecimens and biomarkers; (2) hospital course and acute therapies; (3) imaging; (4) laboratory tests and vital signs; (5) long-term therapies; (6) medical history and prior health status; (7) outcomes and end points; (8) stroke presentation; and (9) stroke types and subtypes. Each domain was assigned to a subcommittee with a subgroup chair. Each subcommittee was provided with a starting candidate pool of data elements taken by the CDE team from NINDS-funded clinical trials and epidemiological studies, and asked to add to, subtract from, and refine these elements as they believed were warranted. The subcommittees worked to consensus by iterative discussion, without use of Delphi or other formal methods of consensus attainment. Portfolios of CDEs generated by the subcommittees were then circulated for comment to all Working Group members and further revised. A draft final CDE panel was posted on the internet for public

Downloaded from http://ahajournals.org by on March 4, 2022

comment in August and September 2010, and comments were specifically solicited from 18 stroke organizations worldwide, including the World Stroke Organization, the American Stroke Association, the National Stroke Association, the Canadian Stroke Consortium, the European Stroke Organization, the Australasian Stroke Society, the Indian Stroke Association, the Japan Stroke Society, the National Stroke Foundation of Australia, and the Virtual International Stroke Trials Archive (VISTA). A total of 23 comments were received, leading to 12 revisions in the draft CDEs. The CDEs were published on the Web in December 2010.

Results

The first published version of the Stroke CDEs contains 980 data elements covering 9 content domains. Among these, 163 (16.6%) data elements were designated as Core, 808 (82.4%) as Supplemental, and 9 (1%) as Exploratory. The full set of Stroke CDEs, along with Case Report Form modules and Guidelines, are available at www.commondataelements.ninds.nih.gov/Stroke.aspx. Brief reports of key principles and decisions made by the content subgroups in generating CDEs in the 9 domains are provided below.

Biospecimens and Biomarkers

Biospecimens were considered human biological specimens of any tissue or fluid type. A stroke biomarker was defined by consensus agreement of participants as any assay performed on a biospecimen for the purpose of diagnosing stroke, assessing stroke severity, predicting outcomes after stroke, or predicting response to treatment of stroke. Stroke biomarkers have been studied for all of these purposes.[9–12] The subcommittee divided data elements into two broad classes: (1) preanalytical data elements to consider when collecting, storing, or performing assays on biospecimens and (2) core biomarkers for stroke-related research. Special emphasis was placed on preanalytical variables and biospecimens that might be practical to obtain in the context of clinical trials (eg, plasma). A draft list of preanalytical data elements was cross-checked for completeness with published consensus statements.[13,14]

A systematic PubMed literature search yielded 617 articles, including protein-, genotype- and metabolite-based assays. A list of biomarkers was assembled. Individual biomarkers were then classified into 3 groups. Core biomarkers were FDA-approved and available in automated platforms. Secondary biomarkers were non–FDA-approved but had substantial evidence supporting their validity as stroke-relevant biomarkers. Tertiary biomarkers were non–FDA-approved and lacked substantial evidence supporting their validity as stroke-relevant biomarkers. Core biomarkers are most highly recommended, as they generally can be ordered through a CLIA-certified laboratory, are available on automated, high-throughput platforms, and have performance characteristics that are rigorously defined and scrutinized.

Preanalytical variables that were seen as potentially worthy of specific attention in the context of stroke biomarkers research include time from stroke onset to sample collection, stroke type (ischemic versus hemorrhagic), use of medications common to the stroke population, brain and cerebrovascular imaging findings, and stroke-specific outcomes. It was recognized that the value of a biorepository largely relates to the level of detail characterizing sample collection, processing, and storage and the clinical data elements linked to specific samples.

A major challenge in recommending standards for collecting and reporting preanalytical data elements for biospecimens is that what constitutes a crucial data element largely depends on the nature of the biomarker or bioassay. Another general obstacle in recommending standards for biomarkers in stroke trials and epidemiology studies is that the reliability requirements of any biomarker may vary according to the goals of the study (eg, treatment versus prevention) and to the purpose of the biomarker (eg, screening patient for eligibility versus defining subgroups in exploratory analyses).

The first iteration of CDEs includes 11 Core and 31 Supplemental biomarker data elements as well as the recommendation of 14 FDA-approved biomarker assays and 16 non–FDA-approved emerging biomarker assays.

Hospital Course and Acute Therapies

The Hospital Course and Acute Therapies Subgroup identified and defined data elements common to trials and health outcome studies of acute stroke therapies. The Subgroup incorporated many of the data elements from GWTG-Stroke and Coverdell, supplemented with elements found to be commonly collected by NINDS-funded stroke studies according to a review of case report forms from 34 studies. The resulting recommendations specify 55 CDEs spanning the following areas: IV Thrombolytic Therapy, Intra-arterial (IA) Thrombolytic Therapy, Intrasinus Thrombolytic Therapy, Other Thrombolytic/Reperfusions, In-Hospital Treatments and Discharge Data.

The Subgroup's CDE development work was complicated by the fact that the data collected about the hospital course and acute therapies administered are usually dictated by the specific research question(s), the type of study, the type of stroke, and the study population (eg, adult versus pediatric) being studied. It is virtually impossible to create a menu of CDEs that will satisfy all the possible permutations. The CDEs in this topic area will undoubtedly need to evolve over time as new treatments for stroke are introduced and become commonplace. Currently the recommendations of this Subgroup include few CDEs that would be considered pediatric-specific elements and this is an area for additional work in the next version of the Stroke CDEs.

Imaging

The imaging working group divided the stroke neuroimaging CDEs into 3 subdomains: (1) brain parenchymal imaging (further subdivided into ischemia and hemorrhage sections), (2) perfusion and penumbral imaging, and (3) vessel imaging. For the parenchymal and perfusion/penumbral imaging sections, modalities covered included computed tomography (CT) and MRI, the modalities most commonly available in the acute hospital setting. For vessel pathology, imaging modalities included catheter angiography, MR angiography, CT angiography, and sonography (carotid Doppler, transcranial Doppler, and transcranial color-coded sonography). Pediatric elements were included in each section.

The group chose to focus on CDEs that were likely to be common to most cerebrovascular studies and trials that

Downloaded from http://ahajournals.org by on March 4, 2022

incorporate neuroimaging. The group decided to avoid prescribing specific recommendations regarding imaging parameters or sequences and instead incorporated a limited number of fields covering technical parameters used in each acquisition. Moreover, for topics that included multiple approaches to definitions or fields in flux (eg, defining ischemic penumbra), the group opted to be inclusive of various definitions rather than make a single recommendation. Wherever possible, validated rating scales in common use were included in the CDEs. For items requiring clinical information (eg, symptomatic hemorrhagic transformation), the CRFs were designed to be linked to data provided by other subgroups (eg, NIH stroke scale scores). Future iterations may incorporate additional perfusion imaging modalities (SPECT, PET, etc) and/or additional elements or rating scales. The CDE specifications focused on data elements rather than picture archiving formats (eg, Digital Imaging and Communications in Medicine, DICOM). The first iteration of CDEs included 285 neuroimaging elements.

Laboratory Tests and Vital Signs

The group generated a list of common measures that would be available by routine examination of study participants, such as blood pressure, pulse, height and weight, and those measures that would only be available in limited clinical situations, such as intracranial pressure. Laboratory studies that could provide information regarding stroke risk, stroke mechanism, stroke outcome and the efficacy and effectiveness of interventions to prevent or treat stroke were compiled.

The compilation of a list of laboratory tests and vital signs for use in stroke studies was complicated by the fact that the important laboratory tests and vital signs to be in a study are usually dictated by the type of study (prevention versus treatment versus epidemiological), the type of stroke (ischemic, intraparenchymal hemorrhage, subarachnoid hemorrhage, and venous), and the study population (adult versus pediatric) being addressed. Further, the number of laboratory studies available to characterize study participants is nearly infinite and the relative importance of each tends to change as new research emerges. Any document outlining the necessary laboratory profile for a stroke study must thus be considered a work in progress and be routinely updated to account for advances in the field. In this first iteration, the work group did not harmonize with the LOINC (Logical Observation Identifiers Name and Codes) and UCUM (Unified Codes for Units of Measure) metrics, but will consider doing so in the future.

The first iteration of CDEs included 56 laboratory test and vital sign elements.

Long-Term Therapies

The Long-Term Therapies Subgroup focused on measures useful in studies of secondary stroke prevention and rehabilitation. Wherever possible, the group focused on measures already collected for other purposes or ones that incurred manageable collection burden. The group prioritized measures that could be used in both adults and children but acknowledged that there would be unique measures for each group. A total of 27 long-term therapy data elements were included in the first CDE generation.

For secondary prevention, the group suggested measures that evaluate whether risk factors are detected, whether a treatment plan is prescribed, the adherence to a prescribed plan, and if objective and subjective measures of risk factors actually improve. Substantial effort was spent on health-related behaviors, including diet, smoking, exercise, and medication adherence. Also included was assessment of the severity of clinical disturbance associated with recurrent stroke and whether adherence to risk factor reduction affected the severity. For rehabilitation, the focus was on variables describing the sophistication and intensity of the clinical setting, content and amount of treatment, and timing of rehabilitation treatments after stroke. Tracking of assessment for rehabilitation, impact of payer on treatment, and durable medical equipment use were also included.

Several challenges and areas for refinement were identified. The need for practical, objective, and quantifiable measures of health-related behaviors was repeatedly raised. Assessment of longer-term or repeated interventions to lessen stroke related disability was difficult; such interventions included bracing, surgical procedures, botulinum toxin injections, electric stimulation, and special education settings for children. Other complexities included the tracking of changes in prescribed cardiovascular risk factor plans and tracking of delayed stroke prevention procedures such as cardiac and neurovascular surgical procedures in children and carotid procedures in adults.

Medical History and Prior Health Status

The Medical History and Prior Health Status group addressed subject health before study enrollment. The group reviewed CRFs from 34 NINDS-funded studies as well as several well-known epidemiological studies (eg, Framingham Heart Study, Multi-Ethnic Study of Atherosclerosis, Northern Manhattan Study) and identified health history data elements commonly collected across these studies. The group also aligned data elements as much as possible with GWTG–Stroke and Coverdell and incorporated the few data elements from these related efforts pertaining to medical history and prior health status.

The resulting recommendations included 225 data elements, grouped into 7 subdomains: Demographics, Medical History (including elements related to stroke comorbidities and risk factors), Medication History, Social/Lifestyle History (ie, use of alcohol, tobacco, and illicit drugs), Family History, Pregnancy/Perinatal History (includes elements relevant for neonatal stroke studies), and Pre-Stroke Functional Status. The Subgroup classified 58 elements as Core. They also labeled several of their elements as pediatric-specific to assist investigators who plan to study stroke in neonates (up to 1 month) and/or children (1 month up to 18 years).

Outcomes and End Points

Selecting the most common data elements for stroke outcome is a difficult task, given the heterogeneity of stroke etiology, symptoms, severity, and recovery. Despite these complexities, the common data elements outcomes workgroup used several strategies to select the best measures. The group incorporated the World Health Organization, International

Downloaded from http://ahajournals.org by on March 4, 2022

Classification of Functioning, Disability and Health (ICF); required that measures had sound psychometric properties of reliability, validity, and responsiveness, and established clinically important differences; and selected measures that covered the range of stroke related symptoms, including motor, sensory, perceptual, language, cognition, and emotion. The group ensured that the selected assessments of activities and functional limitations included measures not suffering from floor and ceiling effects. Well-established global ratings, the modified Rankin for functional status and the EuroQOL for quality of life, were included. A flexible battery for cognitive assessments was incorporated, including items capturing deficits in executive function, a common manifestation of vascular cognitive impairment, consistent with NINDS and Canadian Stroke Network harmonization standards.[15] Also included were recommendations and standardized methods for stroke adjudication from the Reasons for Geographic And Racial Differences in Stroke (REGARDS) study. All measures selected for Core measures are well established; however the work group identified the development of five exploratory measures that might be considered as appropriate core measures in the future.

The first version of CDEs includes 167 Outcomes and End Point data elements. Among these are 36 standardized measures, including 9 designated as Core.

Stroke Presentation

The Stroke Presentation Subgroup identified and defined data elements for stroke (ischemic and hemorrhagic) and transient ischemic attack (TIA) patients that would be commonly available at presentation to a hospital or health provider. The Subgroup's CDE development work was complicated by the fact that the data sources available for adult ischemic stroke trials far outnumbered those available for hemorrhagic stroke, TIA or pediatric strokes of any type.

Wherever adequate definitions existed, they were adopted with a citation of the source. Where these definitions needed further specification or clarity, the root definition was retained, but a modifier was added to allow it to be collected in the CDE but still compared with its parent term in other trials or registries. Rarely, the CDE subgroup proposed a new element and definition, such as brief assessment of gait or monocular visual loss. The group steered away from capturing highly specific clinical signs and symptoms which might be important in isolated trials and collected in a customized manner for that purpose. The goal was to identify elements that would be useful to many individual studies and also be valuable at the level of reuse—study interoperability and data aggregation.

Specific aspects of the neurological examination were drawn from subscales within the NIH Stroke Scale, Glasgow Coma Scale, and other less well-known specialty scales as needed. Pediatric or neonatal modifiers were included where appropriate or available. A limited data set of prehospital elements was included given the importance of these factors in acute intervention trials or those studying patterns of access to care.

Overall, 61 elements were defined, with a large number of additional candidate elements deferred for further research and validation before they were considered appropriate for inclusion in the CDE. The selected CDEs captured data in the following areas: prehospital/emergency medical system course; hospital arrival/admissions description; diagnosis; stroke symptoms/comorbid events; pediatric hospital admission symptoms/acute stroke presentation; neurological/physical findings—NIH stroke scale; and additional neurological/physical findings.

Stroke Types and Subtypes

Classification of stroke types and subtypes is a fundamental activity in epidemiological, genetic, clinical trial, and outcome research as well as clinical care. Since the advent of brain imaging in the 1970s, the classification of stroke types has combined elements of tissue-based evaluation (brain imaging or pathology) as well as clinical presentation (time-based symptoms). However, although the diagnosis of hemorrhagic stroke has been tissue-based since the CT era, classifications of ischemic stroke and TIA in clinical research studies have been primarily time-based with a 24-hour dividing line between ischemic stroke and TIA. The evolution and widespread availability of MR brain imaging, particularly diffusion imaging, demonstrated that the classification of ischemic stroke must start with a tissue-based definition and use a time-based definition when brain imaging was unavailable or not revealing. This recommendation was recently part of an ASA Scientific Statement regarding transient ischemic attack.[16]

The group reviewed the literature for the most frequently used stroke classification schemes for the major stroke types, ischemic stroke subtypes, and pediatric stroke. The identified CDEs provide a single classification system for the major stroke types. This definition includes both tissue-based and time-based clinical information. For ischemic stroke subtypes, several classification systems are detailed with their strengths and weaknesses because the decision to use a given classification system will depend on the population under study, the depth of diagnostic testing in a given population, and goals of a given study. Finally, classification of pediatric stroke deserved a separate section because the causes of pediatric stroke, particularly in the perinatal period, are very different from causes in adults. Systems for classification of pediatric stroke are in evolution and probably will be further clarified in upcoming revisions to the CDEs.

The initial iteration of CDEs includes 62 Stroke Types and Subtypes elements and 18 stroke classification schemes.

Discussion

The Stroke CDEs are designed to assist researchers in the design, implementation, aggregation, and interpretation of translational, clinical, and population research. The publication of the first generation of Stroke CDEs is an important milestone in a larger undertaking to promote the adoption of CDEs as scientifically vetted, standardized, building blocks of clinical cerebrovascular research. The Stroke CDEs are expected to be dynamic tools that are curated and updated as needed, so that they best serve the needs of investigators and the public health of US and global citizens. The NINDS has invited feedback from the cerebrovascular disease research

Downloaded from http://ahajournals.org by on March 4, 2022

community, both in the United States and internationally, to iteratively improve these data tools. Program announcements for new NINDS clinical trials (PAR-11–173 and PAR-10–199) request applicants use the CDEs in their CRFs and data management systems when appropriate, and 2 NINDS trials are already piloting use of the CDEs: the Stroke, Hyperglycemia Insulin Network Effort (SHINE) Trial and Thrombolysis in Pediatric Stroke (TIPS). Feedback from implementation of CDEs in actual clinical trials and epidemiological studies will be crucial to refine the data elements to improve their clarity and utility.

The CDEs have the potential to interface efficiently with electronic medical records, which increasingly permit capture of structured data embedded within routine care processes. With appropriate ethics committee oversight and privacy protections in place, data from electronic medical records may be used to populate data fields of research case report forms, reducing duplicative data elicitation and entry processes at the individual site level.

Across sites and studies, use of controlled and defined data can facilitate data pooling and development of neurovascular, and general medical, informatics networks in the United States and worldwide. The importance of harmonizing stroke clinical research elements has been recognized internationally as well as in the United States, and initiatives are underway to maximize alignment of the CDEs with international cerebrovascular registries, including VISTA, the Safe Implementation of Treatments in Stroke (SITS) Registry, and the World Health Organization International Classification of Disease revision 11 Stroke Working Group. The Stroke CDEs join already deployed NINDS CDEs in the areas of epilepsy,[7] traumatic brain injury,[17] spinal cord injury,[18] and Parkinson disease.[5] Additional NINDS CDE initiatives are under way for multiple sclerosis, Friedreich ataxia, Huntington disease, amyotrophic lateral sclerosis, frontotemporal dementia, congenital muscular dystrophy, and headache. The NINDS CDEs are being harmonized with similar initiatives in other disease states, including the National Cancer Institute's Cancer Data Standards Registry and Repository (caDSR) and the FDA's standardized data collection for cardiovascular clinical trials.

Across the working groups, several limitations and challenges emerged in development of the stroke-specific CDEs, most notably difficulties in generating CDEs intended to cover the diversity of patients represented under the disease category of stroke and fundamental distinctions between pediatric and adult stroke. Despite this heterogeneity the group was able to achieve consensus on a substantial body of CDEs.

The NINDS is currently assembling a Stroke CDE Oversight Committee composed of national and international members from academic institutions, industry, and government agencies. The Stroke CDE Oversight Committee will meet annually to review comments received from those using the CDEs and will be responsible for updating the CDE recommendations based on experience as well as the evolving research landscape.

The stroke-specific CDEs will be a success if they are widely adopted by NIH-funded trials and epidemiological studies and an even greater success if their further development and implementation are facilitated by the global stroke research community. Wide use could foster collaboration and data sharing across institutions and studies, reduce misleading signals arising from data definition discrepancies among studies, and generate large patient samples adequately powered to resolve hitherto unaddressable clinical issues. The Working Group hopes that broad deployment of the stroke-specific CDEs will enable more efficient development of novel prevention and acute treatments for stroke and a more rapid advance toward reducing the human burden of cerebrovascular disease.

Working Group Members

The list of all NINDS Stroke CDE Working Group Members is available as online-only Supplemental Data (http://stroke.ahajournals.org).

Sources of Funding

The Common Data Element Project is funded by the National Institute of Neurological Disorders and Stroke, National Institutes of Health (contract No. N01-NS-7-2372).

Disclosures

J.L.S. reports the following: Research grants: NIH, AGA, EV3, and Concentric; consultant/advisory boards: ALA, EV3, Lunseck, Concentric, Talecris, and BrainsGate; institution/employer: UCLA. J.O. reports the following: Institution/employer: NIH employee. O.B. reports the following: Research grant. J.B. reports the following: Research grants: IMS III Trial, UC SPOTRIAS Center, FIA study, C-32 Cerebrovascular Fellowship Training, Genetic and Environmental Risk Factors for Hemorrhagic Stroke, Comparison of Hemorrhagic and Ischemic Stroke Among Blacks and Whites, IRIS Trial, CREST, COSS, and SWISS; honoraria: Genentech, PhotoThera, and Oakstone Medical Publishing. A.W.D. reports the following: Research grant. P.D. reports the following: Institution/employer: Wake Forest Baptist Medical Center. M.S.V.E. reports the following: Research grants: BMS Sanofi Partnership and diaDexus; other research support: NIH/NINDS; speakers bureau: BMS Sanofi Partnership and Genentech; expert witness: GlaxoSmithKline and Organon. K.J. reports the following: Research grant: NIH-NINDS U01NS064498. C.S.K. reports the following: Research grants: NS057401, NS069763, and NS044378. J.F.M. reports the following: Research grant: NINDS Sign U01NS069208.

References

1. World Health Organization. Burden of Disease Statistics. April 11, 2011. http://www.who.int/healthinfo/bod/en/index.html.
2. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL. Global Burden of Disease and Risk Factors. Oxford: Oxford University Press; 2006.
3. Johnston SC, Hauser SL. Neurological disease on the global agenda. *Ann Neurol*. 2008;64:A11–A12.
4. Johnston SC, Rootenberg JD, Katrak S, Smith WS, Elkins JS. Effect of a US National Institutes of Health programme of clinical trials on public health and costs. *Lancet*. 2006;367:1319–1327.
5. National Institute of Neurological Disorders and Stroke. NINDS Common Data Elements. April 11, 2011. http://www.commondataelements.ninds.nih.gov/default.aspx.
6. Simera I, Moher D, Hirst A, Hoey J, Schulz KF, Altman DG. Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. *BMC Med*. 2010;8:24.
7. Loring DW, Lowenstein DH, Barbaro NM, Fureman BE, Odenkirchen J, Jacobs MP, et al. Common data elements in epilepsy research: development and implementation of the NINDS epilepsy CDE project. *Epilepsia*. 2011;52:1186–1191.
8. National Institute of Neurological Disorders and Stroke. General CDE Standards. April 11, 2011. http://www.commondataelements.ninds.nih.gov/General.aspx.

Downloaded from http://ahajournals.org by on March 4, 2022

9. Laskowitz DT, Kasner SE, Saver J, Remmel KS, Jauch EC. Clinical usefulness of a biomarker-based diagnostic test for acute stroke: the Biomarker Rapid Assessment in Ischemic Injury (BRAIN) study. *Stroke*. 2009;40:77–85.
10. Whiteley W, Chong WL, Sengupta A, Sandercock P. Blood markers for the prognosis of ischemic stroke: a systematic review. *Stroke*. 2009;40:e380–e389.
11. Nafria C, Fernandez-Cadenas I, Mendioroz M, Domingues-Montanari S, Hernandez-Guillamon M, Fernandez-Morales J, et al. Update on the Serum Biomarkers and Genetic Factors Associated with Safety and Efficacy of rt-PA Treatment in Acute Stroke Patients. *Stroke Res Treat*. 2011;2011:182783 [Epub June 9, 2011].
12. Downing GJ. Biomarkers and surrogate endpoints in clinical research: definitions and conceptual model. In: Downing GJ, ed. Biomarkers and Surrogate Endpoints: Clinical Research and Applications: Proceedings of the NIH-FDA Conference held on April 15–16, 1999. New York: Elsevier Science; 2000:1–9.
13. Rai AJ, Gelfand CA, Haywood BC, Warunek DJ, Yi J, Schuchard MD, et al. HUPO Plasma Proteome Project specimen collection and handling: towards the standardization of parameters for plasma proteome samples. *Proteomics*. 2005;5:3262–3277.
14. Teunissen CE, Petzold A, Bennett JL, Berven FS, Brundin L, Comabella M, et al. A consensus protocol for the standardization of cerebrospinal fluid collection and biobanking. *Neurology*. 2009;73:1914–1922.
15. Hachinski V, Iadecola C, Petersen RC, Breteler MM, Nyenhuis DL, Black SE, et al. National Institute of Neurological Disorders and Stroke: Canadian Stroke Network vascular cognitive impairment harmonization standards. *Stroke*. 2006;37:2220–2241.
16. Easton JD, Saver JL, Albers GW, Alberts MJ, Chaturvedi S, Feldmann E, et al. Definition and evaluation of transient ischemic attack: a scientific statement for healthcare professionals from the American Heart Association/American Stroke Association Stroke Council; Council on Cardiovascular Surgery and Anesthesia; Council on Cardiovascular Radiology and Intervention; Council on Cardiovascular Nursing; and the Interdisciplinary Council on Peripheral Vascular Disease: the American Academy of Neurology affirms the value of this statement as an educational tool for neurologists. *Stroke*. 2009;40:2276–2293.
17. Thurmond VA, Hicks R, Gleason T, Miller AC, Szuflita N, Orman J, et al. Advancing integrated research in psychological health and traumatic brain injury: common data elements. *Arch Phys Med Rehabil*. 2010;91:1633–1636.
18. Biering-Sorensen F, Charlifue S, Devivo MJ, Grinnon ST, Kleitman N, Lu Y, et al. Incorporation of the International Spinal Cord Injury Data Set elements into the National Institute of Neurological Disorders and Stroke Common Data Elements. *Spinal Cord*. 2011;49:60–64.

Downloaded from http://ahajournals.org by on March 4, 2022