Predicting future resource use & risk of hospitalization for a general population in NHS England:

Adapting US models & potential lessons for the US

Stephen Sutch

Johns Hopkins Bloomberg School of Public Health

To be presented at The Predictive Modeling Summit
Washington, DC, November 14, 2014

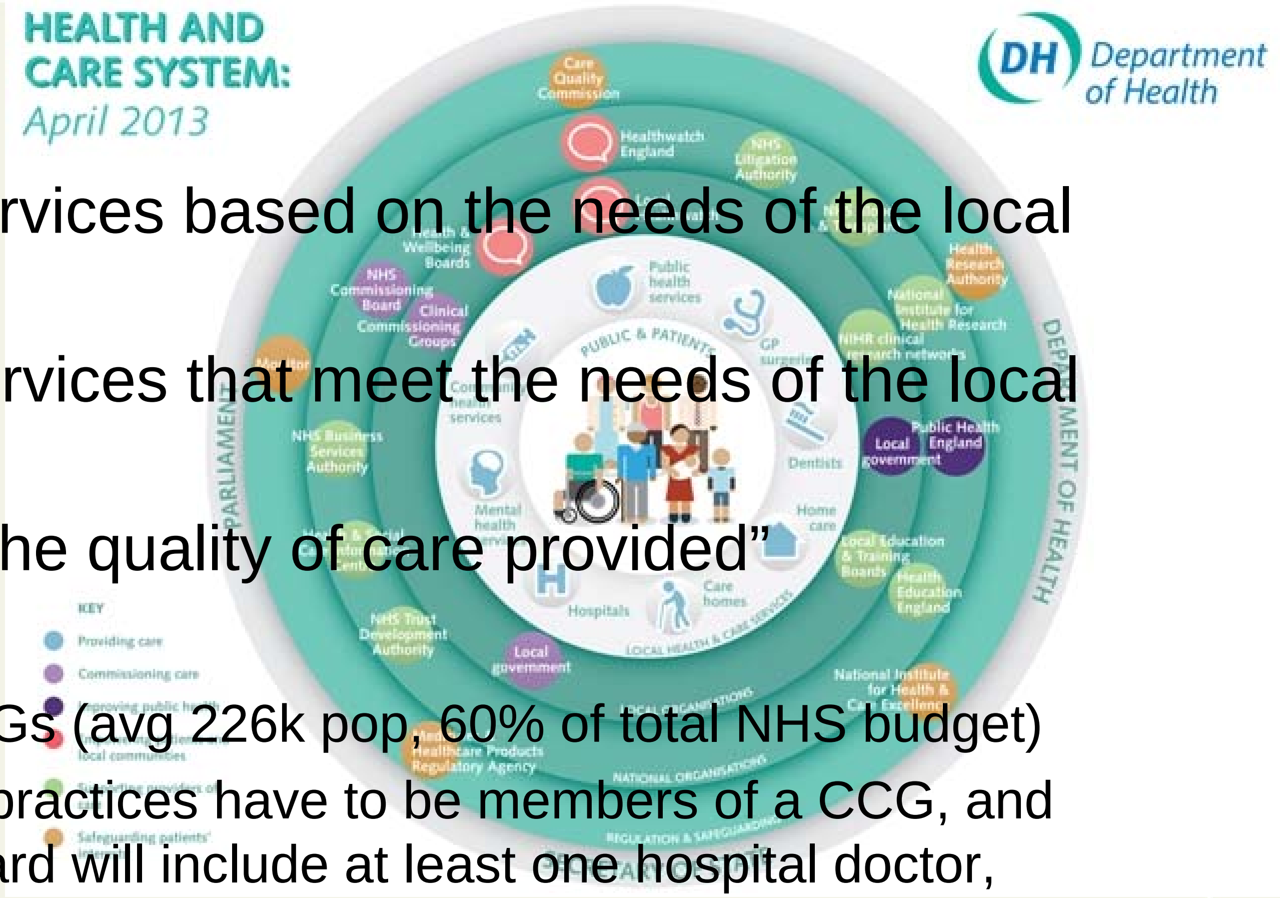
Introduction

- A number of models are available in the US and the UK which predict the risk of hospitalisation, from general and insured populations
- Multiple purposes e.g. screening of patients for Case Management Programs, screening for Disease Management Programs, organisational profiling, and assessing financial risk.
- Response to health policies to reduce unnecessary hospital admissions, Pay for Performance (P4P) measures, Risk stratification tool requirements
- A need to support populations in avoiding hospital admissions that are both expensive and a patient safety risk.

Historic Use of Models in England

- Existing predictive models in the ACG System were based on US data, rescaled on local data
- Early work at Imperial College and UCL showed the applicability of the ACG System to NHS data.
- In 2006, Johns Hopkins University and the Kings Fund created predictive models from NHS data.
- Leeds City PCT showed existing models in ACG System could match and exceed the performance of the Combined Predictive Model (CPM).
- Currently used in NHS to create lists of individuals for clinical review, care management to prevent unnecessary hospital admissions.

Role of Clinical Commissioning Groups (CCG)

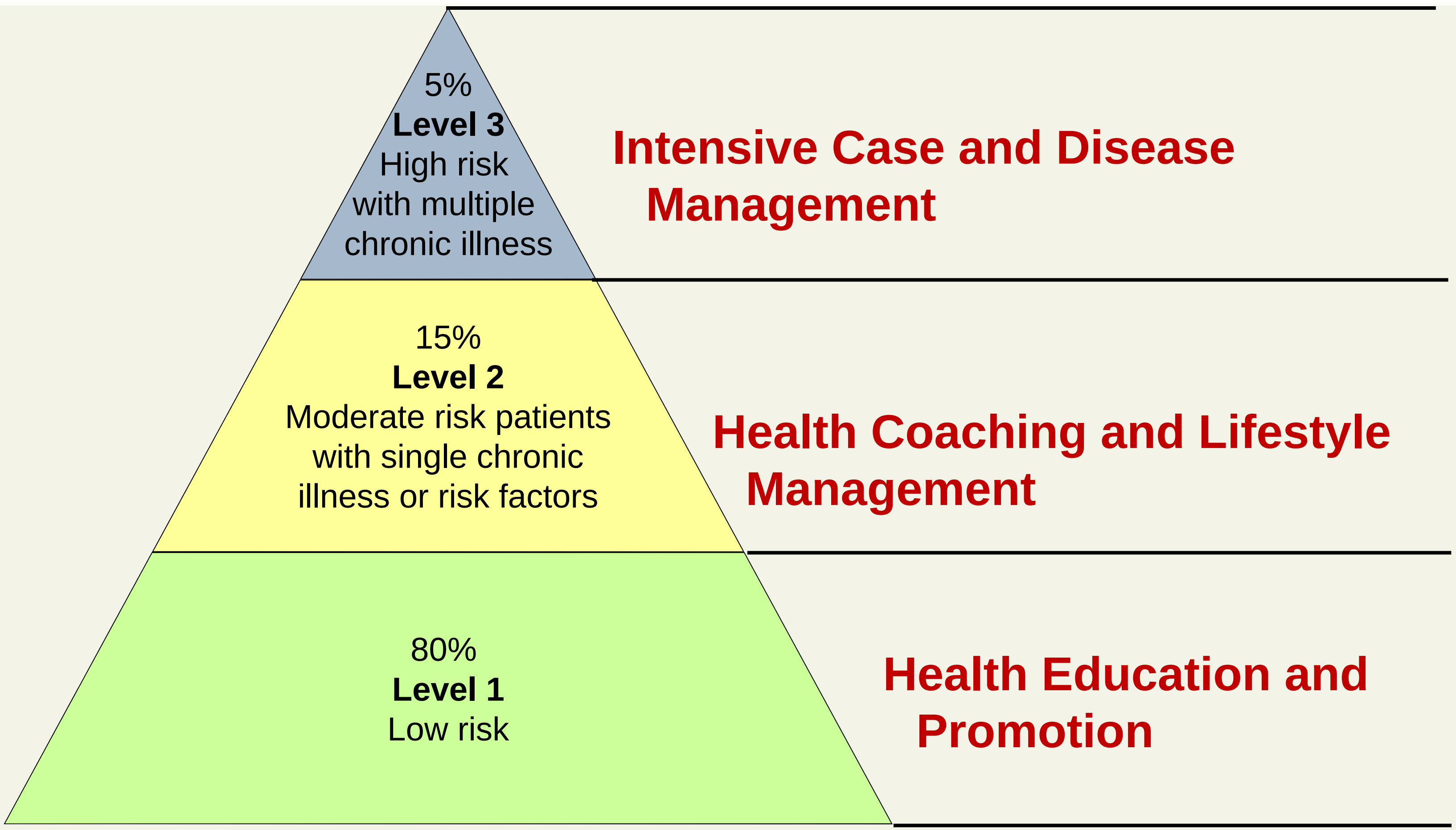


- “Planning services based on the needs of the local population”
- “Securing services that meet the needs of the local population”
- “Monitoring the quality of care provided”

- 2013 - 211 CCGs (avg 226k pop, 60% of total NHS budget)
- “All GP *(PCP)* practices have to be members of a CCG, and every CCG board will include at least one hospital doctor, nurse and member of the public.”

Source: http://www.patient.co.uk/

Using Predictive Modeling to Assign Persons Within the Care Management Pyramid



ACG System predictive models used to generate an outreach “list” for GPs, care management nurses / Community Matrons

Care Management List for Comm_Demo.acgd

Cases | Report Options

Patient Id	Age	Sex	Total Cost	Rescaled Total Cost Resource Index	Probability High Total Cost	Probability IP Hospitalization	Unique Provider Count	Rx Gaps	Hospital Dominant Count	Chronic Condition Count	Frailty Flag
411316*141931615	75	M	9,667.77	23.37	0.95	0.50	3	0	1	11	☐
88493968*8211951	55	M	158,168.03	23.73	0.95	0.50	6	0	5	9	☐
411443*16193717	69	M	75,228.16	26.85	0.95	0.90	4	2	4	23	☐
88494700*3141950	57	M	281,738.71	25.79	0.95	0.74	5	0	5	10	☑
412111*161961117	45	M	38,582.56	29.98	0.95	0.57	4	3	3	18	☐
88495119*951945	61	M	21,723.69	25.70	0.95	0.86	7	2	2	12	☐
413644*14195814	48	F	82,005.22	28.04	0.95	0.99	6	1	2	18	☑
6214215*14194916	57	M	211,501.05	28.30	0.95	0.78	5	8	2	22	☑
414137*16194556	61	F	28,766.14	22.63	0.95	0.62	5	4	1	17	☐
6215421*141952328	54	F	110,792.19	22.99	0.95	0.79	5	1	2	19	☐
414447*1419431218	62	F	70,029.86	22.96	0.95	0.43	4	1	3	23	☐
6221565*13197135	35	F	54,057.82	23.05	0.95	0.53	4	0	2	20	☐
414474*141939514	67	M	40,321.38	23.43	0.95	0.90	12	8	0	22	☑
6227552*1619421128	63	M	72,871.18	23.69	0.95	0.64	3	0	3	21	☐
416153*141920127	85	F	19,894.41	25.20	0.95	0.85	7	0	2	25	☑
6241673*1619501127	55	M	121,288.81	38.24	0.95	0.93	4	0	2	34	☑
416334*1619371130	68	M	66,142.02	31.58	0.95	0.95	9	0	5	29	☑
6261654*16194335	63	F	86,538.62	29.92	0.95	0.85	12	11	4	26	☐
416423*141956116	49	F	81,431.14	25.72	0.95	0.97	6	1	1	26	☐
6262121*141942115	64	M	125,962.55	42.68	0.95	0.90	5	0	2	35	☐
417446*1719861027	19	M	220,751.96	38.46	0.95	0.94	8	0	5	30	☐
6645414*141938331	68	F	96,138.29	22.96	0.95	0.92	9	0	3	19	☑
444232*1419351228	70	F	76,588.26	37.92	0.95	0.70	7	2	4	21	☐
6672724*1419501226	55	F	11,361.46	23.86	0.95	0.36	1	0	1	12	☐
444441*1619271018	78	F	67,767.97	29.27	0.95	0.71	3	2	3	12	☐

Comprehensive Patient Clinical Profile (summary)

Comprehensive Patient Clinical Profile Report - Patient Id: 7442522*16195151

Age	55	Gender	M
PCP Id	5212*11	Product	HMO
Resource Utilization Band	5	Local Weight	9.55

Model

DxRx-PM - total cost - lenient dx -> total cost

DxRx-PM - rx cost - lenient dx -> rx cost

Prior Costs

Total Cost	$ 26,951
Rx Cost	$ 2,700

Predictive Values

Probability High Total Cost	0.65
Predicted Total Cost Range	$30,000-$40,000
Probability High Rx Cost	0.44
Predicted Rx Cost Range	$2,000-$3,000
High Risk Unexpected Pharmacy	N

Coordination of Care

Chronic Condition Count	12
# Unique Providers Seen	2
# Specialty Types Seen	2
No Generalist Seen	Y
% Visits Provided By Majority Source of Care	67
Frailty Flag	N

Utilization

Outpatient Visits	55
ER Visits	3
Inpatient Admissions	0
Major Procedure Performed	Y
Dialysis Service	N
Nursing Service	N

Likelihood of Hospitalization

Hospital Dominant Count	3
Probability Hospital Admission (6 mos)	0.04
Probability Hospital Admission (12 mos)	0.07
Probability ICU/CCU Admission	0.01
Probability Injury-related Admission	0.02
Probability Long-term Admission (12+ days)	0.01

Condition Profile with Pharmacy Adherence

Condition	Present?	CSA	MPR	# Refill Gaps	Untreated
Age-Related Macular Degeneration	NP				
Bi-Polar Disorder	NP				
Congestive Heart Failure	NP				
Depression	ICD				
Diabetes	ICD				
Glaucoma	ICD				
Human Immunodeficiency Virus	NP				
Disorders of Lipid Metabolism	Rx				
Hypertension	ICD				Y
Hypothyroidism	ICD				
Immunosuppression/Transplant	NP				
Ischemic Heart Disease	NP				
Osteoporosis	ICD				Y
Parkinson's Disease	NP				
Persistent Asthma	Rx				
Rheumatoid Arthritis	NP				
Schizophrenia	ICD				Y
Seizure Disorders	NP				
COPD	ICD				
Chronic Renal Failure	NP				
Low Back Pain	ICD				

NP = Not Present, ICD = ICD Indication, Rx = Rx Indication, BTH = ICD and Rx Indication, TRT = Treated with Pharmacy

High Impact Conditions

EDCs		Rx-MGs	
GAS02	Inflammatory bowel disease	RESx020	Respiratory / Chronic Medical

Moderate Impact Conditions

EDCs		Rx-MGs	
CAR14	Hypertension, w/o major complications	CARx040	Cardiovascular / Disorders of Lipid
END02	Osteoporosis	GASx010	Gastrointestinal/Hepatic / Acute Minor
END06	Type 2 diabetes, w/o complication	GASx060	Gastrointestinal/Hepatic / Peptic Disease
MUS14	Low back pain	GSIx020	General Signs and Symptoms / Pain
NUR03	Peripheral neuropathy, neuritis	GURx010	Genito-Urinary / Acute Minor
PSY01	Anxiety, neuroses	MUSx010	Musculoskeletal / Gout
PSY07	Schizophrenia and affective psychosis	PSYx050	Psychosocial / Acute Minor
PSY09	Depression	RESx040	Respiratory / Airway Hyperactivity
REN02	Fluid/electrolyte disturbances		
REN03	Acute renal failure		
RES02	Acute lower respiratory tract infection		
RES04	Emphysema, chronic bronchitis, COPD		

Low Impact Conditions

EDCs		Rx-MGs	
ADM02	Surgical aftercare	GSIx030	General Signs and Symptoms / Pain and
ADM05	Administrative concerns and non-specific	INFx020	Infections / Acute Minor
ADM06	Preventive care	ZZZx000	Other and Non-Specific Medications
ALL03	Allergic rhinitis		

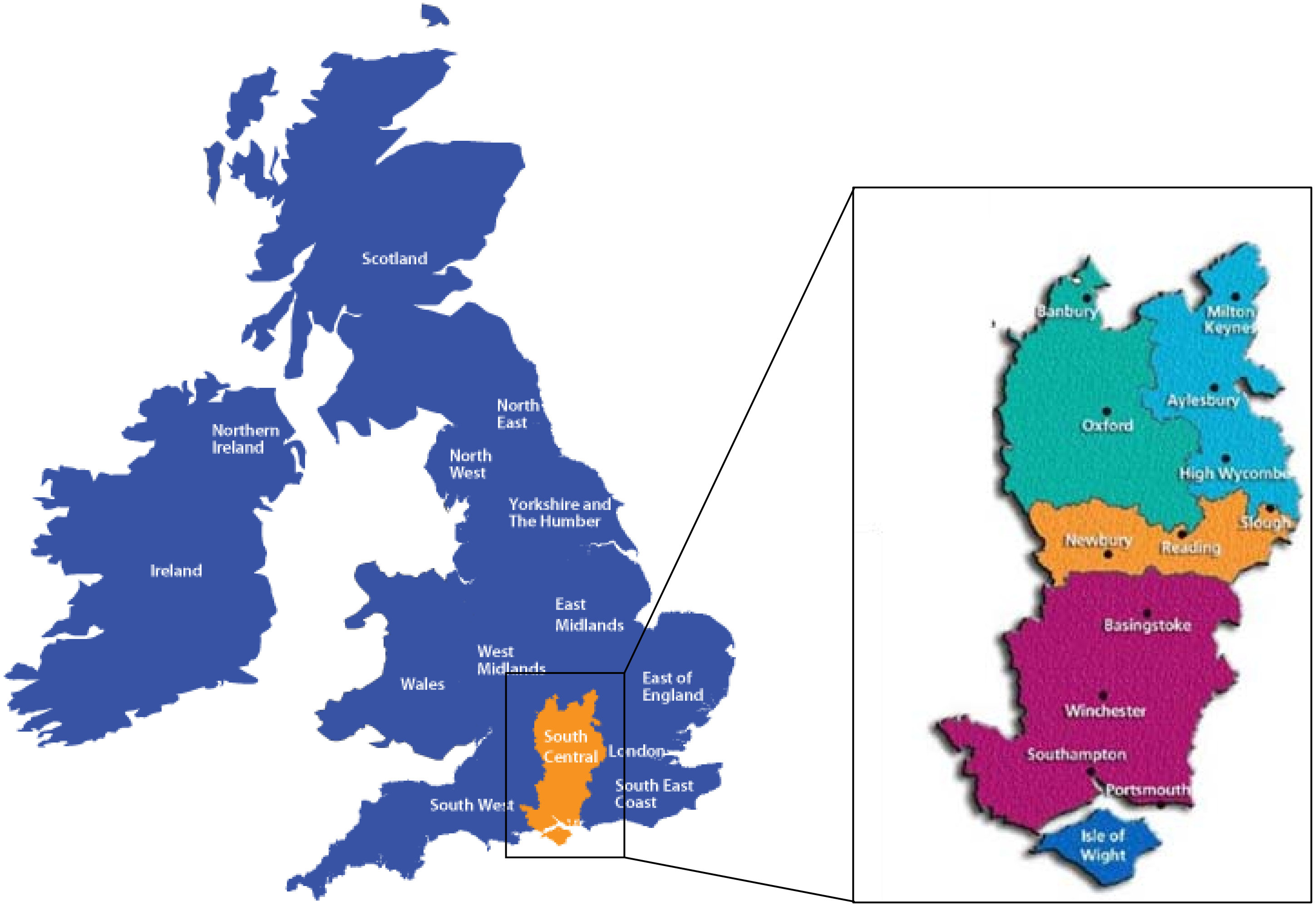
Example Clinical Process

- Identify at risk patients – ACG risk profiling tool
- Core medical team review
  - Identify problems, Action list, Suitability for further interventions
- Personalized care plan
  - Discussion and delivery of care plan, Coded and scanned to records
- Follow-up
  - Clinical review (named clinician), Date of review, Response to interventions

Source: Cricket Green Medical Practice Model

acg JOHNS HOPKINS

The South Central Region of the NHS



9 primary care trusts (PCTS)

510 GP practices clustered into 20 CCGs

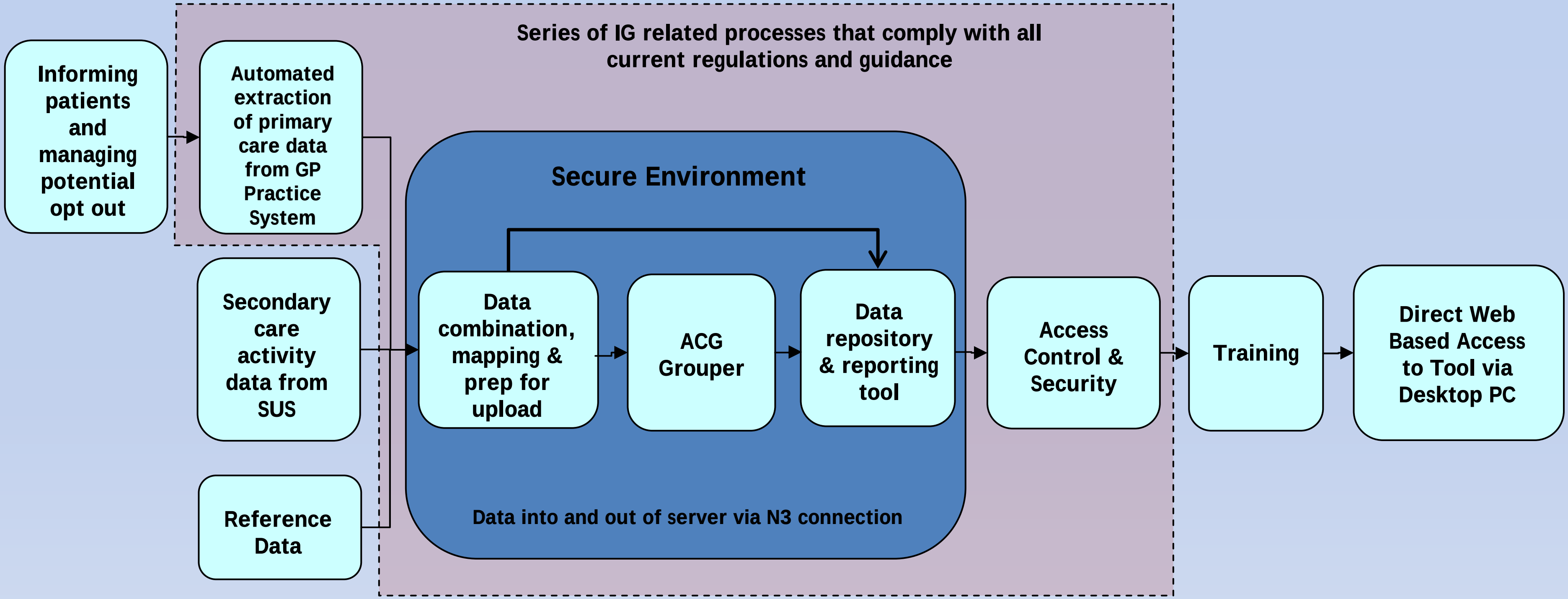
4 million population

PCTs currently responsible for commissioning of services

ACGs in use in approximately half of GP practices

Implemented ACG Solution

NHS South Central
Primary Care Trust Alliance



- A complex end-to-end infrastructure that took over 9 months to put in place but:
  - It addresses all of the issues/concerns/requirements of our stakeholder group particularly around the issue of transferring, storing and sharing data, particularly primary care data
  - Primary care data extraction – a complex and resource intensive process - is undertaken by a specialist company rather than PCT staff
  - End users have access to a user-friendly graphical interface on their desktop
  - It only takes 4-6 weeks from a GP practice opting in and having access to ACG information

Method

- Aim: apply the ACG System variables as independent variables in year 1, to predict patient outcomes in year 2
- Two main dependent (outcome) variables,
  - total cost in year 2 (Linear Regression)
  - hospitalization in year 2 (Logistic Regression)
- Objectives
  - create predictive models from English NHS data
  - validate those models (split half validation)
  - compare with the existing US-based models
  - recommend a model for application England.

Risk Factors in the Johns Hopkins Predictive Model

Results (1)

- Data: 663,797 individuals in year 1
- extracted from primary care practices which had completed and approved a consent process.
- Secondary care data was added from hospital data for cases where patients had also received hospital services.
- linear regression to predict future (year 2) total patient expenditure, R-Square 27.5% untrimmed
- R2 8.8% age/gender, 22.4% US based models
- With prior cost and utilisation variables added the model's performance increased to 30.9%

Future annual cost - NHS England 2013 R Squared Results

	Age / Gender	ACG w/o Prior	ACG w Prior	ACG US All Age
Train	.0910	.2792	.3010	.2238
Validation	.0902	.2745	.2943	.2260

Hospitalisation Prediction - C-Statistics NHS England 2013

	12 Month Admission	6 Month Admission	>12 day LoS	Unplanned Admission	Multiple Emergency
Train	0.795	0.814	0.915	0.781	0.854
Validation	0.795	0.815	0.904	0.781	0.852

Results (2)

- logistic model to predict unplanned hospitalization
  - C-Statistic 0.78
  - Directly related to measure used in P4P program for PCPs (NHS QoF)
    - Reduction in avoidable hospital admissions
  - “Emergency Admissions” (3.74%)

Risk of Unplanned Admission (3.74%) Sensitivity / PPV, NHS England 2013

Split	Cut Pt	Sens	Spec	PPV	NPV
50%	.0198	84.6%	51.3%	6.3%	98.90%
90%	.0790	43.66%	91.28%	15.99%	97.71%
95%	.1227	29.91%	95.95%	21.90%	97.30%
98%	.2048	16.56%	98.55%	30.31%	96.88%
99%	.2874	10.30%	99.35%	37.71%	96.68%
99.5%	.3817	6.00%	99.71%	43.89%	96.54%

Discussion (1)

- The results show a statistically significant improvement over the existing models available in the ACG System implemented in the UK NHS, consistent with similar projects carried out in Sweden and Spain
- The original US models still provided good sufficient estimates that have been proven to be robust in a number of countries over several decades.

Conclusion (1)

- Casemix classifications reduce data complexity and provide robust measures of multimorbidity. The models work well in explaining the top 1% and 5% of data, but also perform well in discriminating risk “lower in the population pyramid” to identify potential emerging risk.
- Current emphasis on identifying the highest risk individuals, there is an increased interest in recognising earlier and **emerging risk**, where more preventative methods can be informed such as chronic disease self-management programs.

Conclusion (2)

- A standard set of independent variables were used in the models. Additional variables could be used in future models such as BMI, Smoking Status, and social care data.
- Alternative models can produce higher results by using current utilisation and costs measures, however these models would increase bias to individuals already accessing healthcare services to the detriment of those with low current access.
- Including prior utilisation and prior cost measures as independent variables also creates perverse incentives to increase resource use."

Discussion (1)

- **Intermediate Classification**
  - Form a set of independent variables from 1000s of input variables
  - Dependent Variable, move from Any admission to unplanned/emergency/preventable
- **Additional Variables, Data**
  - Additional variables could be used in future models such as BMI, Smoking Status, and social care data.
- **Alternative models needed**
  - Historic utilization can produce higher results but bias to individuals already accessing healthcare
  - creates perverse incentives to increase resources
  - Dependent variable, Unplanned admissions

Discussion (2)

- Creating alternative Views
  - Concurrent v Prospective (Performance measurement v Planning)
  - Individuals, Populations
- Longitudinal data, Changing Risk
  - Increasing, decreasing, see-sawing
  - Real-time alerts
- EHR and Social Data
  - Data linkage, assessments, labs
  - Patient data - Health Status, Behaviour, Self-Assessment (e.g. SF12/36, EQ5D, PAM, HRA, PHQ9)
  - Selection Bias (Non-response, Exclusion bias)

Opportunities for Learning more….

- Web Site:
  - www.acg.jhsph.edu
- Contact:
  - Steve Sutch, Dir. Product Management, ACG International
    ssutch1@jhu.edu

Results - Hospitalisation

- logistic model to predict future hospitalisation
  - C-Statistic 0.80
  - age/gender model 0.67
  - current US model 0.75
  - For purposes of generating lists of high risk individuals applying a cut-point such that 1% of the population are designated as “positive”, the model showed a positive predictive value of 65.46%