CHEMM Intelligent Syndromes Tool

Evaluation and Validation Plan

Rev 0

June 30, 2014

Prepared on behalf of
The National Library of Medicine, National Institutes of Health
U. S. Department of Health and Human Services
6707 Democracy Boulevard, Suite 510
Bethesda, MD 20892

Version Control

Version	Date	Action	Change tracking	Approved by
0	06/25/2014	Effective date	New	Pertti (Bert) J. Hakkinen, Ph.D. Acting Head, Office of Clinical Toxicology

Contents

1 INTRODUCTION

The Chemical Hazards Emergency Medical Management–Intelligent Syndromes Tool (CHEMM-IST) is a prototype decision support tool for first responders trying to identify the chemical that patients in a mass casualty incident have been exposed to. It was developed by experts in medicine and emergency response under the sponsorship of the National Institutes of Health's (NIH) National Library of Medicine (NLM).

CHEMM-IST focuses only on severe cases and assumes victims have been exposed via airborne routes. Potential toxic effects are assumed to have resulted from:

- Inhalation
- Skin deposition

Since CHEMM-IST is currently a prototype, it should not be used for patient care. However, once thoroughly tested and validated by a wide range of potential users, it is intended for use by basic life support (BLS) and advanced life support (ALS) providers as well as hospital first receivers.

This document describes an evaluation and validation plan for CHEMM-IST that, once completed, will move CHEMM-IST from its current state as a prototype to a product ready for use in an operational response environment.

This is intended to be a living plan. The approach will be modified as appropriate based on experience during its implementation. In addition, elements of the plan will be reapplied as the CHEMM-IST tool is modified and expanded to address additional toxidromes. This plan will be formally managed through a change management process.

2 HOW CHEMM-IST WORKS

CHEMM-IST asks users a series of questions regarding patient status to determine the most probable toxic syndrome (toxidrome) the patient was exposed to. It uses a decision model to constantly update the probability ranking of the toxidromes as the end user answers questions about victim symptoms.

Results are displayed for each syndrome, showing the likelihood of the syndrome's involvement. The syndrome with the highest probability (displayed as a number between 0 and 10, where 10 is the highest) is most likely the syndrome involved. CHEMM-IST also provides reference information, including a description of the toxidrome, recommendations for acute patient care, information from primary resources, and additional resources.

3 PLAN OVERVIEW

To verify its readiness for operational use, CHEMM-IST will be evaluated by subject matter experts (SMEs) and the user community and validated through tests of its functionality and the results it reports, conducted by a wide range of users over a number of scenarios.

The evaluation and validation plan addresses:

- **Methodology validation**: To determine if the methodology is scientifically sound
- **Tool verification**: To determine if the tool accurately reproduces the results of the methodology
- **Functionality evaluation**: To determine if the tool reliably operates under intended conditions
- **Transition test**: To determine if the functional system performs as well as currently accepted approaches under simulated field conditions
- **User acceptance testing**: To determine if the target end users can operate the tool effectively and with satisfaction
- **Publication**: To publicize and communicate the results of the validation and evaluation
- **Outreach**: To proactively share the results of the validation and evaluation with the stakeholder communities

Details appear in the following sections.

NLM and an advisory group will establish an independent testing team to conduct validation tests under the plan. The team will be independent in that no members will be NLM or NIH staff. The independence of the test team will help to assure the end user community that the tests will be conducted without bias—that the NLM is not evaluating itself or its own tool.

The test team will be built around the skill sets needed to conduct the validation tests. An individual member will bring one or more required skill sets to the team. Desired skill sets will include:

- Algorithm development and evaluation
- Software engineering
- Software code development
- Software code testing
- Software performance testing
- User testing
- Statistical analysis
- Test documentation

One or more team members will conduct each set of validation tests, depending on the mix of skill sets required for the test.

4 METHODOLOGY VALIDATION

The methodology validation will examine the scientific basis for CHEMM-IST—the methods and principles on which the tool's technical approach is based.

It will focus on:

- **Syndrome descriptions**: The information about toxidromes provided to users, including the toxidrome description, recommendations for acute patient care, information from primary resources, and additional resources
- **Syndrome characteristics**: The medically observable symptoms associated with each syndrome
- **Scoring and decision model**: The model used to predict syndrome probability based on combinations of known or observed syndrome characteristics

The CHEMM-IST methodology will be validated using three independent reviews (not internally conducted or influenced by the NLM) by:

- An SME advisory group
- A professional organization
- A peer-reviewed publication

Information about the CHEMM-IST methodology will be provided to reviewers in formats that support efficient, comprehensive reviews. Syndrome descriptions will be organized similarly to the CHEMM-IST web page. Syndrome characteristics will be provided in organized tabular format, preferably as an Excel spreadsheet. The scoring and decision model will be fully documented in a form that allows reviewers to understand and follow its logic.

Methodology validation will involve:

1. **Establishing a CHEMM-IST advisory group**. NLM will establish an advisory group that will include government agencies with related missions and expertise (such as the Department of Homeland Security and the National Institute for Occupational Safety and Health). Other SMEs and interested organizations, such as academic institutions and professional organizations, may be included as well. The advisory group will also provide longer-term advice on the development and deployment of CHEMM-IST.

2. **Conducting an SME review of the CHEMM-IST methodology**. The advisory group will examine the CHEMM-IST methodology for validity of the information and approach. The review will include syndrome descriptions, syndrome characteristics, and the scoring/decision model. The advisory group will document its conclusions and them submit to NLM.

3. **Conducting a professional organization review of the CHEMM-IST methodology**. NLM and the advisory group will select one or more professional organizations that represent the interests of the user community for CHEMM-IST (BLS and ALS providers as well as hospital first receivers) to conduct an independent review of the methodology. Candidate organizations include, but are not limited to:

    - The National EMS Management Association
    - The International Association of EMS Chiefs
    - The National Association of Emergency Medical Technicians
    - The National Association of EMS Physicians
    - The International Association of Fire Chiefs
    - The Society for Risk Analysis
    - The American Medical Society of the Uniformed Services
    - The Medical Toxicology Association
    - The American Academy of Pediatrics, Disaster Activity Council and Emergency Section
    - The National Association of Children's Hospitals (attn. Mike Anderson)
    - The American Congress of Obstetricians and Gynecologists

    The selected organization will conduct a review similar to that performed by the advisory group. The organization will document its conclusions and submit to NLM and the advisory group.

4. **Publishing CHEMM-IST methodology in peer-reviewed publication(s)**. NLM or others interested in the program will submit the CHEMM-IST methodology for publication to one or more peer-reviewed professional journals. Candidate journals include, but are not limited to, the:

    - *Journal of Medical Toxicology*
    - *Journal of Toxicology*
    - *International Journal of Toxicology*
    - *Journal of Applied Toxicology*
    - *Military Medicine*
    - *Journal of American Medical Association*
    - *American Medical Informatics Association*
    - *American Medical Health Informatics Association*
    - *Journal of Emergency Medical Services*
    - *Frontiers in Disaster Medicine*
    - *Disaster Medicine and Public Health Preparedness Journal*
    - *Risk Analysis*

- *Journal of Clinical Pediatric Emergency Medicine*
- *Journal of Pediatrics*
- *Journal of the American Medical Association*

Papers will be prepared according to the procedures of the selected publication.

5 TOOL VERIFICATION

Tool verification will determine if CHEMM-IST accurately reproduces the results of the CHEMM-IST methodology.

CHEMM-IST will be verified through:

- A joint review by SMEs and independent software engineers to verify algorithms against the methodology
- White box testing, a method of analyzing software to test its internal structures, to verify that the computer code accurately reproduces the algorithms

Algorithm verification will be completed before white box testing and will involve:

1. **Providing review materials**. NLM will provide the algorithms and methodology to the advisory group for comparison and review.

2. **Conducting an SME review of the CHEMM-IST algorithms**. The advisory group and software engineer will examine the CHEMM-IST algorithms and compare them to the CHEMM-IST methodology.

3. **Documenting the findings**. The advisory group will document the review and submit to NLM.

White box testing will involve:

1. **Establishing an independent test team**. The CHEMM-IST advisory group will establish an independent test team to coordinate, oversee, analyze, and document the results of the tool verification.

2. **Developing test criteria**. The CHEMM-IST advisory group will develop test criteria and document how each objective will be tested. Test criteria will include objectives, scenarios, and environmental control specifications.

3. **Conducting tests**. The test team will use the test criteria to analyze CHEMM-IST's internal structures.

4. **Documenting findings**. The test team will collect, compile, and analyze the results and output from the tests, will document the results, and submit to NLM and the advisory group.

6 FUNCTIONALITY EVALUATION

Functionality evaluation will determine if CHEMM-IST reliably operates under intended conditions.

It will involve:

1. **Establishing an independent test team**. The CHEMM-IST advisory group will establish an independent test team to coordinate, oversee, analyze, and document the results of the functionality evaluation.

2. **Developing a test plan and criteria**. The advisory group will develop test criteria to ensure that testing is robust and logical. Criteria will be developed for:
   a. Load testing
      Load testing will determine the behavior of CHEMM-IST under expected load conditions. The test team will schedule specific test times and collect data on the behavior of CHEMM-IST, including response time and environment and database issues.
   b. Stress testing
      Stress testing will determine the behavior of CHEMM-IST under extreme (beyond design) load conditions and will also determine system breaking points. The test team will schedule specific test times and collect data on the behavior of CHEMM-IST. Data collected will include breaking points, modes of failure, and error handling.
   c. Capacity testing
      Capacity testing will determine the maximum number of simultaneous users that CHEMM-IST can support while meeting performance criteria. The test team will gather performance data while the system is accessed by varying numbers of users. The data will be analyzed to project the number of users at which key performance metrics fall below criteria levels.
   d. Browser testing
      NLM will specify the intended browsers. Browser testing will ensure CHEMM-IST is fully functional on these targets.
   d. Platform testing
      NLM will specify the intended hardware and software (operating systems and plug-ins) with which end users will access CHEMM-IST. Platform testing will ensure that CHEMM-IST can be accessed and operated from these platforms.
   e. Timing testing

Timing testing will determine if CHEMM-IST can perform adequately under a variety of different connection speeds.

3. **Performing the tests**. The test team will gather performance data for comparison to the performance criteria.

4. **Analyzing the results**. The test team will compare the performance data to the performance criteria, develop summary statistics, and develop findings and conclusions.

5. **Documenting the findings**. The test team will document the functional evaluation and results and submit to NLM and the advisory group.

7 TRANSITION TEST

The transition test will determine if CHEMM-IST is ready for field use.

This test will examine whether CHEMM-IST, when applied by end users in simulated field conditions, performs as well as a currently accepted approach to toxidrome recognition. CHEMM-IST will be compared to the mannequin-based simulation training for toxidrome recognition conducted by the Medical Simulation Center at the University of Virginia School of Medicine. The outcomes will be both objective (a comparative performance of the two methods) and subjective (a "closed beta" type end user assessment of the tool's usefulness).

Transition testing will involve:

1. **Establishing an independent testing team**. The CHEMM-IST advisory group will establish an independent team to coordinate, oversee, analyze, and document the results.

2. **Creating test scenarios**. NLM, the CHEMM-IST advisory group, and the Medical Simulation Center will create two test scenarios for each of the four CHEMM-IST toxidromes. The scenarios may be those already in use by the Medical Simulation Center or new scenarios modeled after the Medical Simulation Center approach. Each scenario will describe either a fictional situation or a case study involving the exposure of one or more persons to an unknown chemical. Scenarios will give users enough information about the state of the patient to allow them to determine the likely syndrome.

3. **Identifying end users**. The advisory group will identify two volunteers from each response discipline targeted for CHEMM-IST: BLS providers, ALS providers, and first receivers.

4. **Conducting the tests**. End users will execute each scenario. They will be presented with the characteristics and cues, then apply either the mannequin-based approach or CHEMM-IST (but not both) to identify the toxidrome. Tests will be organized to ensure that all scenarios are tested with each tool by each discipline and that end users use each

approach an equal number of times. The test team will monitor the tests and record the results and the time to decision. The end users will also provide feedback on the tests and complete a survey regarding the performance and value of CHEMM-IST.

5. **Analyzing the results**. The test team will collect and collate the results of the tests, then analyze the results, developing statistics regarding accuracy and speed of each approach. It will also collate and summarize the end user feedback and survey results.

6. **Documenting the results**. The test team will document the tests, results, and analysis and submit to NLM and the advisory group.

8 USER ACCEPTANCE TESTING

The end user acceptance test will determine if the target end users can use CHEMM-IST effectively and with satisfaction. CHEMM-IST will be evaluated in a full operational environment using simulated test scenarios.

This type of external testing is also termed an "open beta test." The outcomes will be both objective (performance of the tool against expected results) and subjective (an end user assessment of the tool usefulness for operations).

End user evaluation will involve:

1. **Establishing an independent test team**. The CHEMM-IST advisory group will establish an independent test team to coordinate, oversee, analyze, and document the results.

2. **Creating an automated data collection process**. The CHEMM-IST development team will create a data collection process that automatically records the scenario played, the toxidrome selected, and the duration of the session.

3. **Creating a user satisfaction survey**. The CHEMM-IST development team will create an online satisfaction survey to evaluate end user satisfaction with CHEMM-IST.

4. **Creating a test page**. The CHEMM-IST development team will create a page within the CHEMM-IST site to facilitate scenario play by end users, the gathering of test data, and the conduct of surveys.

5. **Publishing the test scenarios**. The CHEMM-IST development team will publish the test scenarios on the test page.

6. **Conducting tests**. NLM will announce the end user evaluation to its user community, then monitor and support the evaluation. The test will be operational for two months and

then close. The test site will then become a scenario-based training site for CHEMM-IST.

7. **Conducting surveys**. After each scenario play, the end user will be encouraged to complete a brief online survey concerning their user experience and the value of CHEMM-IST for operational use.

8. **Analyzing results**. The test team will collect and collate the results of the tests and analyze the results, developing statistics regarding accuracy and speed of the new approach. The survey results will also be statistically analyzed.

9. **Documenting the results**. The test team will document the tests, results, and analysis and submit to NLM and the advisory group.

9 PUBLICATION

NLM will publish the results of the validation and evaluation to fully and formally document the process and findings.

Program results will be published in:

- NIH reports
- Conference and professional meeting presentations
- Peer-reviewed publications

Publication efforts will involve:

1. **Documenting interim and final results**. NLM will document the results from all stages of the validation and evaluation program and prepare a summary of the program along with its final results. The reports will be made available to the end user community, government agencies, and academia.

2. **Presenting the CHEMM-IST validation and evaluation program and results at professional conferences and meetings**. NLM will select professional organizations and meetings that serve and reach end user communities, the professional toxicology community, interested government agencies, and academia. Professional organizations may include, but not be limited to:

    - The National EMS Management Association
    - The International Association of EMS Chiefs
    - The National Association of Emergency Medical Technicians
    - The National Association of EMS Physicians

- The International Association of Fire Chiefs
- The Society of Toxicology
- The Society of Environmental Toxicology and Chemistry

3. **Publishing papers about the CHEMM-IST validation and evaluation program in peer-reviewed publication(s)**. NLM will publish papers and articles about the CHEMM-IST validation and evaluation program in one or more peer-reviewed professional journals. Candidate journals include, but are not limited to, the:

    - *Journal of Medical Toxicology*
    - *Journal of Toxicology*
    - *International Journal of Toxicology*
    - *Journal of Applied Toxicology*
    - *Military Medicine*
    - *Journal of American Medical Association*
    - *American Medical Informatics Association*
    - *American Medical Health Informatics Association*
    - *Journal of Emergency Medical Services*
    - *Frontiers in Disaster Medicine*
    - *Disaster Medicine and Public Health Preparedness Journal*
    - *Risk Analysis*
    - *Journal of Clinical Pediatric Emergency Medicine*
    - *Journal of Pediatrics*
    - *Journal of the American Medical Association*

    Papers will be prepared according to the procedures of the selected publication(s).

10 OUTREACH

NLM will socialize the results of the validation and evaluation with the end user community, government agencies, academia, and the public. Outreach activities will include, but not be limited to:

1. **Conducting outreach activities through, in collaboration with, or consistent with the NIH Office of Communications and Public Information.** Through programs coordinated by the NIH Office of Communications and Public Information, NIH communicates matters of science and health to patients, families, scientists, industry, teachers and students, health professionals, and the press. NIH outreach programs employed may include, but not be limited to:

    - NIH Health Information (health tips, podcasts, and fact sheets on diseases and many science/health topics)

- Research Matters (puts breaking science in context, in an understandable format)
- The NIH YouTube channel (hosts videos and interviews about health and science)
- The NIH Twitter Feed, @NIHforHealth (provides news releases and special announcements in 140 characters or less)
- The NIH Facebook page
- Research articles posted at the National Library of Medicine

2. **Socializing the CHEMM-IST validation and evaluation program and results at professional conferences and meetings**. NLM will select professional organizations and meetings that serve and reach end user communities, the professional toxicology community, interested government agencies, and academia. NLM will host informational booths, provide handout literature, and conduct other outreach activities at these meetings. Professional organizations may include, but are not limited to:

    - The National EMS Management Association
    - The International Association of EMS Chiefs
    - The National Association of Emergency Medical Technicians
    - The National Association of EMS Physicians
    - The International Association of Fire Chiefs
    - The Society for Risk Analysis
    - The American Medical Society of the Uniformed Services
    - The Medical Toxicology Association
    - The American Academy of Pediatrics, Disaster Activity Council and Emergency Section
    - The National Association of Children's Hospitals (attn. Mike Anderson)
    - The American Congress of Obstetricians and Gynecologists

11 SCHEDULE

The notional schedule for the CHEMM-IST evaluation and validation is summarized in the following table and visualized in the following schedule chart.

Component	Duration (months from start)		Notes
	Start	End	
Methodology validation: review	0	3	
Methodology validation: publication	0	12	
Tool verification	4	6	Begins after advisory group review of methodology is complete

Component	Duration (months from start)		Notes
	Start	End	
Functionality evaluation	7	9	Begins after tool verification is complete
Transition test	10	12	Begins when functionality evaluation is complete
User acceptance testing	13	15	Begins when transition test is complete
Publication: NIH report	16	18	Begins when transition test is complete
Publication: peer-reviewed papers and professional meetings	16	24+	Begins when user evaluation is complete
Outreach: Office of Communication and Public Information activities and professional meetings	16	24+	Begins when user evaluation is complete

Task	Quarter							
	1	2	3	4	5	6	7	8
Methodology validation: review								
Methodology validation: publication								
Tool verification								
Functionality evaluation								
Transition test								
User acceptance testing								
Publication: NIH report								
Publication: peer-reviewed papers, professional meetings								
Outreach: outreach activities and professional meetings								

12 ABBREVIATIONS AND ACRONYMS

ALS	Advanced life support
BLS	Basic life support
CHEMM-IST	Chemical Hazards Emergency Medical Management– Intelligent Syndromes Tool
EMS	Emergency Medical Service
NIH	National Institutes of Health
NLM	National Library of Medicine
SME	Subject matter expert