

Knowledge Management -- Data Mining a Healthcare Database

Brian L Rucco

Knowledge Management improves organizational performance. It is a discovery of experiences by the organization. It is a strategic goal.

An intellectual asset, Knowledge Management is developed by transforming the experiences -- data files, maintained by the organization. It provides insights gleaned across organizational functional lines to reveal process and continuous improvements

The Building Blocks for Healthcare Business Intelligence

The image displays four overlapping Microsoft Excel spreadsheets, each showing a different data table. The spreadsheets are titled 'Physician Information', 'Demographic Information', 'Clinical Information', and 'Financial Information'. The 'Physician Information' table lists MR#, EVENT#, Specialty, AttendingMD, AdmitMD, DischMD, and Surgeon. The 'Demographic Information' table lists MR#, EVENT#, AGE, GENDER, ZIPCODE, AdmitDate, and DischDate. The 'Clinical Information' table lists MR#, EVENT#, MDC, DRG#, LOS, GLOS, DischargeUnit, MedServ, 1stDx, AdmitDx, WorkDx, 2ndDx, 3rdDx, 1stTx, 2ndTx, and 3rdTx. The 'Financial Information' table lists MR#, EVENT#, FinType, CarrierPlan1, Subscriber#, CarrierPlan2, totCharges, ExpectdPymt, TotPayAmt, FinbillDate, Acc't Loc, Pay1, Pay1Date, Pay2, Pay2Date, Adjust1, and AdjustDate. The spreadsheets are overlaid on a green background with three large medical crosses (blue, green, and purple).

MR#	EVENT#	Specialty	AttendingMD	AdmitMD	DischMD	Surgeon
554662	4055935	142	430035	430035	430035	430054
235442	4055938	143	585775	785775	585775	
665534	4055944	142	235468	235468	235468	
204555	4055955	143	522266	522266	522266	
546633	4055965	140	665435	665435	665435	

MR#	EVENT#	AGE	GENDER	ZIPCODE	AdmitDate	DischDate
554662	4055935	64	M	10152		
235442	4055938					
665534	4055944					
204555	4055955					
546633	4055965					

MR#	EVENT#	MDC	DRG#	LOS	GLOS	DischargeUnit	MedServ	1stDx	AdmitDx	WorkDx	2ndDx	3rdDx	1stTx	2ndTx	3rdTx
554662	4055935	5	291	6	5	3N	CAR	428.1		414.8					
235442	4055938	4	190	4	4.7	4W	PUL	496			790.22	290.2			
665534	4055944	5	280	4	3.4	HS	CAR	410.9							
204555	4055955	2	203	3	2.6	HN	PUL	406							
546633	4055965	4	203	3	2.6	HN	PUL	406							

MR#	EVENT#	FinType	CarrierPlan1	Subscriber#	CarrierPlan2	totCharges	ExpectdPymt	TotPayAmt	FinbillDate	Acc't Loc	Pay1	Pay1Date	Pay2	Pay2Date	Adjust1	AdjustDate
554662	4055935	M	555001	A45975		\$54,325	47,800	43,500	9/15/xx	AR	43,500	10/22/xx				
235442	4055938	C	720550	10-985-45	4-7755-03	\$37,855		38,450	11/12/xx	AR						
665534	4055944	B	3010210	-267455		\$42,600				AR						
204555	4055955	W	735110	02-564-766		\$35,224				AR						
546633	4055965	M	555001	A76833		\$24,685				AR						

Carrier Plan Codes

- 301010 Blue Cross Plan
- 555001 Medicare
- 655002 Medicare + HOM Plan
- 800001 Medicaid
- 735110 State Workers Comp

Financial Type

- M - Medicare
- C - Commercial Plan
- B - Blue Cross Plan
- W - Workers Comp

Acct Location

- AR Acc't Rec
- BD Bad Debt

merging data across functional areas creates a repository of service experiences applied to data mining tasks for root cause analysis and gain information within performance indicator stats



Core
Components
When Data
Mining a
Healthcare
Database

- **Database** containing data about every patient encounter performed by the facility. Facility's typically have multiple software applications – so it is key to acquire event linkage information (data elements) to build a single significant file. This enables the intermingling of the patient's demographic, clinical and financial information.
- **Software tools** for extracting detailed data within a statistic and from integrated databases. It is significant to explaining patterns, trends, clusters discovered within the patient encounter database
- **Knowledge Base** of the healthcare industry (public policy thought, financing mechanisms, quality initiatives) to present empirically based solutions. Understanding these relationships provides insights and benchmarks that are cogent for leadership and administrators alike.

Database Essentials

Demographic

- age
- sex
- Zip
- Ethnicity
- Marital status
- Admit Source

Clinical

- 1st-Diagnosis
- working diagnosis
- admitting diagnosis
- co morbidities
- MDC
- DRG#
- Tests/drugs/meds per charge item
- Admit date
- Discharge date
- ED visit date
- Discharge Status

Financial

- financial class
- insurance type
- Insurer contract #
- total charges
- net revenues
- account balance
- Account location
- Amt paid by insurer
- Final Bill Date
- Paid dates
- Payment codes
- Adjustment reasons

Physician

- attending
- surgeon
- ED-MD
- Admit MD
- Consult MD
- Med specialty
- Med license

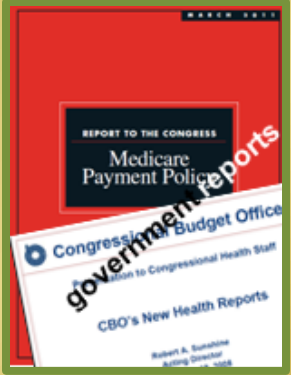


Other Data Content Considerations

- pt account #
- MRN – medical record number
- admits via the ED
- patient safety (infection, falls, med errors) *per* ICD-9
- Never Events
- primary language
- clinic visit history
- ED visits
- nursing station locations
- census tract residence
- employment status
- amburg encounters
- Patient Type
- Readmission Status
- Primary Care MD
- Care Provider IDs
- total cost *per case/per day* estimate
- denied payment days
- payment dates/amounts
- guarantor Information
- expected reimbursement amount
- patient co-pays
- Alternative Level of Care Days
- Total FTE staffing
- CPT codes
- HCPCS codes

linking to external sources =
Meaningful Measurements

statistics	code	target
Actual Acute Discharges	#1	2000
Mcare CMI (all patients)	#5	1.35
Number- Ambulatory-Surg Cases	#1	3,500
Average LOS (per mo discharge)	#4	6.4



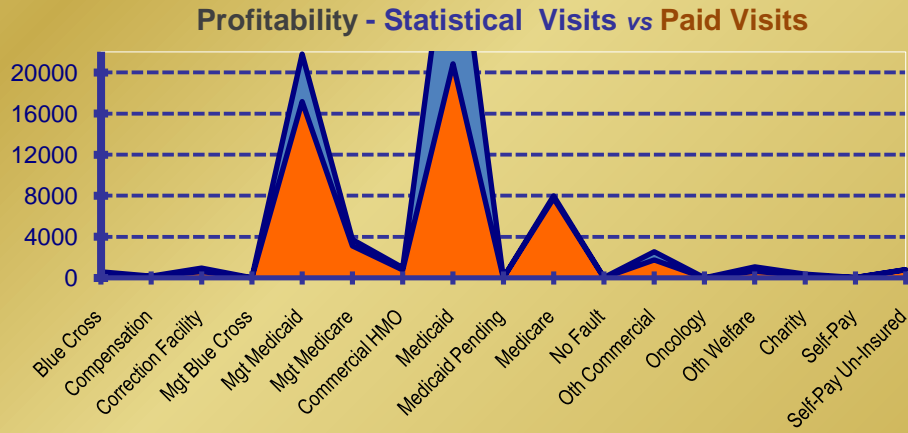
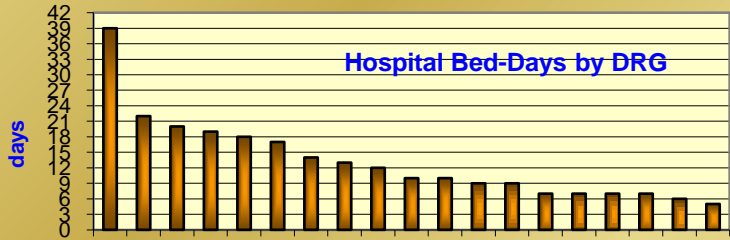
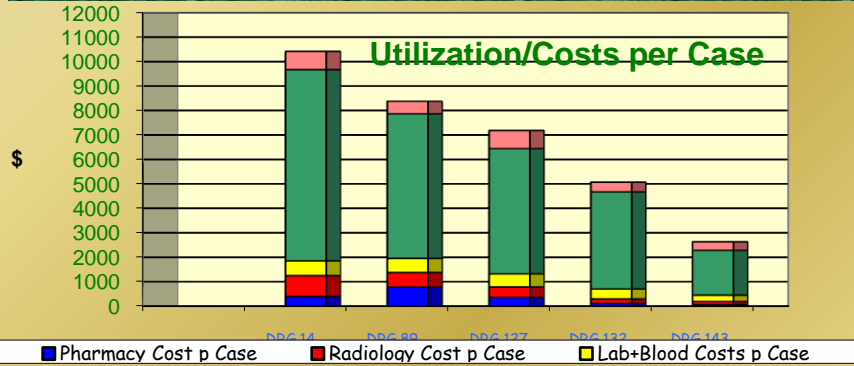
Circulopulmonary Mortality Rate	#5 ORYX	<3.2%
General Surgery Mortality	#5 ORYX	<2.5%
Class I Wound Infection Rate	#3	<2%
Pressure Sore Incidence Rate	#3	<3.43%



096 - Asthma w cc	1,2	3 days
294 - Diabetes	1,2	5 days
090 - Pneumonia w o cc	1,2	5 days



Insights Gleaned When Data Mining a Healthcare Encounter Database



Performance Indicators								
statistics	code	target	Prev FY	Qtr-1	Qtr-2	Qtr-3	Qtr-4	Total
Actual Acute Discharges	#1	2000	20500	4000	5500	6000	5500	21000
Mcare CMI (all patients)	#5	1.35	1.2255	1.355	1.245	1.685	1.695	1.3776
Number - Ambulatory-Surg Cases	#1	3,500	3250	730	850	885	790	3255
Average LOS (per mo discharge)	#4	6.4	6.3	6.6	6.4	6.4	6.3	6.4
quality of care measures								
Circulopulmonary Mortality Rate	#5 ORYX	<3.2%	3.4%	5.50%	8.3%	6.5%	5.0%	7.6%
General Surgery Mortality	#5 ORYX	<2.5%	2.8%	3.50%	4.4%	1.5%	4.6%	3.8%
Class I Wound Infection Rate	#3	<2%	0.3%	0.95%	0.82%	0.00%	0.00%	0.00%
Pressure Sore Incidence Rate	#3	<3.43%	3.0%	2.40%	3.2%	1.6%	1.6%	2.3%
Referred Tests								
- radiology	#8	3400	3300	887	783	780	794	3244
- lab tests	#8	2200	2100	511	496	504	505	2016
DRGs w Pathways								
096 - Asthma w cc	1.2	3 days	95/4.0	10/5.7	17/4.2	5/4.0	10/5.7	42/4.0
294 - Diabetes	1.2	5 days	162/7.5	12/11.5	12/15.6	12/11.5	13/5.7	49/7.6
090 - Pneumonia w o cc	1.2	5 days	77/4.3	9/4.87	10/6.4	9/4.8	10/4.5	38/6.0
Payers								
Mgt Blue Cross	#1	35	0	3	32	10	0	45
Mgt Medicare	#1	5500	36	216	709	3391	366	4682
Commercial HMO	#1	2400	9	100	1002	888	137	2127
Medicaid	#1	3008	624	403	6600	22772	5263	35038
Medicare	#1	15000	129	337	1615	8505	997	11454
Oth Commercial	#1	4575	7	197	995	1967	166	3325
Self-Pay	#1	250	0	129	5	41	23	198