What We All Know

- ...$.30 to $.40 of every dollar spent on health care...is spent on costs associated with “overuse, underuse, misuse, duplication, system failures, unnecessary repetition, poor communication, and inefficiency.”
  IOM 2008

- 3/4 of patients surveyed described their medical encounter as “a nightmare to navigate”
  American Hospital Assoc 2008

- 80,000 hospital deaths and 150,000 ambulatory deaths occur annually from Error
  IOM 2006, 2008
O/R Techniques Have Little Impact

- Demand Forecasting has not reduced ED overcrowding, diversions, “sentinel” events or “boarding”
- Scheduling optimization has not reduced OR under-utilization or staff overtime
- EMR’s have neither reduced error (though they’ve changed the pattern), improved data analysis or improved patient–provider communication
Numerous National Quality Initiatives

- Pharmacy Quality Alliance launched
- Hospital Quality Alliance launched
- Ambulatory Quality Alliance launched
- Creation of The Leapfrog Group
- Creation of Bridges to Excellence
- Deficit Reduction Act mandates expansion of measurement and sets precedent for lack of add-on payment for errors
- Medicare Modernization Act ties hospital market basket updates to quality reporting for 10 measures
- IOM Report To Err is Human: Building a Safer Health System
- JCAHO launches the core measures initiative
- IOM Report Crossing the Quality Chasm
  - Focused on a redesign of health care delivery
  - Called for creation of performance-based payment
- JCAHO launches the ORYX Initiative
- CMS Nursing Home Compare launched
- CMS Home Health Compare launched
- Hospital Compare launched
- IOM Report Performance Measurement Accelerating Improvement
- CMS Roadmap to Quality launched
- Hospital Compare expanded to payment and volume information and HCAHPS patient experience data
- Value-Based Purchasing Report to Congress on the Plan to Implement a Medicare Hospital VBP Program
- CMS Preventable Events
- AHIC Quality Workgroup Approved

...Adopting a Number of Quality Improvement Techniques

- Total Quality Management
- Continuous Quality Improvement
- Benchmarking
- Evidence-based Medicine
- Clinical Practice Guidelines
- Sigma Six
- Toyota Lean
- Pay-for-Performance
...Using “Rearview Mirror” Quality Measures

- % of 2 years olds fully vaccinated
- % of diabetics with HbA1c screenings
- % of depressed patients on medications
- Joint Commission LD 3.10 and MM series
- Press-Ganey Patient Satisfaction Scores

ALL SYSTEM OUTPUTS RATHER THAN DETERMINANTS OF QUALITY
...Yielding

Nelson DE et al. JAMA 2002;287:2659-67
...Begs The Question

- How can a national system that leads the world in technological and biological innovation be so much less than the parts?
The Answer

Medicine lacks a *science* of health care quality

- Incoherent Health Service Research Agenda
- Initiatives treat quality as a black box
- Unable to describe the relationship between satisfaction, outcomes and efficiency
Science of Health Care Quality Requires…

- Theory
- Axioms
- Testable Hypothesis
- Field Observations
- Methodologies
- Identification of Key Constraints
The Phenomenon to be Understood

- Patients arrive with
  - Biologic conditions
  - Health beliefs
  - Health experience
  - Health goals/desires

- Providers Arrive with
  - Knowledge of human biological systems
  - Knowledge of treatment options
  - Knowledge of legal and ethical constraints
  - A sense of what is possible
Characteristics Shared with Other Service Industries

- Co-production of value
- Perishable, heterogeneous, and intangible (but measurable)
- Inputs are human capital and knowledge
- Production and consumption are inseparable (but the value derived have different temporal gradients)
- Requires significant levels of mass customization and service provider judgment
Characteristics Unique to Health Care

- Consumer has little choice
- Consumer has little ability to objectively judge the quality of the service provided
- Timing of service can be critical
- Consumer satisfaction *NOT* meaningfully defined the same as in other industries
- Severe step-wise cost functions
This Dialectic Process:

(Axioms or First Principles)

1. Is dynamic and transformative
2. Information intensive
3. Often time-sensitive
4. Predicting and matching service demand to service capacity is critical
5. "Outputs" are a customized (but categorical) patient and provider mutual understanding of what is possible.
Implications

- A core Service Science tenet posits that clients must take an active role as co-producers of the knowledge based solution to achieve a successful outcome
  - Technically correct care that is either not wanted or not understood erodes the “value” of clinical encounter
- Difficult to specify the terms of the interaction a priori
- Must agree on the expectations and scope of the consultation
  - Which will be dynamic and evolutionary
  - Knowledge transfers will be “sticky”
Service Science Challenges

- Enabling technologies for smarter information use
- Reducing and managing uncertainty
- Quality assurance in real time
- When should variability be reduced—and would it not
- Meaningful measurement systems
- Modeling and predicting processes
- Satisfization of resource use
Mutual Understanding Theory of Health Care Quality Frontiers

▪ The state of what is possible is defined as the health care quality potential
▪ This potential must be dynamic to meet changes in medical knowledge, the health system, and patient resources
▪ The dialectic process is evolutionary, as beliefs, knowledge and perceptions are altered by the process
▪ The interactions amongst the variables in this system are knowable and quantifiable and produce the quality potential frontier
Developing Objective Functions for:

Provider Variables
- Practice Setting
- Arrival Patterns
- Informational Needs
- Provider Personality

Patient Variables
- Patient Health Status
- Patient Health Beliefs
- Patient Circumstances
- Patient Personality

Provider-Patient Interaction

Individualized Desired Outcomes
Beginning the Mathematical Model

- Definition patient understanding array $PU$, and provider understanding array $DU$.
- Definition of element variables within understanding array $A$.
- Definition of component transfer function vector $Z_{Ci}^P(PU)$ and $Z_{Ci}^l(PU)$
- Definition of component inequality constraints (Potential/Frontier) $P V_{Ci}(A)$
- Definition of component current value $C V_{Ci}(A)$
Key Determinants or Inputs

Ratio of Patient-Provider time to all other time in Service

– Internal Parameters
  • (Service Capacity = staffing, exam rooms, scheduling schemas)

– External Parameters
  • (Service Demand = stochastic for ED, deterministic for Specialty, mixed for Primary)

– Information Parameters
  • (External for Specialty, Internal for ED mixed for Primary)
Developing a Mathematical Model

- Principle Component Analysis
  - Outcomes *(the Gap between what is possible and what has occurred for individual patients)*
  - Patient satisfaction *(rooted in likelihood of compliance with agreed-to treatment recommendations)*
  - Efficiency *(ratio of patient-provider time to all other time in clinic)*
Examples form the Field
Field Tests

- Direct Observations in
  - Primary care (Pediatric, Emergency Departments, Family and GIM clinics)
  - Rural and Urban
  - Specialty care (e.g. Nephrology, Orthopedics, Cardiology, HIV)
  - Large (V.A. and Kaiser), Medium (15-60 providers) and small (>3 providers) clinics
Field Study 1: Matching Service Demand to Service Capacity in an Emergency Department

- Nearly half of U.S. hospitals report ED crowding on a daily basis
- Seriously ill and injured ED patients can wait 12-48 hours for an inpatient bed
- > ½ million inbound ambulances diverted in 2007
- “Crowding” / imbalance in patient/staff ratio responsible for 35% of “sentinel events”
Emergency Department Constraints

- Must accept all takers
- Legal mandates (with financial penalties) on staffing
  - 1 RN per 3 low acuity
  - 1 RN per 2 medium acuity
  - 1 RN per 1 high acuity
- Joint Commission Requires ED’s to examine Patient flow (JCAHO LD.3.15 leadership standard)
- Severe Step-wise cost functions
ED Benchmarks

- Arrival to triage: 6 min
- Triage to room: 23 min
- Arrival to MD: 53 min
- Arrival to admit: 253 min
- Arrival to discharge: 139 min
- Lab turnaround: 49 min
- Radiograph turnaround: 47 min
- LWOT: 2.3%
- Pts per MD hour: 2.25 pts/hr
Annual Service Demand Mapped to Week of Year

Time Series Plot of Weekly Total
Average Hourly Demand
Mean ED Wait Times by # of Admits

Mean ED DC Wait Time by # ADMITS > 3 hrs / day

# of ADMITS with Order-Dispo Time > 3 hrs / day

Mean ED Wait Time for DC Patients (hrs)
Patient’s Waiting Time vs. Doctor’s Idle Time
Ambulance Diversion – Impact of Delayed Admits
Predictive Modeling

- AutoRegressive Moving Average (ARIMA)
- Exponential Smoothing Models
- Poisson regression Models
  - All assume Gaussian Process
  - But there is obvious interaction in ED Demand
Modeling Inputs

- Data Mining 24 months of patient demand by:
  - DX
  - Patient Characteristics
  - Were patients idled within ED
  - Were staff idled within ED
  - “Feeder” institution data (e.g. County Mental Health discharges)
  - Readmission within 24 hours of discharge
Bending Light: Winters Methods Added to an ARIMA Model

Winters' Method Plot for Number
Additive Method

Variable
- Actual
- Fits
- Forecasts
- 95.0% PI

Smoothing Constants
- Alpha (level) 0.2
- Gamma (trend) 0.2
- Delta (seasonal) 0.2

Accuracy Measures
- MAPE 8.193
- MAD 12.874
- MSD 274.840
When Staffed to Predicted Demand

- Reduction of 24 hour readmits reduced by 32%
- Non-compliance staffing ratio reduced from 16% to 0%
- Diversions reduced from 16% to 3%
- Non-Emergent patients referred to FDQC’s
Field Study 2: Service Demand Modeling and Primary Care Screening Ordering

– It is standard of Care to Provide Diabetic Pt’s routine, scheduled screening tests
  – Eye Exam
  – Ha1C
  – Lipid Screening
  – Foot Exam

▪ National Average is about 73% for any 1 item-less then 45% for all items
▪ Significant impact on Patient Health
▪ Major Pay-for-Performance Measure
The Moving Parts

Finite appointment slots and fixed individual provider times

– Permitting indiscriminate visits with any GIM physician is expensive
  • takes longer
  • May not identify unique characteristics of patient—e.g. antibiotic response

– Requiring Urgent visits with PCP only is not feasible

– GIM exists within a larger Medical Group
  • There is an ED
The Two Common Models
The Moving Parts: Capacity Allocation for Two Stream Demand

Scheduled and Urgent

- Each has different patient requirements for timeliness and continuity
  - Scheduled is Deterministic
  - Urgent is Stochastic
- Each has different cost-structures and revenue
Analyzing Demand-Urgent Care

- Size of provider panel
- Complexity of panel
  - Age
  - Chronic condition
  - Health Expectations
  - Copay structure
- Sub-set Patient demand Characteristics
An Example from an Academic Medical Center in Dan Diego

- **Group 1:** Women < 50 yrs of age with no chronic condition
  - Low demand and mostly for urgent care appointments

- **Group 2:** Anyone over 69 with chronic heart disease
  - High demand - mostly for scheduled appointments with a seasonal spike in
Type and Volume of Visits-2009

Healthy Women < 50

Chronic Illness > 69
Decision Process: Appointment Booking Policies

1) Strategic - Build the Patient-Physician relationship through access
2) Tactical, reserve XX daily slots for open access patients and anticipate future appointments up to the remaining slots
3) Operational level, appointments for daily open access demand given dynamically
Third Model: Partial Flexibility
Chained Scheduling

\[ D_P^1 \quad (a) \quad N_P^1 \quad (b) \]
\[ D_O^1 \quad (c) \quad (e) \quad N - N_P^1 \]
\[ N - N_P^2 \quad \text{Physician 1} \]

\[ D_O^2 \quad (d) \quad (f) \]
\[ N_P^2 \quad \text{Physician 2} \]

\[ D_P^2 \]
Mathematical Modeling

- $N$ is the total daily capacity
- $N_p$, decision variables (aka performance and philosophical constraints)
- $D_p$ demand for scheduled
- $D_o$ demand for urgent care

Creates

- $\min N_p \ c_p E[D_p - N_p]^+ + \ c_o E[D_o - (N - \min(N_p, D_p))]^+$
# Results of Improved Scheduling: Adherence to Policy

<table>
<thead>
<tr>
<th>Activity</th>
<th>Baseline</th>
<th>Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient reports having been reminded of appt</td>
<td>31%</td>
<td>85%</td>
</tr>
<tr>
<td>Contact information verified at check-in</td>
<td>74%</td>
<td>92%</td>
</tr>
<tr>
<td>Patient visit terminated before seeing a Provider</td>
<td>3%</td>
<td>0%</td>
</tr>
<tr>
<td>Patient chart “pulled” by check-in/pre-exam</td>
<td>60%</td>
<td>86%</td>
</tr>
<tr>
<td>Health history taken</td>
<td>26%</td>
<td>86%</td>
</tr>
<tr>
<td>Influenza vaccination rate</td>
<td>43%</td>
<td>68%</td>
</tr>
</tbody>
</table>
Results of Improved Scheduling: Patient Throughput

- Checkin: 1, 3
- Pre-Exam: 5, 6
- PT-Provider time: 28, 24, 16
# Results of Improved Scheduling: Rate of Screenings

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ha1c</td>
<td>67%</td>
<td>93%</td>
</tr>
<tr>
<td>Lipid screening</td>
<td>63%</td>
<td>91%</td>
</tr>
<tr>
<td>Eye Exam</td>
<td>43%</td>
<td>84%</td>
</tr>
<tr>
<td>Foot Exam</td>
<td>37%</td>
<td>89%</td>
</tr>
</tbody>
</table>
Results of Improved Scheduling: Patient Satisfaction

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registration</td>
<td>3.0</td>
<td>4.2</td>
</tr>
<tr>
<td>Waiting Room</td>
<td>3.0</td>
<td>4.4</td>
</tr>
<tr>
<td>Waiting for Provider</td>
<td>3.0</td>
<td>4.2</td>
</tr>
<tr>
<td>Doc</td>
<td>4.0</td>
<td>5.0</td>
</tr>
</tbody>
</table>
Field Study 3: Operational Conditions Affecting Patient Understandings of Treatment

- 13% of Pediatric Transplant “fail” annually
  - failure to comply with treatment recommendations
- Traditional Studies of Non-Compliance Focus on difficult to change factors
  - Patient SES
  - Provider practice type
  - Patient-Provider communication style
- Little analysis of the structural determinants of “understanding”
Input: Uncontrolled Patient Arrival Patterns Determine Wait Times

- Traditionally assume patients arrive randomly (bell shaped curve peaking at appointment time)
  - Each patient’s “timeliness” is independent of the other’s
  - “Timeliness” is within patient control
  - Consistent arrival patterns throughout the day
Tardiness (GM): Johnson $S_U$ fits well

Method: Moment matching
\[ \gamma = -0.3099 \quad \delta = 1.432 \]
\[ \lambda = 24.5789 \quad \xi = -19.02 \]

Goodness-of-fit tests
K-S statistic \[ \chi^2_{\text{GOF}} = 0.1110 \]
Significance probability \[ p = 0.0296 \]

Graphs showing the empirical and theoretical cumulative distribution functions (CDF) and probability density functions (PDF) of patient tardiness for the Johnson $S_U$ distribution.
Change in Patient Understanding of Provider Recommendations

- Pre:
  - 100%: 3%
  - <50%: 7%
  - >50: 90%

- Post:
  - 100%: 28%
  - <50%: 62%
  - >50: 10%
Scheduling Optimization goals

1. Revenue
2. Patient throughput
3. Patient access
4. *Fairness*
Results of Naïve Modeling

- Schedule using *Easy* First Five rule gives 10.8% decrease in total cost over present scheduling system.
- If classification of patients into easy and hard is implemented, it can reduce total cost by more than 10% on average.