

Case Study #3: Sentinel Initiative October 19, 2016

Lesley Curtis, PhDDuke University







Sentinel Initiative

- Response to 2007 FDA Amendments Act mandate to create an active surveillance system
- Mini-Sentinel pilot
- Continuous access to electronic healthcare databases
 - Access data from 25m individuals by July 2010
 - Access data from 100m individuals by July 2012

Calendar No. 270

110TH CONGRESS 1ST SESSION H.R. 2900

IN THE SENATE OF THE UNITED STATES

July 16, 2007 Received; read twice and placed on the calendar

AN ACT

To amend the Federal Food, Drug, and Cosmetic Act to revise and extend the user-fee programs for prescription drugs and for medical devices, to enhance the postmarket authorities of the Food and Drug Administration with respect to the safety of drugs, and for other purposes.

Be it enacted by the Senate and House of Representa-

т





Not-so-secret ingredients

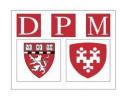
- Engaged partners
- Attention to data quality
- Reusable tools





Sentinel Partner Organizations

Lead - HPHC Institute



Data and scientific partners











Scientific partners

















America's Health Insurance Plans









Bringing engaged partners together

- Centralized vs distributed systems
- Distributed data system is preferred because
 - Data sits behind data partner's firewall
 - Data remains under local control
 - Only minimally necessary info is shared in a given analysis
 - Patient privacy and proprietary interests are preserved





Not-so-secret ingredients

- Engaged partners
- Attention to data quality
- Reusable tools





Sentinel Common Data Model

Administrative

Enrollment						
Person ID						
Enrollment start & end dates						
Drug coverage						
Medical coverage						
Medical record availability						

Demographic
Person ID
Birth date
Sex
ZIP code
Etc.

Dispensing							
Person ID							
Dispensing date							
National drug code (NDC)							
Days supply							
Amount dispensed							

Encounter
Person ID
Service date(s)
Encounter ID
Encounter type & provider
Facility
Etc.

Diagnosis
Person ID
Service date(s)
Encounter ID
Encounter type & provider
Diagnosis code & type
Principal discharge diagnosis

Procedure
Person ID
Service date(s)
Encounter ID
Encounter type & provider
Procedure code & type
Etc.

Clinical

Lab Result
Person ID
Result and specimen collection dates
Test type, immediacy & location
Logical Observation Identifiers Names and Codes (LOINC ®)
Test result & unit
Etc.

Person ID Measurement date and time Height and weight Diastolic & systolic BP Tobacco use & type Etc.	Vital Signs
Height and weight Diastolic & systolic BP Tobacco use & type	Person ID
Diastolic & systolic BP Tobacco use & type	Measurement date and time
Tobacco use & type	Height and weight
	Diastolic & systolic BP
Etc.	Tobacco use & type
	Etc.

Registry

Death	Cause of Death	State Vaccine		
Person ID	Person ID	Person ID		
Death date	Cause of death	Vaccination date		
Source	Source	Admission Type		
Confidence	Confidence	Vaccine code & type		
Etc.	Etc.	Provider		
		Etc.		



Sentinel Common Data Model

Administrative								
• 193 mil								
Person ID	Person ID	Person ID	Person ID	Person ID	Person ID			
Enrollment start and and answer end and answer end answer end and answer end answer end and answer end and answer end and answer end and answer end and answer end answer end and answer end answer end and and answer end and answer end and answer end and answer end and and answer end and answer end and answer end and answer end and and answer end and answer end and answer end and answer end and and answer end and answer end and answer end and answer end and and answer end and answer end and answer end and answer end and and answer end and answer end and answer end and answer end and a	million persor	n-years of obse	ervation time					
Drug coverage	Sex	National drug code (NDC)	Encounter ID	Encounter ID				
39 milli	• 39 million currently accumulating data provider Encounter type & provider							
Medical record availability		Amount dispensed	Facility	Diagnosis code & type	Procedure code & type			
 4 billior 								

• 5.5 billi	ion unique er	ncounters			
	Vital Signs				
Person ID	Person ID		Person ID	Person ID	Person ID
Populatio	ns with well-o	defined longit	tudinal pers	son-time for w	hich _{nation date}
most med	lically attende	ed events are	known		Admission Type
Logical Observation Identifiers Names and Codes (LOINC ®)	Diastolic & systolic BP				Vaccine code & type
	Tobacco use & type	•	Etc.	Etc.	
Ability to	<mark>obtain elect</mark> ro	onic or paper	medical re	cords	



Rigorous data checking and characterization

~1500 data checks with each refresh

0bs	ENCTYPE	ADATE	COUNT	PERCENT				0bs	px_codetype	enctype	COUNT	PERCENT
1	AV	2000	7030952	5.1370					09			0.0001
į	ΑŬ	2001	7454699	5.4466					09 09	AV	3891384 940211	0.2061 0.0498
3	ΑŬ	2002	8014346	5.8555	0bs	RXDATE	N	2 3	09 09	ED IP	7716848	0.498
4	ΑŬ	2003	8261199	6.0358	1 .			3	09 09	IS	168596	0.4088
5	ΑŬ	2004	8251011	6.0284	1	2000JAN	75816	;	09	0A	510196	0.0270
ă	ΑŬ	2005	8857635	6.4716	2	2000FEB	68872	2	C2	AV	4906255	0.2599
7	ΑŬ	2006	9576674	6.9969	3	2000MAR	240058	5	C2	ED	325738	0.0173
İ	ΑV	2007	10240959	7.4823	4	2000APR	248527	8	C2	IP	392155	0.0208
) š	ΑV	2008	11831682	8.6445	5	2000MAY	261254	9	C2	is	18219	0.0010
10	ΑV	2009	13785025	10.0716	6	2000JUN	258289	10	C2	0A	222605	0.0118
l ii	ΑV	2010	14499322	10.5935	7	2000JUL	241145	liĭ	C3	AV	212648	0.0113
12	ΑV	2011	14988289	10.9508	8 9	2000AUG 2000SEP	260316 252799	liż	C3	EĎ	5276	0.0003
13	ED	2000	193108	0.1411	10	2000SEP	260813	13	C3	ΪΡ	7755	0.0004
14	ED	2001	213180	0.1558	l ii	2000NOV	254161	14	C3	is	269	0.0000
15	ED	2002	231296	0.1690	12	2000DEC	259611	15	C3	ÖÄ	2030	0.0001
16	ED	2003	232122	0.1696	13	2001JAN	275314	16	C4	ΑV	1364119936	72.2580
17	ED	2004	230756	0.1686	14	20015FEB	242270	17	Č4	ED	95271865	5.0466
18	ED	2005	266406	0.1946	15	2001MAR	278558	18	C4	ĪP	50242438	2.6614
19	ED	2006	291381	0.2129	16	2001APR	260591	19	C4	IS	3914519	0.2074
20	ED	2007	314060	0.2295	17	2001MAY	268647	20	C4	OA	27959691	1.4810
21	ED	2008	343936	0.2513	18	2001JUN	267520	21	HC	ΑV	252901204	13.3963
22	ED	2009	400500	0.2926	19	2001JUL	257699	22	HC	ED	14811325	0.7846
23	ED	2010	414312	0.3027	20	2001AUG	279320	23	HC	IP	8125355	0.4304
24	ED	2011	451881	0.3302	ží	2001SEP	251170	24	HC	IS	1600478	0.0848
25	IP	2000	432504	0.3160					T HC	0A	31067795	1.6457
26	IP	2001	477466	0.3 <mark>Obs</mark>	Age_gr	oup	COUNT	PERCENT	ND	ΑV	16692216	0.8842
27	IP	2002	517710	0.3					ND	ED	639229	0.0339
28	IP	2003	543660	0.3 1	0.1 0-1		602059	1.4996	ND	IP	147970	0.0078
29	IP	2004	543692	0.3 2	02. 2-4		1376997	3.4298	ND	IS	12924	0.0007
30	IP	2005	587863	0.4 3	03. 5-9		2553188	6.3595	ND	0A	819916	0.0434
				4		4 Yrs	2638462	6.5719	OT	ΑV	194765	0.0103
				5	05. 15-1		2135457	5.3190	OT	ED	374	0.0000
				6		l Yrs	1670742	4.1615	OT	IP	2607	0.0001
				(l4 Yrs	14770481	36.7906	OT	IS	1367	0.0001
				8	08. 45-6		11221814	27.9515	OT	0A	348	0.0000
				'3	09. 65-7		1854092	4.6182				
				10	10. 75+	ırs	1324163	3.2982				



Why check after every refresh?

- Underlying data sources are dynamic
- Verify compliance with the common data model
- Identify changes in Data Partners' data sources or transformation processes
- Identify problems and/or differences in Data Partners' data transformation methods





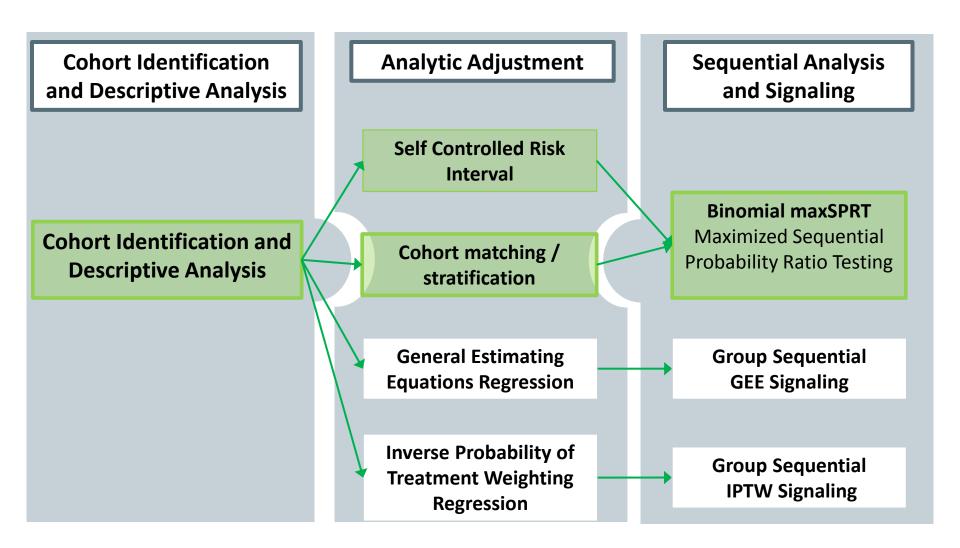
Not-so-secret ingredients

- Engaged partners
- Attention to data quality
- Reusable tools

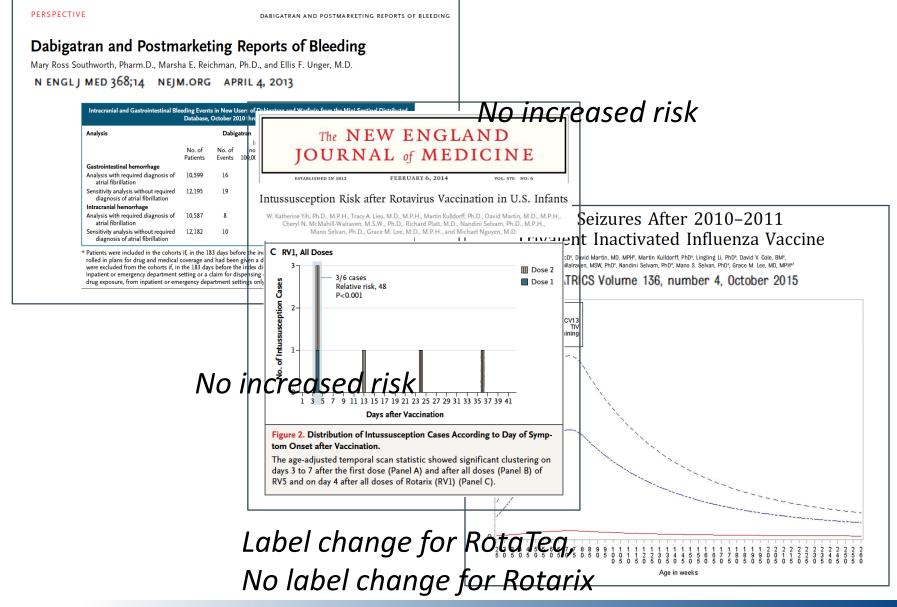




Reusable Rapid Query Tools









Impact / Dissemination

- 4 FDA drug safety communications
- 48 Methods reports / white papers
- 70 Peer-reviewed articles
- 137 Assessments of products, conditions, productoutcome pairs



Thank you