

National Practice Assessment: Immunization Information Systems Patient De-duplication

Grant, F.J., Myerburg, S., Wain, J.A., et al. (2012). National Practice Assessment: Immunization Information Systems Patient De-duplication. ***

Abstract

Background: Immunization Information Systems (IIS) are computerized databases that collect vaccination histories for populations within a defined geography. IIS represent an important area of national focus as they are a core component of public health meaningful use. With the increasing use of electronic data interchange, IIS are receiving data from a greater number of external sources. The number of patient duplicate records is increasing. Duplicate patient records undermine data quality and the ability to accurately depict immunization status. This national practice assessment supported the activities of the 2012 The Centers for Disease Control and Prevention (CDC) IIS Patient-level De-duplication Expert Panel. The study results provide a snapshot of IIS patient de-duplication practices.

Methods: A national practice assessment (NPA) was conducted to gather information about patient-level IIS de-duplication software procedures, tools, problems, and practices. The NDA was developed using Delphi techniques. The NPA collected both quantitative and qualitative indicators regarding current IIS de-duplication practices and future plans. The target population was state and territorial immunization information system (IIS) (N=55) implementations. Statistical results were analyzed utilizing SPSS. Qualitative data were analyzed using thematic analysis, expert panel interpretation, and recursive abstraction.

Results: Highlights from the NPA revealed that respondents (n=43) manage IIS containing an average of 4.7 million records and that the causes for patient duplicate records include typographical errors and missing and incomplete data from public immunization providers and hospitals. There are wide variations in the IIS patient de-duplication capabilities and methods in use. De-duplication logic is largely unique and integrated into each instance of IIS software. 48.8% of respondents indicated that replacement or substantial revision of patient de-duplication software would be accomplished within the next 12 to 36 months. Among those most satisfied with their automated processes, the greatest successes have been found in leveraging combinations of deterministic and probabilistic methods. Expert panel discussions of the NPA

findings provided a strong suggestion that the current reliance on human review methods is necessary and beneficial but may become unsustainable as the volume of external data increases. To reduce the need for human review, stronger and more uniform approaches to IIS external provider data submission will be required.

Key Discussion Points

- IIS contain information on over 18 million U.S. children (75%) under six years of age (CDC, 2008).
- IIS are of central importance in Stage 1 and Stage 2 meaningful use initiatives as incented by the Centers for Medicare and Medicaid Services (CMS).
- The goal is to move towards universal adoption of electronic health records (EHRs) supported by a national infrastructure of electronic data interchange which will reduce cost, improve speed, accuracy, and ultimately the quality of patient care.
- Duplicate patient records pose a threat to efficiency and accuracy of IIS operations much like other forms of poor data quality.
- Duplicate records can inhibit the ability to determine patient immunization status and impact patient care. Immunization forecasting may not be correct and may endanger disease protection.
- Duplicate records and poor data quality can adversely affect IIS and public health system credibility.
- IIS patient-level de-duplication methods and processes vary widely.
- Human manual data review represents a significant workload. While manual data review is currently an essential element of IIS data quality management, increasing volumes and resource limitations constrain reliance on this approach.
- Data quality drives continuous evolution. Simple errors in data entry, typos, and transpositions may or may not be able to be consistently detected, depending upon their nature.
- Data coming from both public and private immunization providers contribute equally to duplicate records prob-

lems.

- Improvements in immunization provider data quality represent a complex and important area; however, a focus on basics such as automated methods for preparing data for comparison purposes and cleaning garbage from names and addresses, detecting typos, name variations, and misspellings may produce a large pay-off.
- IIS should have formal documentation and a set of rules that everyone performing manual de-duplication should follow. It was noted that operator experience does improve manual de-duplication accuracy.
- Merging and unmerging. No record matching system is perfect, and at some point all systems will erroneously link together records that belong to separate patients. It is best to plan for the capabilities needed to swiftly unmerge erroneous merges, rather than pretending that they will not occur.
- As a best practice, it is better to have duplicates than to have mixed records, or badly merged records. The expert panel believed that it would be very helpful for all persons to have an understanding of the algorithms used for automated patient matching and de-duplication (how it works, why some matches are found and others are not, what problems to look for, etc.).

Limitations of National Practice Assessment

- Few previous NPAs have specifically examined IIS patient-level de-duplication practices.
- Public Health agencies were the key participants in this investigation. Because of time and resource constraints, it is not always possible to respond or respond with complete information.
- Further study of national IIS patient de-duplication challenges, issues, and practices is needed.

Conclusion

- A sustained investment is required to continuously improve IIS patient matching and patient-level de-duplication capabilities.
- IIS implementations need to closely monitor ONC and other governmental policy decisions regarding what levels of accuracy will be required for patient de-duplication. The establishment of national standards for accuracy could significantly impact patient de-duplication process and related operations.

- Results of the NPA are being used to inform the work of the CDC Patient-level De-duplication Expert Panel. The work of this group is scheduled to conclude in early 2013.

Address for correspondence

Jennifer Wain, Program Manager, Northrop Grumman Contractor to CDC. Jua7@CDC.gov, 011-678-595-9771.

Frederic Grant, PhD, MPH, MBA. Senior Public Health Advisor & Contractor to CDC. Chief Scientist, Public Health, Northrop Grumman Corporation, 011-770-262-7593. fred.grant@ngc.com or hlc4@CDC.gov.

* Expert Panel Coauthor Information: Michael Berry / RI; Gerry Bragg / MI; Nathan Bunker / UT; Shaun Grannis, MD / IN; Rick Hall, PhD / NJ; Steve Jarvis / CO; Brian Jorgage / PA; Mary Beth Kurilo, MPH, MSW / OR; Linda Leubchow / MN; Christie Levy / MS; Megan Meldrum / NY; Chris Pratt / UT; Helen Redfield / TX; Cecile Town / NM.

** Technical Writing/Editing: Celia Toles, Northrop Grumman Corporation.

This study conformed to CDC Guidance on Scientific Integrity.

This handout is being made available courtesy of The Centers for Disease Control and Prevention.

As a contractor to the CDC, the Northrop Grumman Public Health Division works with federal, state, local governments and private agencies to contribute to public health's evidence-based practice. Northrop Grumman's heritage of supporting public health system development spans more than 15 years.