Leveraging Statewide Immunization Registry Data Assets

Erich Daub, Scientific Technologies Corporation
Buzz Conn, Scientific Technologies Corporation
Michael L. Popovich, MSSE, Scientific Technologies Corporation
Ruben Tapia, MPH, Louisiana Immunization Program Director
Quan Le, RN, Louisiana Immunization Program Registry Manager

Scientific Technologies Corporation

January 29, 2010

JANUARY 29, 2010

Leveraging Statewide Immunization Registry Data Assets

Erich Daub, Scientific Technologies Corporation
Buzz Conn, Scientific Technologies Corporation
Michael L. Popovich, MSSE, Scientific Technologies Corporation
Ruben Tapia, MPH, Louisiana Immunization Program Director
Quan Le, RN, Louisiana Immunization Program Registry Manager

Overview

... to test the feasibility and value of integrating a decision-support toolset into a state immunization registry.

This white paper describes the value of using a business intelligence tool to access information within an immunization registry. Statewide registries that capture population-based immunization records contain significant amounts of health information specific to vaccine preventable diseases. Scientific Technologies Corporation (STC), in partnership with the State of Louisiana and Space-Time Research (STR), undertook a demonstration project to test the feasibility and value of integrating a decision-support toolset into a state immunization registry.

... creating data queries in minutes to assess risk, improve the management of resources, and monitor rapidly changing conditions...

The demonstration project showed that ad hoc reporting and easy access to all data in an immunization registry creates an information environment that offers users and decision makers data that previously had not been readily available. Furthermore, it demonstrated the possibility of creating data queries in minutes to assess risk, improve the management of resources, and monitor rapidly changing conditions such as the recent Novel H1N1 event.

Summary of State Immunization Registry Assets and the Need

... Immunization histories of Louisiana residents were rapidly made accessible to physicians and health care providers throughout the U.S.

In the early 1990's the Centers for Disease Control and Prevention (CDC) established a national vision for immunization tracking systems. The vision included a challenge to implement statewide immunization registries that would allow physician access to patient vaccine histories regardless of where the immunization had been received. Evidence of the progress toward these goals was demonstrated in the aftermath of Hurricane Katrina (2005) in which immunization histories of Louisiana residents were rapidly made accessible to physicians and health care providers throughout the U.S.[1]

As state registries have continued to develop over the past decade, the value of their data has grown. States are collecting more patient records than ever before. Electronic links to medical record systems are capturing patient data at the time of the immunization encounter. The extent of the user community with access to the data is equally expanding. Table 1 illustrates the number of patients, immunization records, and active users for six states. These state registries represent nearly sixteen million patient records.[2] Nationwide, it is estimated that over 100 million online patient and 800 million immunization records exist and that this number is growing at a rate of 10% - 20% per year.

State	Patients	Immunization Records	Registry Users
Arizona	4,000,000	39,000,000	8,200
Idaho	680,000	6,500,000	6,400
Louisiana	3,000,000	33,000,000	17,500
Maine	1,000,000	4,400,000	5,000
Washington	6,000,000	44,000,000	18,000
West Virginia	1,000,000	7,000,000	2,500

Table 1: Number of Patients, Immunization Records, and active users in six states

In addition to expanding immunization histories, some states have initiated efforts to add health parameters such as test results for newborn screening and lead, obesity, and asthma risk indicators. Novel H1N1 influenza brought further attention to immunization population-based Electronic Health Records (EHR) for retaining non-traditional vaccines, thus accelerating the growth of patient data and users.

Throughout this evolution, demand for information is increasing, yet the ability for immunization program staff to easily access and report data is becoming more difficult. Current data access technologies supporting registries most often rely on custom-developed SQL queries written within database applications. Existing standard reports integrated within registry systems are not easily adaptable to real-time demands for information.

STC recognized the value of a decision-support/business intelligence tool to "unleash the power" of registry information.

STC recognized the value a decision-support/business intelligence tool would have to health agencies. If such a tool could be easily integrated and utilized, it could "unleash the power" of registry information. The requirements of the tool were established to include:

- Must support rapid on-demand statistical and analytical reporting.
- Must allow for the creation of de-identified data sets to support detailed study and analysis.
- Must support the dissemination of information to various groups and categories of users.
- Must provide tools to support spatial analysis.
- Must be used effectively by non-programmers.

For a demonstration project, STC selected a suite of analytical tools available from Space-Time Research (STR).[3] The goal of the project was to evaluate the effectiveness of decision-support tools in both day-to-day operations and in support of an event.

In addition, a technical assessment was undertaken to determine the requirements to integrate decision-support tools, such as the STR product suite, with a statewide immunization registry that adhered to U.S. CDC standards.[4]

Demonstration Pilot – Louisiana

To fully understand the value and the integration requirements, an actual dataset was utilized. STC has worked with the Louisiana Department of Health and Hospitals, which operates the Louisiana Immunization Network for Kids Statewide (LINKS), for the past ten years. LINKS is a centralized web-based immunization registry with population data from individuals throughout the state. The registry supports daily immunization program efforts as well as capacity building for preparedness and event planning.[5]

Louisiana Immunization Network for Kids Statewide (LINKS), is a centralized web-based immunization statewide registry supporting daily immunization program efforts as well as capacity building for preparedness and event planning.

The LINKS system, like all state immunization systems, is a health data asset. LINKS was used heavily during the early stages of the H1N1 planning and throughout the course of the event to support vaccine ordering, distribution, resource tracking, and data collection. Mass immunization tools allowed the state to collect data in mass clinic settings in real- time.

As registry patient records are added, the potential for utilization of the LINKS data increases; thus, the potential value of the information also increases. Pressures to conduct real-time assessments, create on-demand reports, and monitor resources in an event increase data value to the end user if results can be provided quickly. LINKS does include established reports which provide some information, but these reports are not easily or rapidly adaptable to assist staff and state health authorities with information to support real-time decisions. Therefore:

… decisions and outcomes could be achieved more rapidly and productively by registry staff.

- If registry data could easily accept ad hoc queries and analysis by public health professionals, decisions and outcomes could be achieved more rapidly and productively by registry staff.

... a greater analytic power could be achieved leading to more informed decision making.

- If registry data stored in tables could be rapidly converted to different value sets, a greater analytic power could be achieved, leading to more informed decision making. A basic example is the ability to regroup patient ages to any user-defined categories, expanding upon those already hard-coded into the registry.

... a greater range of comparisons could be available to improve planning and to enhance the utility of registry data to immunization and other public health programs.

- If registry data could be combined with data from other sources, a greater range of comparisons could be available to improve planning and to enhance the utility of registry data to immunization and other public health programs. One example would be combining census information with registry data to rapidly project the locations of under-vaccination or areas with high concentrations of providers not participating in the registry. Additional data sources such as Medicaid, Maternal and Child Health, school system, and Vital Records, among many others, expand the potential of registry data to improve population health and control healthcare costs.

... the ability to confidentialize data in reports and control the visibility of small numbers would be an essential feature of an analysis and reporting system.

- If registry data could be published or shared in a manner that conforms to applicable standards of confidentiality, the ability to confidentialize data in reports and control the visibility of small numbers would be an essential feature of an analysis and reporting system.

Demonstration Pilot – Sample Scenarios

Four test scenarios were generated to evaluate the effectiveness of real time decision-support tools. Initially the test data set was created by abstracting a subset of the LINKS data, de-identifying the records, and scrambling the immunization information.[6] Installation of the STR decision-support products[7] to allow access to the test dataset was simplified by providing direct access to the test data, though the optimal approach would be to integrate the tools within the registry application.

A summary of the four scenarios, a few brief examples of the user screens to set up the ad hoc reporting capabilities, and highlights of the output are illustrated on the following pages.

Scenario 1: Average Patient Age Parish Map by Vaccine Type

Under-immunized patients are identified, strategies developed, and programs implemented to increase coverage.

Age appropriate immunization coverage requires summarizing specific vaccine and vaccine groups by patient age and comparing this to recommended immunization schedules. Under-immunized patients are then identified, strategies developed, and programs implemented to increase coverage. Most often this type of analysis is targeted to specific state regions, sub-populations, providers, clinics (i.e., community and rural health centers), or enrolled member bases of health plans.

… when determining age appropriate risk, the results are geographically provided in maps to quickly assess the high-risk areas.

This scenario required the use of the decision-support tool to query the sample LINKS dataset for DtaP/DT/Td/Tdap vaccines and to determine the age appropriate risk. Presenting the results geographically provided maps for senior staff to quickly assess the high-risk areas. Figures 1 through 3 illustrate the three-steps in creating a query to search the database for the desired information. Figure 4 illustrates the resulting output – appropriate coverage by Parish for the specified vaccine series.[8]

Initially, the user opens the SuperVIEW launch page and selects the Report from the list on the left hand panel, or clicks on one of the thumbnails.

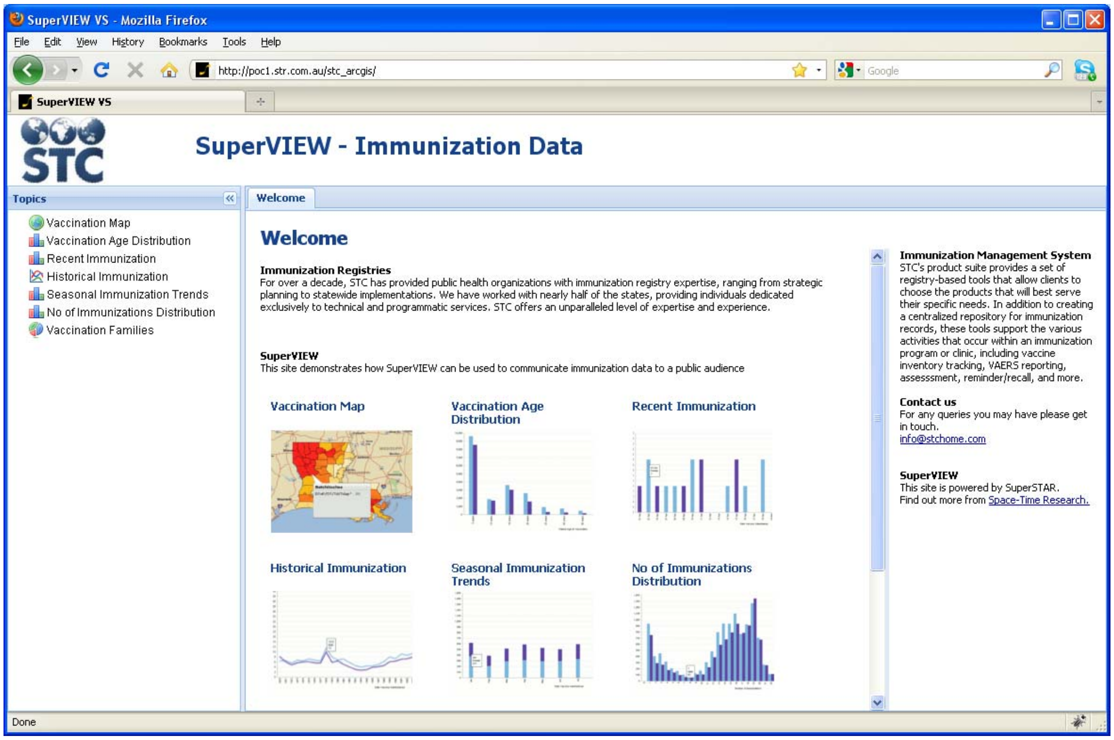


Figure 1: Step 1, Launching SuperVIEW and Selecting a Report Type

The report opens with a default view, showing “Average Age at Vaccination” across all time periods, and all vaccines (Figure 2).

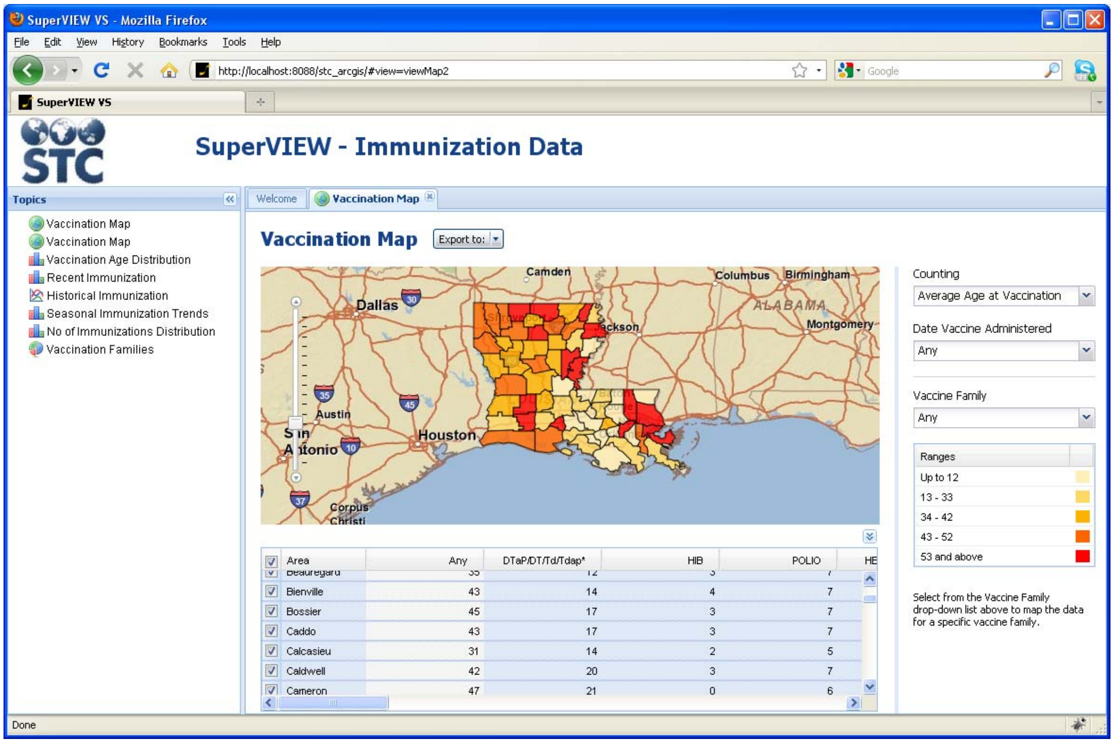


Figure 2: Step 2, Default Report View

Next, the user selects the time period of interest from the drop-down lists on the right-hand side, and which vaccine family to filter on (Figure 3).

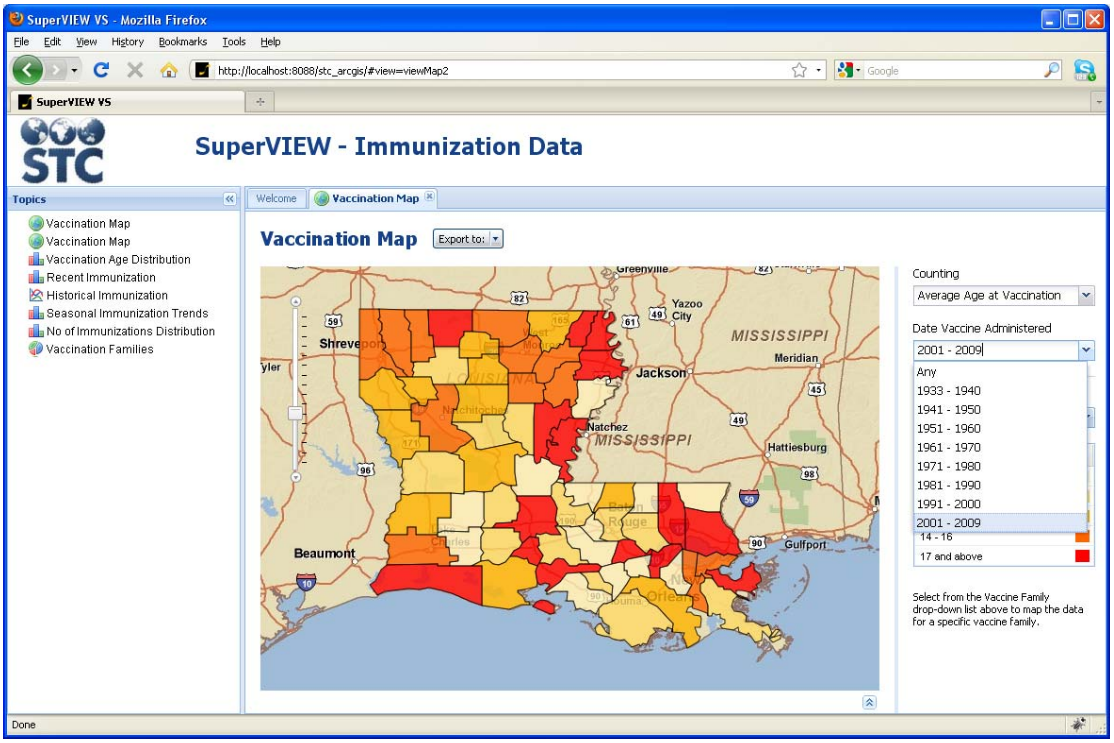


Figure 3: Step 3, Specifying Time Period and Vaccine Family

A resulting updated map is displayed (Figure 4) which can be output in a number of formats. The map can display with or without its companion data table.

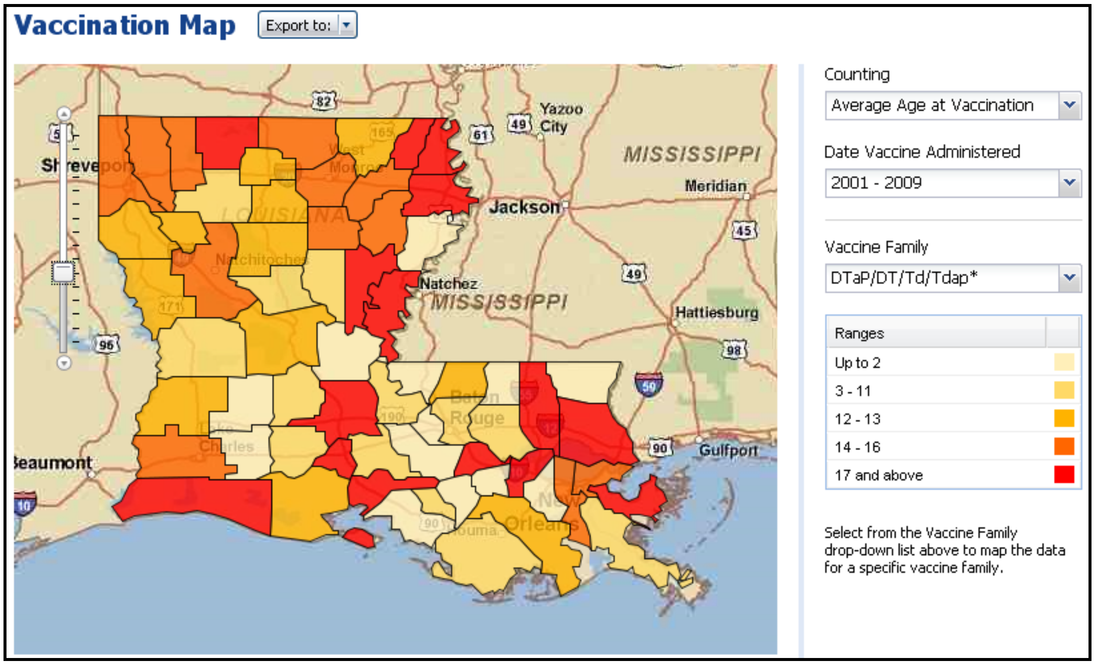


Figure 4: Result of Query to Determine Age Appropriate Coverage for DTaP/DT/TD/Tdap by Parish

Scenario 2: Performing Statistical Analysis with Limited Data

Ensure patient confidentiality with STR decision-support tools.

The second scenario was an ad hoc query for a statistical analysis, with the twist that limited data in some cells might result in patient confidentiality issues if included in the report. To ensure patient confidentiality the STR product was used to create a set of rules that were applied when limited data existed. The rules that can be implemented with the STR decision-support tools[9] are robust and include the ability to use rounding, suppression, and secondary suppression of data to protect the identity of individuals.

Figure 5 illustrates the user's initial ad hoc request creating the data rules.

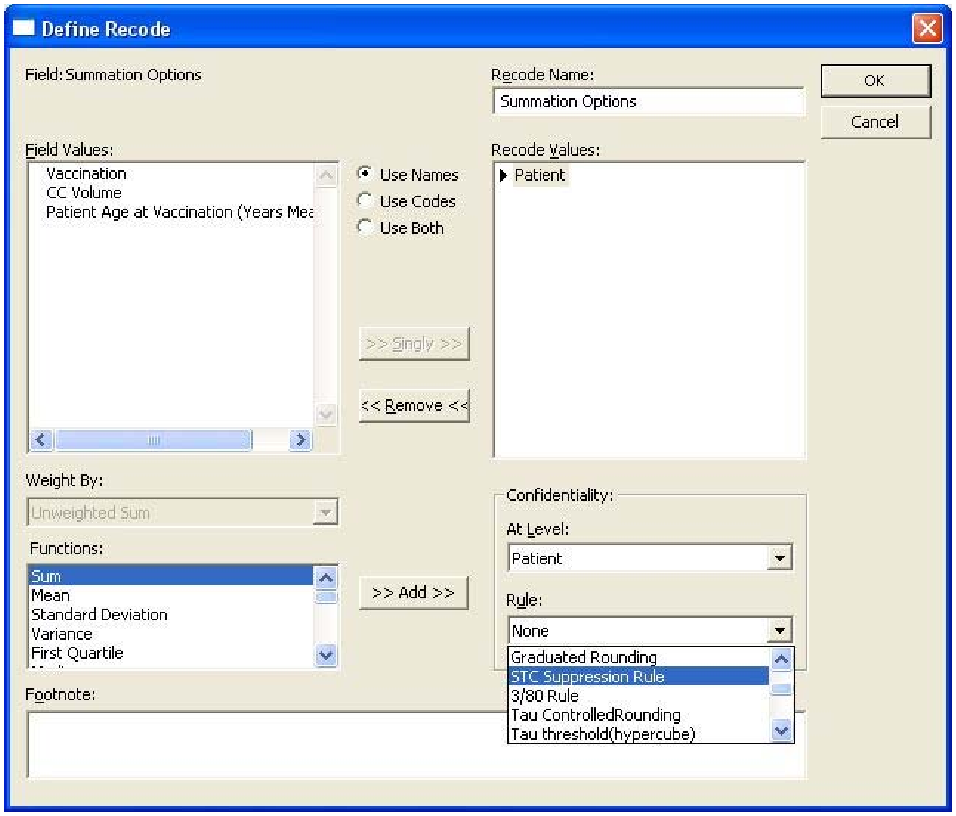


Figure 5: Selecting the Confidentiality Level and Rule

Table 2 illustrates the before and after tables revealing the results of the data query before application of the “protection rules” and after. The cells displaying “..C” denote that the patient data numbers are small enough, per the established threshold rules, that if they were included in this report confidentiality could be compromised.

Raw results

Ethnicity and Vaccine by Birth Year and Gender
Caddo Parish
Number of Patients

	1994		1995	
	Female	Male	Female	Male
Hispanic or Latino				
DTaP/DT/Td/Tdap*	15	11	9	17
HEP-B 3 DOSE**	15	11	9	18
VARICELLA	8	5	4	12
HEP-A	3	1	3	3
FLU	2	-	1	2
Not Hispanic or Latino				
DTaP/DT/Td/Tdap*	576	613	549	559
HEP-B 3 DOSE**	574	601	539	555
VARICELLA	294	332	328	349
HEP-A	28	37	25	31
FLU	22	26	23	28

Results after privacy protection applied

Ethnicity and Vaccine by Birth Year and Gender
Caddo Parish
Number of Patients

	1994		1995	
	Female	Male	Female	Male
Hispanic or Latino				
DTaP/DT/Td/Tdap*	15	11	9	17
HEP-B 3 DOSE**	15	11	9	18
VARICELLA	8	5	4	12
HEP-A	..C	..C	..C	..C
FLU	..C	-	..C	..C
Not Hispanic or Latino				
DTaP/DT/Td/Tdap*	576	613	549	559
HEP-B 3 DOSE**	574	601	539	555
VARICELLA	294	332	328	349
HEP-A	28	37	25	31
FLU	22	26	23	28

Table 2: Illustration of Results from a Standardized Database Query and an Enhanced Query using Data Confident Tools

Scenario 3: Vaccination Trends

... target and monitor the effectiveness of specific programs... are important applications of the data.

Using immunization data to minimize the risk of disease on individuals and to target and monitor the effectiveness of specific programs, such as the uptake of HPV vaccine to high-risk young women, are important applications of the data. The scenario established to test ad hoc reporting called for determining who had received the HPV vaccine. Figures 6 through 8 illustrate the real-time query that was defined to access the registry data and the resulting bar chart.

The user begins by selecting pre-saved reports, or by selecting a database with which to create a new report. In this scenario the LINKS database was selected. The user is presented with a screen showing the list of available reporting fields on the left-hand side (Figure 6).

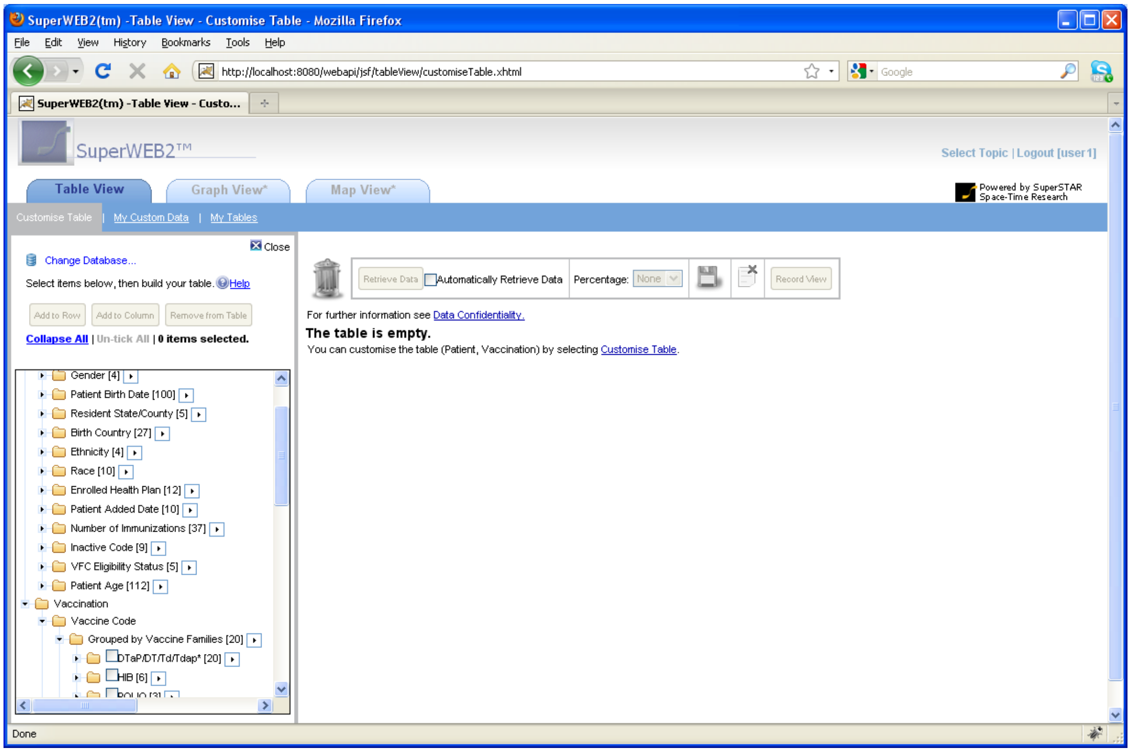


Figure 6: Step 1, Selecting the Desired Database

The user then selects the desired fields and values and chooses where to position them in the table.

In this scenario, a subset of values from the field “Patient Age at Vaccination” was selected and positioned in the Row dimension of the table.

To further filter the results, the user also selects a specific vaccine type – HPV in this scenario – and adds that to the Column dimension of the table (Figure 7).

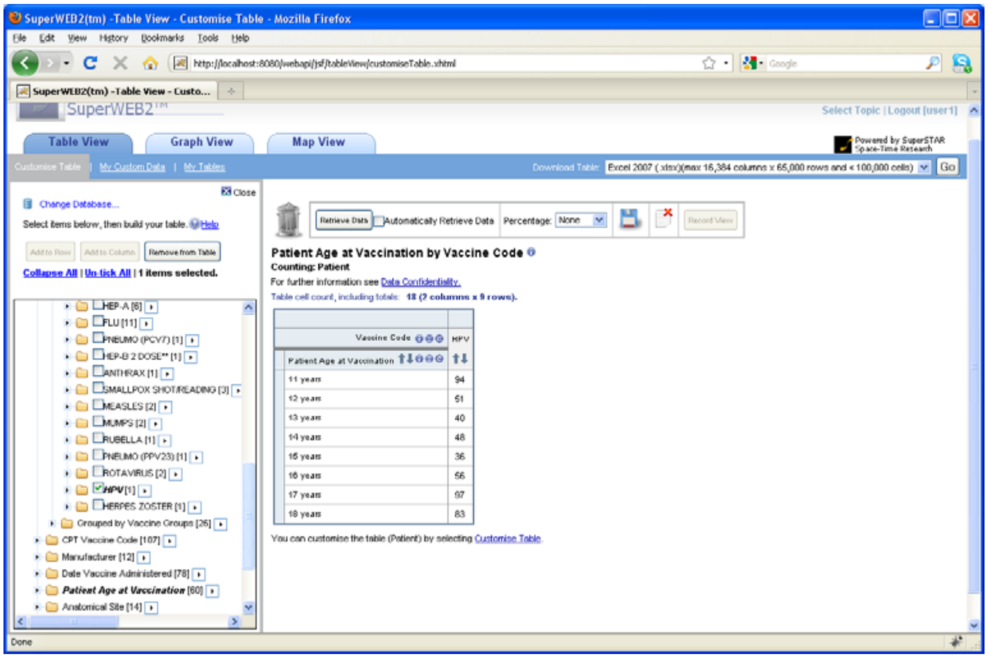


Figure 7: Step 2, Selecting the Desired Fields and Vaccine of Interest

Finally, to represent the results in a chart or graph, the user selects the desired view and the visualization is displayed (Figure 8).

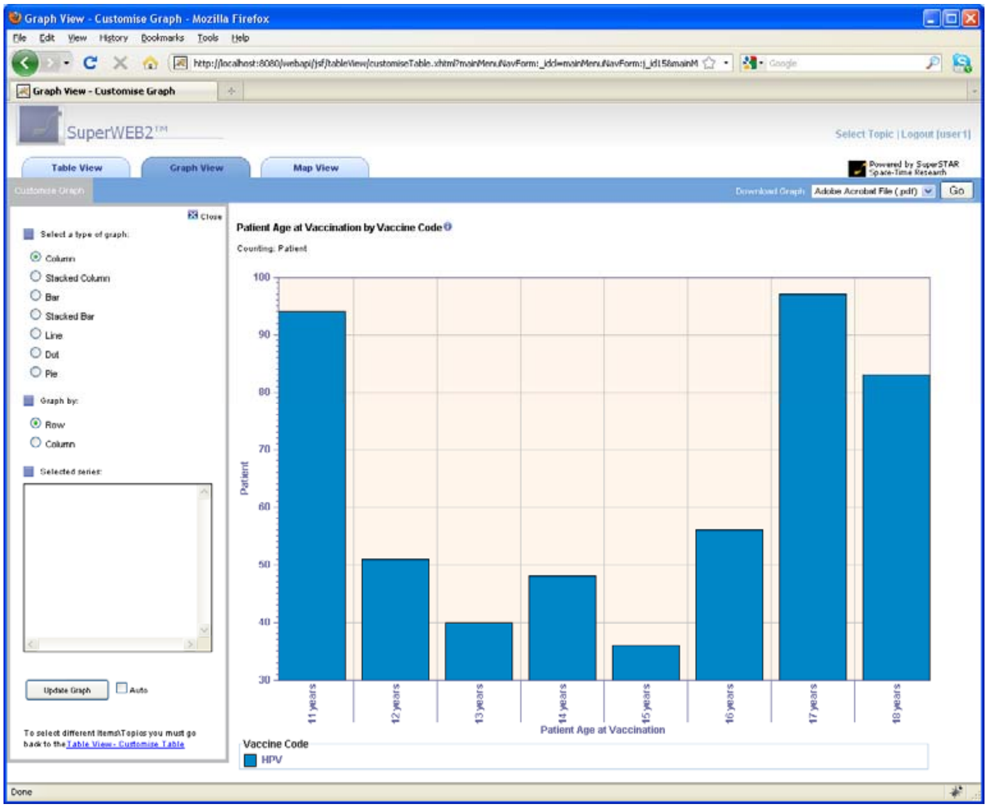


Figure 8: Example Representation of the Number of High-risk Young Women who have received the HPV Vaccine

Scenario 4: Data Quality

... the value of the data is only as good as its accuracy and completeness... test the feasibility of using decision-support tools to monitor data quality.

As data is collected, the value of the data is only as good as its accuracy and completeness. A scenario was developed to test the feasibility of using decision-support tools to monitor data quality. In this scenario, a cross-tabulation was created to verify whether patient vaccination dates preceded their date of birth. The initial step involved the user adding the desired fields and values to a table. In this scenario, a selection of patient birth dates and vaccination dates were included (Figure 9).

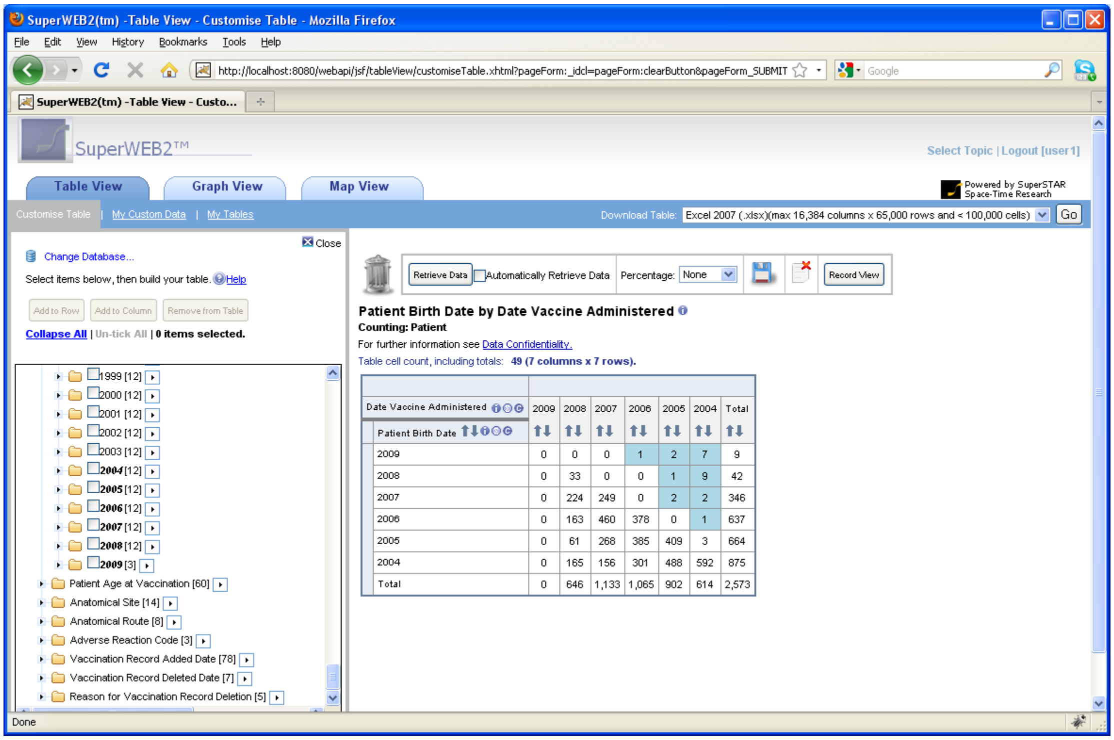


Figure 9: Step 1, Selecting Desired Fields and Values for the Data Table

The user next selects the fields to view from each of the tables. Fields selected in this scenario included "Patient ID," "Gender," "Birth Date," "Vaccine Family," and "Vaccination Date" (Figure 10).

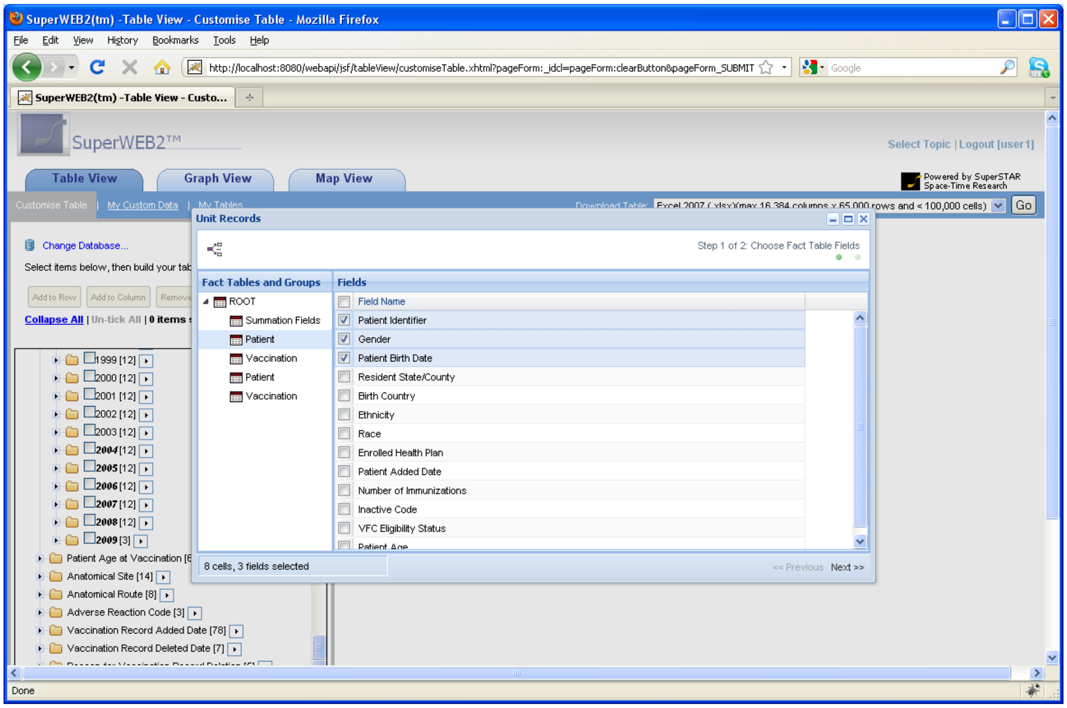


Figure 10: Step 2, Selecting Desired Fields to View

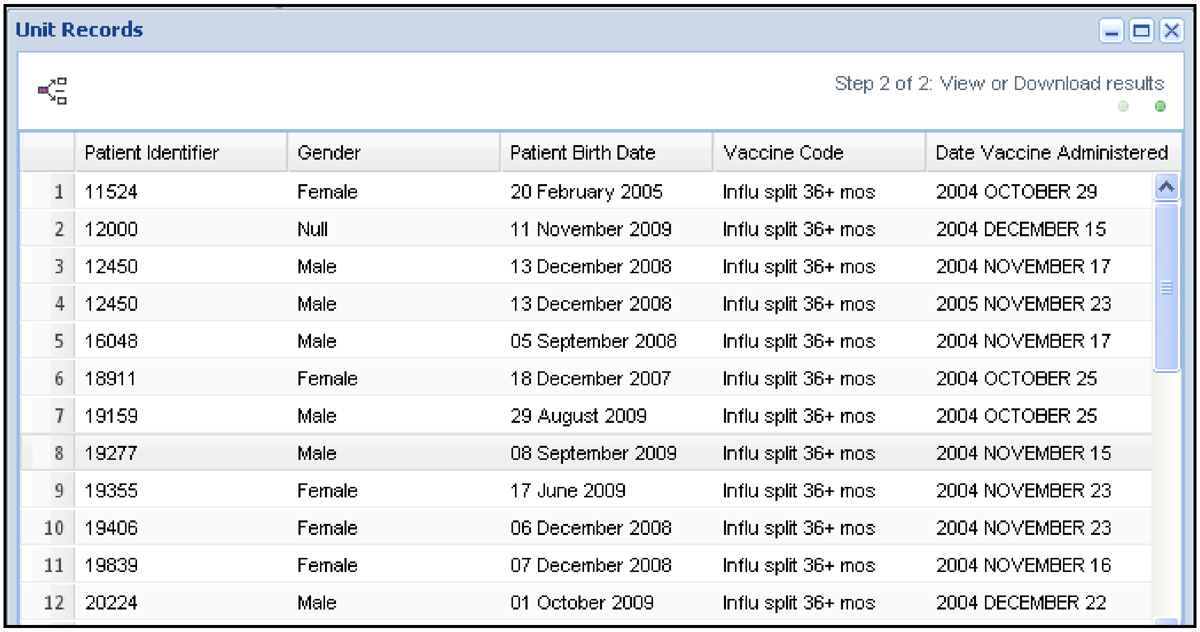
Figure 11 illustrates the output of this cross tabulation, summarizing patients born in specific years, and the date of their most recent immunization. From this output the application color-codes the cells in blue to highlight likely data problems.

Date Vaccine Administered	2009	2008	2007	2006	2005	2004	Total
Patient Birth Date							
2009	0	0	0	1	2	7	9
2008	0	33	0	0	1	9	42
2007	0	224	249	0	2	2	346
2006	0	163	460	378	0	1	637
2005	0	61	268	385	409	3	664
2003	0	146	647	216	348	664	1,136
2004	0	165	156	301	488	592	875
Total	0	792	1,780	1,281	1,250	1,278	3,709

Figure 11: Resulting Output of Data Quality Cross-Tabulation of Date of Birth and Vaccine Date

... easily follow-up and correct information.

The STR decision-support toolset enables a user to interact with the data and request a line listing of the patients that were likely to have data issues. This is illustrated in Figure 12 which lists example records with administration dates prior to the date of birth. This should allow data managers to easily follow-up and correct the information.



Unit Records

Step 2 of 2: View or Download results

	Patient Identifier	Gender	Patient Birth Date	Vaccine Code	Date Vaccine Administered
1	11524	Female	20 February 2005	Influ split 36+ mos	2004 OCTOBER 29
2	12000	Null	11 November 2009	Influ split 36+ mos	2004 DECEMBER 15
3	12450	Male	13 December 2008	Influ split 36+ mos	2004 NOVEMBER 17
4	12450	Male	13 December 2008	Influ split 36+ mos	2005 NOVEMBER 23
5	16048	Male	05 September 2008	Influ split 36+ mos	2004 NOVEMBER 17
6	18911	Female	18 December 2007	Influ split 36+ mos	2004 OCTOBER 25
7	19159	Male	29 August 2009	Influ split 36+ mos	2004 OCTOBER 25
8	19277	Male	08 September 2009	Influ split 36+ mos	2004 NOVEMBER 15
9	19355	Female	17 June 2009	Influ split 36+ mos	2004 NOVEMBER 23
10	19406	Female	06 December 2008	Influ split 36+ mos	2004 NOVEMBER 23
11	19839	Female	07 December 2008	Influ split 36+ mos	2004 NOVEMBER 16
12	20224	Male	01 October 2009	Influ split 36+ mos	2004 DECEMBER 22

Figure 12: Detailed Record Level View of the Data Quality Issues

Results

... power and ease of use of a decision-support tool.

... rapid ad hoc analysis... completed in less than ten minutes.

... data manipulations easily performed... with minimum training.

The value of a tool to access state immunization records to produce ad hoc statistical analysis and reporting was demonstrated in the above scenarios.

The user experience to create the supporting data for the defined scenarios illustrated the power and ease of use of a decision-support tool.

Rapid ad hoc analysis for each of the four scenarios was completed in less than ten minutes.

Data manipulations were easily performed by non-technical individuals with a minimum of training.

... evaluate information anomalies, improve data quality, dynamically control data output while preserving patient confidentiality; increases the value of data in the registry.

These specific STR toolsets included unique decision-support components that afforded user capacities to evaluate information anomalies, to improve data quality, and to dynamically control data output to preserve patient confidentiality. The web-based implementation supported widespread utilization of these analytic, visualization and reporting capabilities, thus increasing the value of the data captured in the registry.

Integrating a decision-support tool within a state immunization registry is recommended. A variety of products exists in this space and should be evaluated based on specific state information technology standards and requirements. As described here, the STR analytical tools proved to be extremely versatile and powerful with a set of enhanced capabilities not found in other systems.

Summary

... rapidly provide decision makers with the knowledge needed to address public health questions.

Historically, registry data has been underutilized. Significant opportunities exist for improving population health by leveraging existing and emerging registry data. The approach and tools described herein made improvements by freeing registry program staff from canned reports and complex “language-based” analytic software restricted to highly trained users. Beyond their immediate benefit to routine programmatic management, the greatest benefit may be the potential of the tools to rapidly provide decision makers with the knowledge needed to address public health questions for which the circumstances surrounding them have yet to be encountered.

... quickly formulate decision-support questions and easily extract results from data.

Patient records in state immunization registries will continue to grow. Satisfying the existing registry operational requirements demands tools to quickly formulate decision-support questions and easily extract results from data. Immunization programs will become more efficient and effective if data can be turned rapidly into information which creates knowledge to determine:

- What populations have been effectively immunized and the characteristics of those at risk?

- Where are available supplies of vaccine and utilization characteristics?
- How effective are providers at interfacing with the registry, how can their data quality be improved, and where best to recruit additional providers?
- How can coverage rates be improved?
- What data within the registry could be better leveraged to improve health outcomes?
- What additional sources of data could be integrated to improve health outcomes?

Investments made now will prove cost effective and create value added benefits...

Over the course of the next two years all registry managers should be examining how to integrate such tools within their current application environments. Investments made now will prove cost effective and create value added benefits by unlocking the information within patient population data to support the continuing campaign to minimize the impact of vaccine-preventable disease on individuals, families, and economies.

[1] Urquhart GA, Williams W, Tobias J, Welch FJ. Immunization Information Systems Use During a Public Health Emergency in the United States. J Public Health Management and Practice. 2007 Sep-Oct;13(5):481-5.

[2] Statistics (rounded) from six STC IWeb and WIR US-based Immunization Registry Partners, December 2009.

[3] Space-Time Research Pty Ltd Level 1, 630 Church Street, Richmond VIC 3123, AUSTRALIA info@spacetimeresearch.com

[4] CDC, NCIRD, "Recommended Core Data Items for Immunization Information Systems (IIS)," Updated May 2008, http://www.cdc.gov/vaccines/programs/IIS/stds/downloads/coredata.pdf

[5] Popovich M, Zhang X, 2009. Chapter 4, Case Exercise: Louisiana Mass Immunization Exercise (2007) and Case Exercise: Empowering Patients, Immunizations and Knowledge-based Personal Health Records. In Wickramasinghe N, Bali RK, Lehaney B, Schaffer JL, Gibbons MC (editors), Healthcare Knowledge Management Primer. Routledge, New York and London.

[6] The STC SIIS Data scrambler was utilized to create a de-identified test dataset reflecting actual patient records.

[7] The STR SuperSTAR toolset was the decision-support system integrated with the test data set.

[8] The STR SuperVIEW product was used to tabulate and present the results using ESRI ArcGIS mapping capabilities.

[9] The STR SuperCROSS component was used to establish the thresholds and present the results.