GroupHealth

Evaluation of existing methods for safety signal identification for the Sentinel Initiative

Jennifer C. Nelson
Biostatistics Unit
Group Health Research Institute

Department of Biostatistics
University of Washington

Acknowledgements

Collaborators

Andrea J Cook, PhD	Group Health Research Institute
Onchee Yu, MS	Departments of Biostatistics,
Lisa A Jackson, MD, MPH	Epidemiology & Medicine
Bruce M Psaty, MD, PhD	University of Washington

Consultants

Martin Kulldorff, PhD	Harvard Medical School, Harvard Pilgrim Health Care
Thomas Fleming, PhD	Departments of Biostatistics and Statistics, University of Washington
William DuMouchel, PhD	Lincoln Technologies Division, Phase Forward

Funding: A Sentinel Initiative planning contract funded by the Food and Drug Administration (HHSF223200S31324P) supported this work.

Outline

GroupHealth

- **Review common signal detection methods**
- **Weigh pros and cons for use in Sentinel**
  - *Large, broad population (*vs pre-licensure trials)
  - *Proactive* (vs spontaneous reports)
  - *Rapidly* available (vs conventional Phase 4)
  - *Longitudinal* versus single-point-in-time analysis
- **Discuss implementation**
- **Identify methodological gaps**

Data Mining

GroupHealth

- **Extracting hidden patterns in large datasets**
- **Simultaneously examine 'all' adverse event & drug pairs for associations**
- **Basic idea**
  - Compute measure of 'disproportionality'
  - Rank all pairs
  - Large values → signal

	Drug of interest	All other drugs	
AE of interest	**a**	**b**	**a+b**
All other AEs	**c**	**d**	**c+d**
	a+c	**b+d**	**N**

Data Mining

GroupHealth

- **Key features**
  - Hypothesis-free
  - Retrospective
  - Analysis at a single point in time
- **Applications**
  - Passive spontaneous report databases
    - WHO adverse drug reactions (ADR) database (Uppsala)
    - FDA AERS and VAERS databases in the U.S.
    - Pharmacovigilance activities by industry
  - Feasibility studies in health care claims data

CDC Home | Search | Health Topics A-Z

MMWR Weekly

September 3, 2004 / 53(34);786-789

Suspension of Rotavirus Vaccine After Reports of Intussusception ---United States, 1999

Please note: *An erratum has been published for this article. To view the erratum, please click here.*

On July 16, 1999, CDC recommended that health-care providers suspend use of the licensed rhesus-human rotavirus reassortant-tetravalent vaccine (RRV-TV) (RotaShield®, Wyeth Laboratories, Inc., Marietta, Pennsylvania) in response to 15 cases of intussusception (i.e., a bowel obstruction in which one segment of bowel becomes enfolded within another segment) among infants who received RRV-TV (*1*). The Vaccine Adverse Event Reporting System (VAERS) monitored for adverse events following licensure of RRV-TV on August 31, 1999. After the recommendation to suspend use of the vaccine, no additional cases were reported (*2*). This report describes the surveillance activities used to identify this vaccine adverse event, the emergency response, and follow-up investigations. Suspension of RRV-TV after the initial cases of intussusception parallels the removal of the Broad Street pump handle in response to John Snow's epidemiologic studies; both were decisive, life-saving public health actions.

VAERS, operated by CDC and the Food and Drug Administration (FDA), is a national passive surveillance system that monitors the safety of vaccines (*3*). Health-care providers, consumers, and vaccine manufacturers are encouraged to report adverse events involving all U.S.-licensed vaccines. During 1998--1999, CDC and FDA monitored VAERS for reports of intussusception and other severe gastrointestinal events among RRV-TV recipients. As a requirement for FDA licensure, the vaccine manufacturer funded a postlicensure phase IV trial of RRV-TV at Northern California Kaiser Permanente (NCKP) to monitor possible adverse events (*4*). Intussusception had been observed at low rates in prelicensure clinical trials, but a causal association with the vaccine was not proven; this information and a request for reporting to VAERS

Sequential testing

GroupHealth

- **Pre-specify target AE-drug pairs (hypotheses)**
  - Null hypothesis: No difference in MI risk
  - Alternative: Increased MI risk among Vioxx recipients
- **Analyze these targets as information accrues**
- **Basic Idea**
  - Each week, count up AEs among drug recipients
  - Count up AEs among comparators
  - Compute statistic to compare risk between groups
  - If difference is too big: STOP, safety problem
  - If no signal at study end: STOP, no problem detected
  - ‘Too big’ chosen to minimize false positive errors

Monthly testing example

Sequential testing

GroupHealth

- **Key features**
  - Formal hypothesis-testing
  - Prospective (rapidly as data accrue)
  - Repeated tests at multiple points in time
- **Applications**
  - Efficacy monitoring in pre-licensure randomized trials
  - Preliminary drug surveillance within HMORN CERT (Center for Education & Research on Therapeutics)
  - Routine (weekly) post-licensure vaccine safety surveillance within CDC's Vaccine Safety Datalink

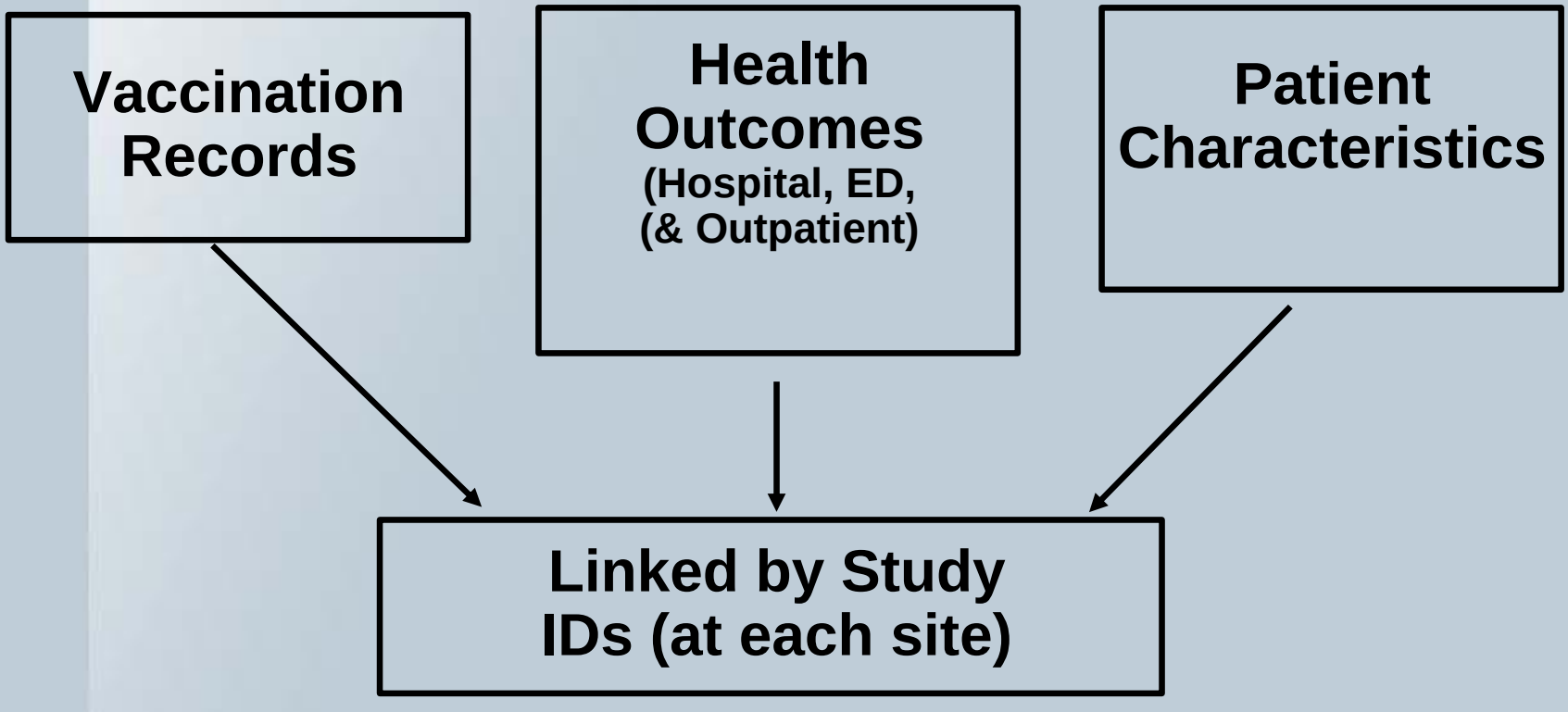
CDC-sponsored Vaccine Safety Datalink project

GroupHealth

- Large: 8.8 million (3% of U.S.)
- Population-based: enrollees from 8 health plans

Vaccine Safety Datalink data systems

GroupHealth



- Standardized electronic health care utilization data is updated weekly
- Has allowed for routine longitudinal monitoring of new vaccines since ~2005 (e.g. Menactra, Rotateq, HPV, Tdap, MMRV, & flu vaccines)

CDC Home | Search | Health Topics A-Z

MMWR Weekly

March 14, 2008 / 57(10);258-260

Update: Recommendations from the Advisory Committee on Immunization Practices (ACIP) Regarding Administration of Combination MMRV Vaccine

On February 27, 2008, new information was presented to the Advisory Committee on Immunization Practices (ACIP) regarding the risk for febrile seizures among children aged 12--23 months after administration of the combination measles, mumps, rubella, and varicella (MMRV) vaccine (ProQuad®, Merck & Co., Inc., Whitehouse Station, New Jersey). This report summarizes current knowledge regarding the risk for febrile seizures after MMRV vaccination and presents updated ACIP recommendations that were issued after presentation of the new information. These updated recommendations remove ACIP's previous preference for administering combination MMRV vaccine over separate injections of equivalent component vaccines (i.e., measles, mumps, and rubella [MMR] vaccine and varicella vaccine).

The combination tetravalent MMRV vaccine was licensed by the Food and Drug Administration (FDA) on September 6, 2005, for use in children aged 12 months--12 years (*1*). MMRV vaccine can be used in place of trivalent MMR vaccine and monovalent varicella vaccine to implement the recommended 2-dose vaccine policies for prevention of measles, mumps, rubella, and varicella (*1*, *2*). The first vaccine dose is recommended at age 12--15 months and the second at age 4--6 years.

In MMRV vaccine prelicensure studies, an increased rate of fever was observed 5--12 and 0--42 days after the first vaccine dose, compared with administration of MMR vaccine and varicella vaccine at the same visit (*3,4*). Because of the known association between fever and febrile seizures (*5*), CDC and Merck initiated postlicensure studies to better understand the risk for febrile seizures that might be associated with MMRV vaccination.

The Vaccine Safety Datalink (VSD),* which routinely monitors vaccine safety by near real-time surveillance using computerized patient data, detected a signal of increased risk for seizures of any etiology among children aged 12--23 months after administration of MMRV vaccine compared with administration of MMR vaccine (many children also received varicella vaccine). When children who received MMRV vaccine were

Pros (+) and cons (-): Data Mining

GroupHealth

+ **Useful way to glean info from spontaneous reports**

- ***Ad hoc* control for false positive errors**

- **Associations for a given AE:drug can be masked**
  - Due to magnitude of multiple testing
  - Other drugs being mined are indicated for that AE

- **Designed for retrospective, single analysis**
  - Not immediately clear if/how to apply prospectively

Pros (+) and cons (-): Sequential Testing

GroupHealth

+ **An established methodology for monitoring efficacy in RCTs**
+ **A special case (weekly testing with a flat boundary) has proven successful for proactive surveillance in VSD**
+ **Offers flexible monitoring plan options**
  - Can vary boundary, frequency of testing, comparator group
  - Tailor choices to desired trade-off between power & timeliness
- **Little info to guide design choices outside efficacy setting**
- **Optimal design for post-licensure safety setting not known**
  - How frequently should testing be performed?
  - What stopping boundary shape is desirable?
  - What statistic best quantifies risk difference?
  - (How) should answers above vary by outcome?

Pros (+) and cons (-): VSD approach
(involves near-continuous testing and a flat boundary)

GroupHealth

+ **Highly frequent testing yields shorter average time-to-detection**
- **Highly frequent testing**
  - May not be feasible (infrastructure, weekly analysis & review)
  - May sacrifice data quality (late-arriving claims, enrollment stable?)
  - Is less powerful than less frequent testing
  - May not be necessary (for rare AE's not found pre-licensure)
+ **Flat boundary (vs. higher early) is powerful at early time points**
- **Flat boundary yields more false positives early when there is greater uncertainty in data**
- **Flat boundary is less powerful overall**

Proposed Implementation strategy

GroupHealth

1. **Perform conflict-free review of all phases of available pre-licensure data**
2. **Develop an approach to identify and prioritize AE:product pairs for proactive post-licensure surveillance (versus some other study design)**
   - What size risk do you want to rule out?
   - Are the data of adequate quality to support surveillance?
3. **Formally and sequentially test these target hypotheses on routine basis**
   - Allows focus on most important, reliably-measured pairs
   - More powerful/efficient than ‘all-by-all’ data mining
   - Explicitly controls for false positive errors

Proposed Implementation strategy

GroupHealth

4. **Supplement routine hypothesis testing with**
   - Data quality checking
   - Risk estimation (to quantify 'how big' increased risks are)
   - Signal confirmation/refutation follow-up analyses
5. **To ensure comprehensive surveillance, apply complementary approaches to avoid missing AE's not anticipated in advance**
   - Data mining
   - Sequential testing for broad, non-specific outcomes
6. **Collect and analyze complementary data sources**
   - Spontaneous reports (AERS, VAERS)
   - Existing large longitudinal cohort studies
   - Meta-analyses of multiple post-licensure studies

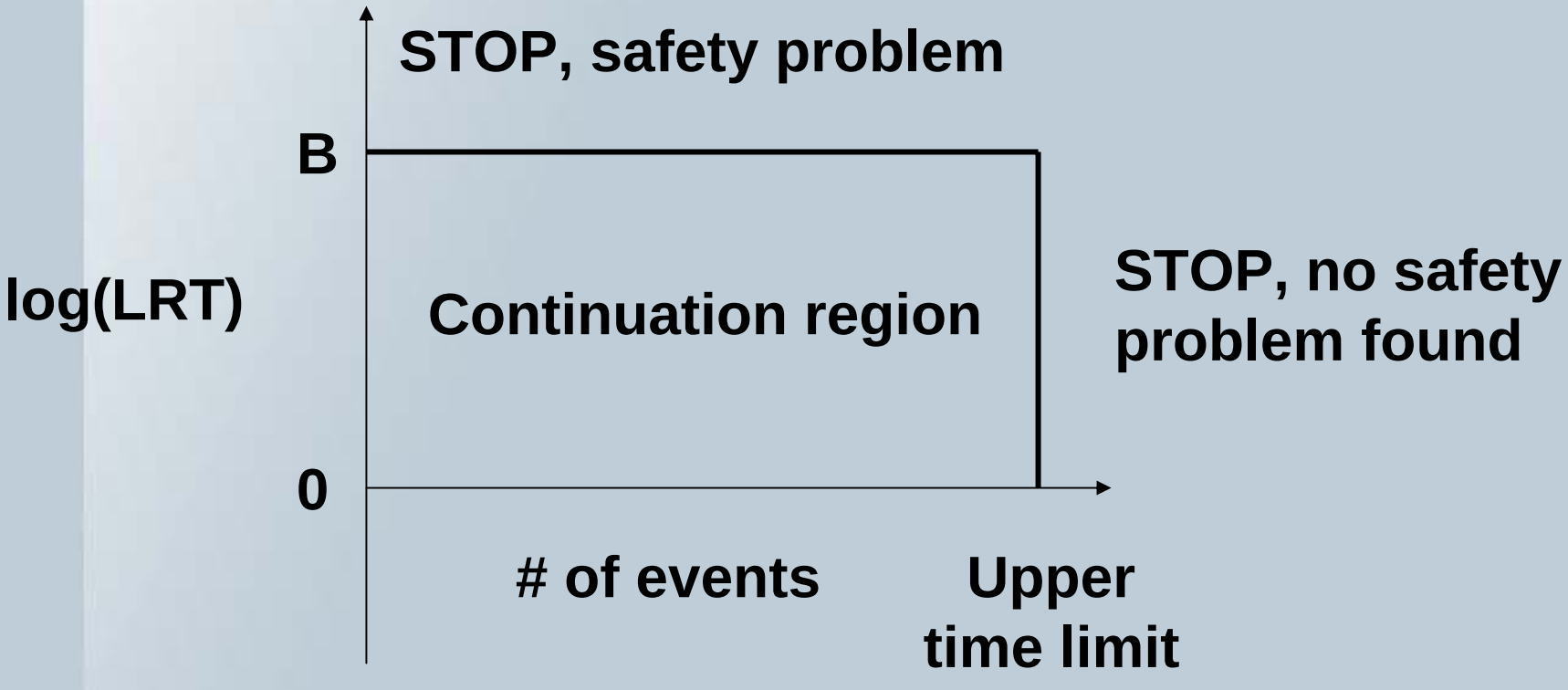
Methodological gaps and future work

GroupHealth

- **Methods to improve data accuracy (ICD-9 code groupings, NLP)**
- **Sequential testing geared for safety outcomes**
  - Better evaluate the performance of sequential design options geared for safety
  - Develop approaches designed for rare events
- **Sequential testing in observational settings**
  - Better account for confounding (match, stratify, adjust, self-control)
  - Better handle complex exposures (time-varying, varying/cumulative dosages)
  - Accommodate missing or late-arriving data
  - Accommodate more complex outcomes (delayed onset AE; not binary)
- **'All-by-all' data mining methods to discover unanticipated events**
  - Modify existing disproportionality approaches for claims data settings
  - Explicitly control for false positives
  - Better account for confounding and drug interactions
- **Better framework to evaluate the performance of signal detection methods**
- **Methods to assess signal robustness (within subgroups, across data sources)**

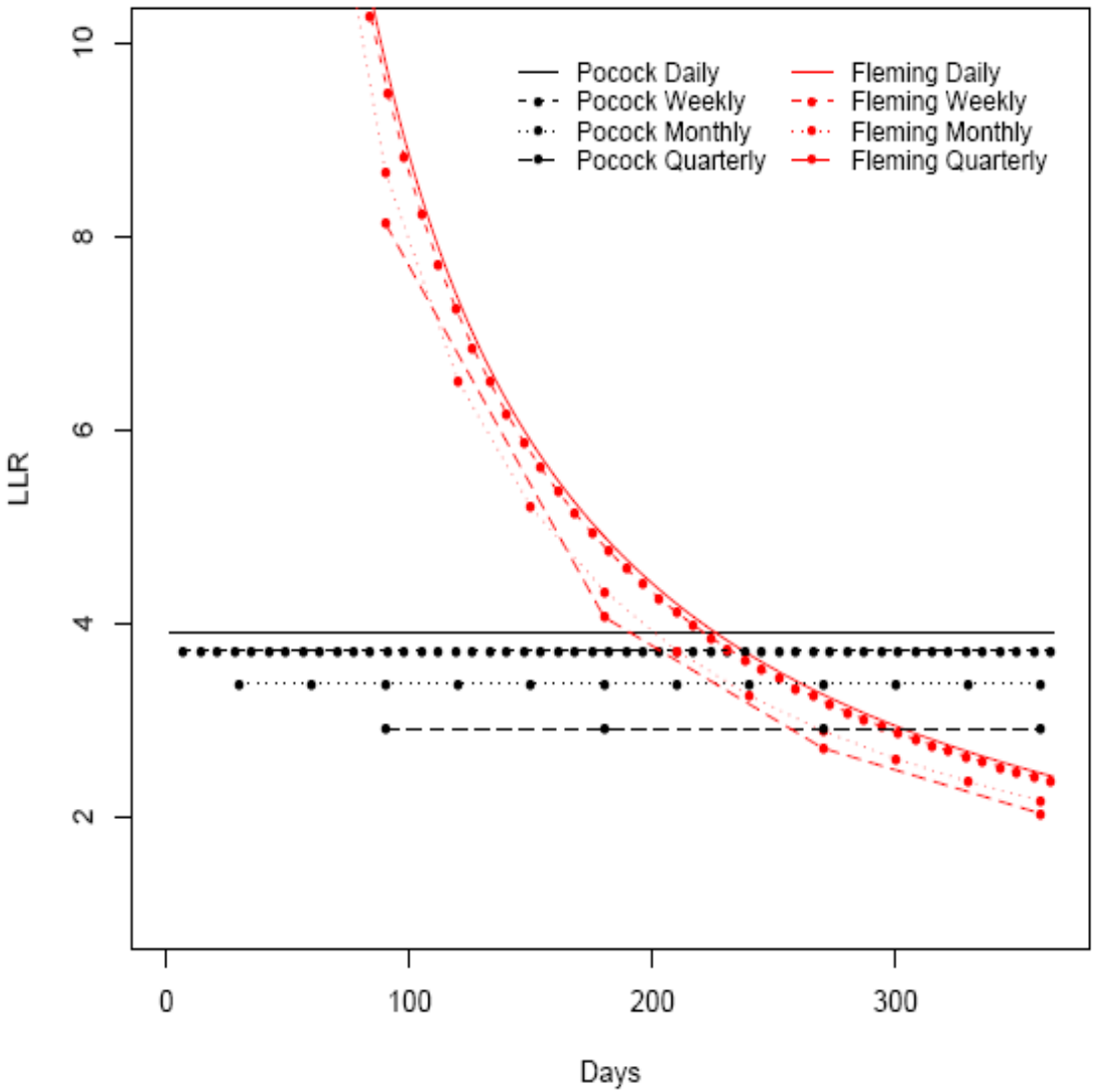
GroupHealth

Example


STOP, safety problem
B
log(LRT)
Continuation region
STOP, no safety problem found
0
# of events
Upper time limit

More stopping boundary examples

GroupHealth



Vaccine Safety Datalink

- Frequency: weekly
- Boundary: flat
- Test statistic: LRT

Typical efficacy trial

- Frequency: quarterly
- Boundary: decreasing
- Test statistic: varies (LRT, RR, risk difference)

Sequential testing – Common methods

GroupHealth

- **Sequential probability ratio test (SPRT):** Wald 1945
- **Cumulative sum chart (CUSUM):** Page 1954
- **Idea introduced for RCTs:** Armitage 1958
- **Stopping boundaries**
  - Pocock 1977, O'Brien & Fleming 1979
- **Alpha-spending functions**
  - Lan & DeMets 1983, Pampallona 1995
- **Bayesian designs:** Fayers 1997, Berry 1993 & 2004
- **Generalized sequential likelihood ratio tests**
  - Lai 1991 & 2004, Kulldorff 2007 (maximized SPRT)

Methods not recommended at this time

GroupHealth

- **Proportional reporting ratio (PRR):**
  - Unstable when sample sizes are small (yielding false signals)
- **Sequential probability ratio test (SPRT)**
  - Tests a simple alternative hypothesis => inaccurate
  - May delay or fail to signal if alternative is incorrect
    - If true *RR=2* and one selects an alternative *RR=10*, then truth (*RR=2*) is closer to the null (*RR=1*) than alternative (*RR=10*); SPRT may not signal
    - If true *RR=10* and one selects an alternative *RR=2,* then the null (*RR=1*) and alternative (*RR=2*) are similar; SPRT takes more data/time to signal
- **Cumulative sum chart (CUSUM) methods**
  - Designed to detect change in risk over time within the same population
  - Sentinel Initiative: detect a difference (not expected to change over time) in risk between a group exposed versus unexposed to a new product
  - Exception: if one wants to detect an elevated safety risk at certain points in time (e.g. due to a problem with a specific drug or vaccine lot)

ENGELBERG CENTER for
Health Care Reform
at BROOKINGS

Sentinel Initiative Public Workshop

Monday, January 11, 2010