Identifying variables associated with physician's clinical cognitive knowledge decay

James H. Crutchfield, Lorie R. Ingraham
Lockheed Martin
Mission Systems and Training
Orlando, Florida
James.H.Crutchfield.Jr@lmco.com,
Lorie.R.Ingraham@lmco.com

Lloyd Werk, Maria Carmen Diaz, Lori K. Handy,
James P. Franciosi, Tim Wysocki
Nemours Children's Health System
Jacksonville, Florida
Lloyd.Werk@, MariaCarmen.Diaz@, Lori.Handy@,
James.Franciosi@, Tim.Wysocki@nemours.org

ABSTRACT

Successful enterprise healthcare management involves making realistic estimates of clinical and support staff performance expectations and assignment of resources, equipment and facilities to meet the anticipated demand for services. A continuous training program is needed to maintain the quality of physician's clinical decision making and to keep staff current with changes in clinical care to meet their performance goals. These training programs require investments of both time and money, which can be disruptive and burdensome, when an extensive training program is delivered indiscriminately across an organization. Many enterprise personnel management systems provide for personalized learning plans, but few organizations customize training schedules and content to the individual experiences and capabilities of physician learners on a continuing basis. A better understanding of individual differences in physician knowledge retention is needed to promote a more efficient and affordable approach to matching training to physician needs.

This study examines de-identified retrospective data from electronic health records to assess physician performance when treating several specific clinical conditions. Diagnosis and management standards were based on clinical practice guidelines and consultation with experts in the field. This enabled the team to validate a generalized forecast model of performance decay based on time since previous encounter for each condition. A measure of individual deviation from the expected forecast model was estimated at each encounter. Individual score divergence was then regressed on variables including characteristics of the physicians that included recent exposure to similar encounters, performance on last encounter, workload, location of service and specialty. Results show how these characteristics were statistically significant ($P<0.0001$) in explaining some of the variation in individual performance. In conclusion, this paper proposes how these variables, which are available from electronic healthcare records and physician credentialing systems, may be used to inform a training management system that can dynamically personalize the training program for the healthcare professional.

ABOUT THE AUTHORS

James Crutchfield, Ph.D. is a Human Factors Design Engineer at Lockheed Martin Mission Systems and Training (MST) in Orlando, FL. He has over 20 years of experience as a social scientist, software requirements analyst, database architect and data analyst in the simulation industry. Recent projects include healthcare quality of care assessment, forecasting and intervention modeling; developing system dynamics models for simulating economic warfare and healthcare intervention strategies; the Cognitive Human Performance Measurement (CHPM) Framework; technical support to the Integrated Training Center simulation model and the Advanced Distributed Learning (ADL) experiment for Lockheed Martin. In a joint project with Nemours, he is developing performance-forecasting models with empirically derived learning retention curves for a US Army research program. Dr. Crutchfield received a Ph.D. in Sociology from Stanford University, with domain concentrations in Social Psychology, the Small Group, and Urban Sociology; a Master of Arts (MA) in Sociology, also from Stanford and his bachelors from Georgia State University.

Lorie R. Ingraham is Principal Investigator of the Decision Support Analytics Lab at Lockheed Martin Mission Systems and Training (MST) in Orlando, FL. She has over 25 years of modeling & simulation (M&S) experience working on complex, real-time software systems, training simulation systems including command & control trainers, virtual and constructive simulators, immersive learning environments, logistics analysis, and enterprise operational analytics. Ms. Ingraham also works on the application of artificial intelligence technologies for operator consoles and semantically-enabled training solutions. Her current responsibilities include design/development of professional preparation learning environments, common enterprise M&S frameworks, mission and business process simulations, and ontology-based technology. Ms. Ingraham received a bachelor's degree in Computer Information Science Engineering from the University of Florida and a Master of Science in Computer Engineering degree from the University of Central Florida.

Lloyd Werk, MD, MPH, FAAP serves as Chief of the Division of General Pediatrics for Nemours Children's Hospital, leads its faculty and programs, and continues to practice pediatric medicine. Within the Nemours Office of Quality and Safety, he works with clinicians and researchers in health services research and quality improvement projects focusing on embedding metrics in the electronic health record, triggering just in time computerized clinical decision support, and providing feedback on clinical performance. Dr. Werk completed his undergraduate degree at Johns Hopkins University in Baltimore Maryland in 1983. He earned a Doctor of Medicine degree from Tel Aviv University in Tel Aviv, Israel in 1987 and a Master of Public Health degree with concentration in Epidemiology & Biostatistics from Boston University in Boston, Massachusetts in 1999. He completed a pediatric residency at Children's Hospital of Buffalo in Buffalo, New York, and practiced primary care pediatrics for several years on Long Island, New York, prior to completing a general academic pediatrics fellowship at Boston University in Boston, Massachusetts. He is an Affiliated Associate Professor of Pediatrics at the University of Central Florida College of Medicine.

Maria Carmen G. Diaz, MD, FAAP, FACEP is a Clinical Professor of Pediatrics and Emergency Medicine at Sidney Kimmel Medical College at Thomas Jefferson University. She is the Medical Director of Simulation for the Nemours Institute for Clinical Excellence and is an Attending Physician in the Division of Emergency Medicine at Nemours/Alfred I. duPont Hospital for Children in Wilmington, Delaware. Dr. Diaz is the author of several peer-reviewed articles and textbook chapters and has lectured nationally and internationally in the fields of Pediatric Emergency Medicine and Simulation.

Lori K. Handy, MD is an Assistant Professor of Pediatrics at the Sidney Kimmel Medical Center at Thomas Jefferson University, and is an attending physician in the Division of Infectious Diseases at Nemours/Alfred I. duPont Hospital for Children where she sees pediatric patients. She is trained as an epidemiologist and focuses her research on improving guideline adherence in the diagnosis and management of pediatric infectious diseases in large healthcare networks. Dr. Handy completed her undergraduate degree at Duke University and Doctor of Medicine degree from the University of Pennsylvania, Perelman School of Medicine. She completed a residency in Pediatrics and a fellowship in pediatric infectious diseases at Children's Hospital of Philadelphia. She is currently completing her Master's Degree of Clinical Epidemiology and Biostatistics at the University of Pennsylvania in Philadelphia, Pennsylvania.

James P. Franciosi, MD, MA, FAAP serves as Chief of the Division of Gastroenterology, Hepatology and Nutrition for Nemours Children's Hospital, leads its faculty and programs in clinical quality and safety efforts, and continues to practice pediatric medicine. He is a clinical researcher focused on clinical outcomes, esophageal conditions, an investigator in the Center for Health Care Delivery Science at Nemours Children's Health System and Co- Investigator for the Nemours site of the PEDSnet Clinical Data Research Network. Dr. Franciosi completed his undergraduate degree at Marquette University and Doctor of Medicine Degree from the Medical College of Wisconsin in Milwaukee, Wisconsin. He completed a pediatric residency and fellowship in pediatric gastroenterology at Children's Hospital of Philadelphia and Master's Degree of Clinical Epidemiology and Biostatistics from the University of Pennsylvania in Philadelphia, Pennsylvania. He is the author of numerous peer-reviewed articles and textbook chapters and has lectured nationally and internationally in the fields of Pediatric Gastroenterology and health outcomes.

Tim Wysocki, Ph.D., A.B.P.P. is Co-Director of the Center for Health Care Delivery Science at Nemours Children's Health System in Jacksonville Florida and Principal Investigator on the research contract that supported this work. He is also Co-Principal Investigator for the Nemours site of the PEDSnet Clinical Data Research Network and on a trial of shared medical decision making for adolescents with type 1 diabetes, both funded by the Patient-Centered Outcomes Research Institute. He also directs an NIH-funded trial of a behavioral intervention targeting enhancement of benefit from continuous glucose monitoring also by adolescents with type 1 diabetes. Dr. Wysocki completed his B.S. and M.A. in Psychology at the University of Texas, his Ph.D. in Psychology at Western Michigan University and a Post-Doctoral Fellowship in Pediatric Psychology at the Kennedy-Krieger Institute at the John Hopkins University School of Medicine. He is Past President of the Society of Pediatric Psychology, on the editorial boards of several journals and a member of the NIH Behavioral Medicine Interventions and Outcomes Study Section.

Identifying variables associated with physician's clinical cognitive knowledge decay

James H. Crutchfield, Lorie R. Ingraham
Lockheed Martin
Mission Systems and Training
Orlando, Florida
James.H.Crutchfield.Jr@lmco.com,
Lorie.R.Ingraham@lmco.com

Lloyd Werk, Maria Carmen Diaz, Lori K. Handy,
James P. Franciosi, Tim Wysocki
Nemours Children's Health System
Jacksonville, Florida
Lloyd.Werk@, MariaCarmen.Diaz@, Lori.Handy@,
James.Franciosi@, Tim.Wysocki@nemours.org

INTRODUCTION

Medical science continues to make remarkable progress but remains limited by the knowledge-capacity of those who apply it. Faced with the demands of caring for larger panels of patients and the exponential growth in medical knowledge, physicians are challenged to retain and remain abreast with the continuous advances in care. Lapses in knowledge can have severe consequences as patients may receive substandard care. Recognizing that human error is inevitable, Makary and Daniel (Makary & Daniel, 2016) recommend a three-way approach that includes "making errors less frequent by following principles that take human limitations into account" (p.2). These have been implemented through institutional methods of crosschecking, accountability and continuous training.

Meeting these needs is not without cost, so that successful enterprise healthcare management requires making realistic estimates of professional and support staff performance expectations and implementing the necessary crosschecking, accountability and continuous training to obtain quality outcomes. Maintaining the quality of physicians' clinical decision making requires a continuous training program to keep staff capable and able to meet their performance goals. Implementation of a systematically applied knowledge retention program could address physicians' knowledge degradation as well as the requirement to continuously re-evaluate and update their training and education programs. However, this type of training approach can be very inefficient if delivered indiscriminately across an organization. Although many enterprise personnel management systems provide for personalized learning plans, few organizations customize training content to the individual experiences and capabilities of physician learners on a continuing basis or on-demand. An understanding of individual differences in physician knowledge retention is needed to promote a more efficient approach to matching education and training programs to physician needs.

This study uses de-identified electronic health records (EHR) to demonstrate how data recorded during patient encounters can be used to assess physician performance characteristics when clinicians treat several specific clinical conditions. These conditions include pediatric obesity, gastroesophageal reflux disease (GERD) and supracondylar fractures of the elbow. Using diagnosis and management standards based on clinical practice guidelines and consultation with experts in the field, measures of required clinical knowledge were developed that could be evaluated through examination of the EHR. Using these measures, a generalized forecast model of knowledge decay based on time since previous encounter for each condition was derived. It was expected that measures of individual deviation from the forecast model might be explained through characteristics of the physician or the environment. An analysis of some of the available individual physician data such as those related to workload and location was performed.

The purpose of this study is to identify relevant data, available from the EHR and other electronic sources, and develop an individualized model of declarative knowledge decay. This information will be used to create a more personalized learning and job support program geared to meet specific physician needs.

BACKGROUND

Ebbinghaus (Ebbinghaus, 1999), in 1885, proposed an initial mathematical model for learning and knowledge decay.

$$R = e^{-\frac{t}{s}} \qquad (1)$$

Where R is memory retention and s is the relative strength of the memory and t is decay time. Ebbinghaus tested the approach by attempting to memorize a sequence of random syllables, then recording the number of errors after different time periods. Although this research was incomplete and tested only on himself, the resulting retention curve is an effective first attempt at a declarative memory retention forecasting model and continues to be referenced in the study of memory.

In subsequent research, several mathematical models of knowledge retention have been proposed as shown in Figure 1. It was determined that, from a theoretical standpoint, an exponential model would best suit this application. We assume that performance of a task will reverse the effects of decay and represent time between opportunities to perform as delta-t (Δt). Our approach slightly extends the Ebbinghaus model to a Δt skill decay model by adding a parameter defining a range for skill decay, a minimum asymptote for performance, and a shape factor that can be determined to include a vector of attributes of the work, worplace and worker. In each case, time has been used as a proxy for some unseen process of memory decay. Individual, knowledge/task, and environmental characteristics can influence the model.

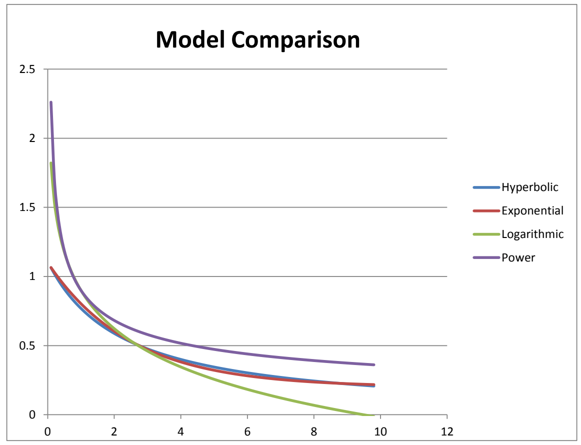

Figure 1. Comparison of alternative knowledge decay models

Theory of Memory Decay

When looking at the process of memory decay itself, it is helpful to identify the process by which memories or skills are lost. Predominant theory suggests that memory failure appears when some event interferes with the recovery process. Levy, Kuhl and Wagner identify five mechanisms of forgetting (Levy, Kuhl, & Wagner, 2010):

1. Forgetting due to failed encoding
2. Forgetting due to disrupted consolidation
3. Forgetting due to retrieval competition
4. Forgetting as a consequence of resolving competition
5. Forgetting due to ineffective retrieval cues

The first two have to do with how memories are acquired and stored and are beyond the scope of this analysis. The latter three are relevant because elements that compete with retrieval and cause competition resolution are observable. For example, retrieval competition refers to when retrieval cues are related to multiple associated memories. The resolution of retrieval competition similarly impedes retrieval as a result of forcing competing memories out of salience, and they remain so even when they become relevant. High volume physician practices result in greater retrieval competition and its effects. These conditions may also reduce the effectiveness of retrieval cues through any number of distracting elements.

Experts have also theorized that memory loss may be the result of a more direct adaptation whereby response sets are activated by specific cues, resulting in the deactivation of others. This provides greater speed and accuracy with the selected response set as might be needed by the situation. As a consequence, other memory traces are deactivated and pushed down in salience. Thus, repeated use of a response set will assist in retaining the memory, whereas lack of use causes the memory to be pushed down out of salience giving the appearance of decay.

This memory loss adaptation is thought to enhance performance because it prevents “rogue control information hijacking behavior” (Altman & Gray, 2008, p. 605) by gradually inhibiting actions that have not been taken for some time. Without this memory loss mechanism, older “rogue” knowledge could suppress the acquisition of new knowledge. However, as physicians see clinical conditions at a rate that can best be described as variations on a Poisson process, their ability to diagnose and treat some less frequent clinical problems can be necessarily degraded through memory loss. Consequently, a memory loss adaptation impedes the complete exercise of some varied and complex tasks.

In recent years, a number of researchers have approached the problem with the intent of understanding memory and retention, usually with regard to the training method used to create the memory. The Army Research Institute (ARI) developed a performance prediction model based on ten characteristics of the work ability being retained (Radtke &

Shettel, 1985). The approach did not take into account characteristics of the individual worker or the workplace situation where work is being performed or not performed. Moreover, it found application on a limited set of military tasks.

APPROACH

This research follows the method of using time as a proxy for the unseen process of forgetting, recognizing that this study design will not determine the psychological processes of memory or determine training for maximum retention. Our approach takes a *post hoc* applied approach to determining some of the characteristics that can impact retention. Taking the research outside of experimental data, we use existing EHR data to estimate knowledge decay curves and rate performers on their expression of declarative knowledge.

Day et al. indicate that "it is imperative that the search for individual differences extend beyond assessments of skill at the conclusion of training" (Day, et al., 2013, p. 283). This research assumes three classes of characteristics that may impact knowledge retention are (1) the work and how it is performed, (2) the worker, whose latent abilities and personal experience impact retention, and (3) the workplace environment, with distractions, collaborative interaction and the availability (or lack thereof) of passive and active job aids. This approach is intended to develop a system for customizing a performance prediction model based on those characteristics that are available and useful in competence sustainment.

To validate the approach, our research team applied it to three clinical conditions, pediatric obesity, gastroesophageal reflux disease (GERD) and supracondylar fractures of the elbow. The study population includes primary care pediatricians in a variety of clinical and hospital settings. The study of supracondylar fractures specifically looks at physicians employed in emergency departments and urgent care centers. The three clinical conditions studied are highlighted in the following sections.

Pediatric Obesity

Pediatric obesity has become of increasing interest due to its impact on individual physical development, psychological development, and long-term health outcomes of the patient (Freedman, 2005) (Shashaj, 2014) (Abdul-Ghani M., 2006) (Wang W., 2011). Best practices conveyed in clinical practice guidelines encourage doctors to educate parents and patients of the health risks of obesity as well as provide counseling on proper nutrition and the necessity for exercise. Among children identified as overweight or obese, specific history and physical examination elements are elicited and interventions around laboratory testing, counseling, consultation and follow-up are advised (Institute, 2012) (Barlow, 2007) (Spear B.A., 2007).

Gastroesophageal Reflux Disease

Gastroesophageal reflux disease (GERD) is a condition in which reflux of gastric contents cause troublesome symptoms and/or complications that may even be life threatening. In contrast, Gastroesophageal reflux (GER) is a normal physiologic process involving passage of gastric contents into the esophagus with or without regurgitation and vomiting. It is not uncommon for infants to experience periods of spitting up as the infant learns the muscular sequences required for swallowing and belching. Parents may be dismayed by the observed behaviors and expect the physician to react by prescribing acid control medications or possibly other tests. Most often, these responses are generally unnecessary. Instruction on proper positioning of the infant, feeding frequency, and formula thickening may be all that are required. Despite dissemination of best practices of care, there remains variation in the decision making around diagnosis and treatment among clinicians (Vandenplas Y., 2009) (Lightdale, 2013). This results in the over prescription of acid suppression medication, particularly evident among primary care providers, where most GER patients are diagnosed and treated.

Supracondylar Fracture

Supracondylar fractures are the most common elbow fractures in children and represent 3 % of all pediatric fractures with an estimated annual incidence of 177.3 per 100,000 children (Houshian, Mehdi, & Larsen, 2001). Emergency departments (ED) and urgent care (UC) staff are the first to assess the degree of the injury and pass on critical information to the orthopedists such as providing an initial assessment of the injury and any associated complications.

This requires an evaluation of the site of injury, accurate examination of the radial, median and ulnar nerve motor and sensory functions, and a thorough evaluation of the relevant vascular systems. The physician must know what to expect and perform these evaluations with accuracy and confidence, placing the patient more at ease and minimizing the pain these examinations may introduce. Yet, there is often a lack of timely documentation of these examinations and information of value for the orthopedist to recommend a treatment plan.

Applying a Decay Model

The knowledge decay model represented by Equation 2 extends the Ebbinghaus model by adding a parameter defining a range for skill decay, a minimum asymptote for performance, and a shape factor that can be determined to include a vector of attributes of the work, workplace and worker:

$$K(\Delta t) = \psi + \rho e^{-\gamma \Delta t} \quad (2)$$

Where:

Δt is the time since last encounter with the condition,
$K(\Delta t)$ is the estimated or forecast level of knowledge at time Δt,
ψ is the minimum level of knowledge below which the practitioner will not descend,
ρ is the range the knowledge will vary, or how much it can be expected to descend, and
γ is the shape factor for the decay function.

Note that the Ebbinghaus model was a standardized form of this equation where ρ equals one (1) and ψ equals zero.

The key problem is the estimation of the level of knowledge "K". Since knowledge is understood as a collection of memory traces, and a complex task involves the application of a range of knowledge, the metric representing "K" should be a function of an array of observable elements.

For each clinical condition studied, an array of observable elements, which are elements searchable in the EHR, are identified. They are combined into a knowledge utility function (KUF) described below, for each of the three conditions. The approach taken here employs review of archival EHR data, rather than directly accessing the physicians and applying a knowledge test. Such a test would be expensive to develop and validate, require exquisite discriminant validity to possess adequate sensitivity, and its re-administration with the same physicians would be fraught with interpretive and logistic problems. Consequently, indicators of key knowledge as expressed in the EHR are likely to function as well, or better, than attempts to measure physician knowledge directly via examinations and repeatedly over time.

METHOD

The study used specific methods for specifying indicators and the KUF, formulating the model and evaluating individualized performance indicators. Patient records were de-identified prior to analysis and abstracted for key knowledge utility function elements. Similarly, physician identities were also de-identified, hampering the present research to a limited degree, but providing necessary security for their practice and privacy. This research was approved and monitored by the appropriate institutional review boards.

Knowledge Utility Function (KUF)

The development of the KUF followed a process with six stages. These stages are:

- Survey a panel of physicians to select suitable targeted clinical problems (TCP).
- Interview subject matter experts (SME) on appropriate TCP diagnosis, assessment, treatments and other concerns for the clinical problems.
- Analyze the interview transcripts to identify and categorize candidate elements of declarative knowledge.
- Develop measurable indicators for the presence of the expression of candidate elements of declarative knowledge in the EHR.

- Select specific indicators for time-based analysis.
- Apply algorithms and scale techniques to generate KUF scores.

The development stages are further described in the following sections.

1. Survey a panel of physicians to select targeted clinical problems (TCP)
The panel consisted of the diverse group of practicing physicians on the research team. They included pediatricians working in research, general and subspecialty pediatrics, ED and hospital settings. The panel initially selected nine conditions with three reported in this paper: obesity, GERD and supracondylar fractures.

2. Interview subject matter experts (SME)
Two subject matter experts were individually consulted, providing narrative information on existing clinical practice guidelines and best practices and any perceived gaps in decision making during diagnosis, evaluation and treatment.

3. Analyze the interview transcripts
Transcripts were analyzed to identify candidate elements of declarative knowledge. Clinical practice elements were identified, categorized and listed in a spreadsheet. Physician SMEs were asked to review the spreadsheets to edit and down-select to elements that are key indicators.

4. Develop measurable indicators
To develop indicators for the presence of the expression of candidate elements of declarative knowledge in the EHR, data analysts constructed queries to identify declarative knowledge in the selected EHR records. The study set for obesity and GERD focused on patient encounters in 2013. The data for elbow fractures was selected from 2015 records due to the availability of medical records and sequence in study submissions. The resulting variables indicated the presence or absence of selected elements. Missing values were interpreted as the absence of required elements.

5. Select specific indicators for analysis
Some elements had limited expression in the data, with less than ten occurrences. This is likely due to a gap between the accepted practice of the clinicians and the expectations of experts. The unit of analysis was the clinical encounter. The number of observations obtained for the three clinical conditions, obesity, GERD and supracondylar fracture were 20,264; 4,530; and 192, respectively.

6. Apply scale techniques to generate KUF scores
A correlation analysis was used to identify the inter-item correlation and estimate a Cronbach's alpha coefficients. This is the estimate of measure of how closely related a set of items are as a group and a measure of scale reliability. For obesity, GERD and supracondylar fractures the standardized alphas were 0.75, 0.29 and 0.76, respectively. Principal components analysis (PCA) was used to identify weights to the individual elements which enabled them to be summed to a KUF score. Score results for the three conditions are interpreted as follows:

- **KUF Elements for Obesity**
  The practice elements identified for obesity included the presence of notations about obesity, as well as dietary and physical activity counseling in encounter notes.

- **KUF Elements for GERD**
  The practice elements identified for GERD include the presence of notations on spitting up, notes about feeding refusal and arching, notes about choking and gaging and whether or not the patient was asked about projectile vomiting.

- **KUF Elements for Supracondylar Fracture**
  The practice elements identified for elbow fracture included notes on the vascular examination, presence of pulses, and an evaluation radial, median and ulnar nerve motor and sensory function.

Formulating the Model

The time between encounters with the targeted condition is assigned a range identifier named Δt, which is defined regardless of whether the encounters are for the same or different patients. As previously mentioned, principal components analysis was used to construct KUF scores for the individual encounters for each of the clinical conditions. The explained proportion of variance for the correlation matrix for the first component of obesity was 0.58. The resulting eigenvector for key practice elements "nutrition notes", "obesity notes", "exercise notes" and "blood pressure" were 0.55, 0.54, 0.47 and 0.43, respectively. The explained proportion of the correlation matrix for the first component of GERD was 0.32. The resulting eigenvector for "spitting up", "arching", "projectile vomiting", and "choke gagging" were 0.63, 0.56, 0.45 and 0.31, respectively. The explained proportion of the correlation matrix for the first component of elbow supracondylar fracture was 0.61. The resulting eigenvector for "Vascular Exam Normal", "Ulnar Motor Norm", "Ulnar Senses Norm", "Radial Senses Norm" and "Radial Pulse present" were 0.12, 0.57, 0.57, 0.57 and 0.10, respectively.

When plotted against Δt, it was observed that KUF variation was substantial at all values of Δt. Yet for larger values of Δt, the values clustered progressively lower. This indicated that, though variation in performance occurs at all delay periods, we know that some of this variation was due to decay, and some was due to other factors leading to nonadherence to the reporting standards recommended by our SMEs.

To obtain an estimate of decay independent from adherence issues, it was observed that since the highest values decreased for the values of larger Δt, these were likely due to decay. The upper fifth percentile KUF scored encounters at each Δt were selected out for model development. This approach allows us to keep the case count distribution resulting from the patient arrival rates while evaluating primarily encounters where adherence was attempted. A non-linear equation using the formula described above was fit, for each condition, using maximum likelihood estimation. This provided an expected knowledge decay curve based on our more adherent providers. The KUF residuals, computed as differences from the expected curve, for all providers, not just the upper 5th percentile, may then be regressed against the individual factors.

Evaluating Individualized Performance Indicators

The resulting performance curve characterizes the expected score distribution for the upper fifth percentile performers at each Δt. To establish a score against which individual indicators could be evaluated, the residual difference between the expected high score and actual scores was computed for each encounter. Referencing the performance model of Equation 2, it is possible solve for γ by using the standardized form of the equation, where ψ is set to zero, and ρ is set to one. Solving for γ (the shape function), we find it is equal to the log of the KUF score divided by negative Δt.

The definition of γ describes a vector of individual performance indicators, so we can solve for the factors on these indicators using linear regression. The resulting weights are useful in constructing individualized performance forecasts as they relate to the decay curve. However, γ is computationally a very low value, so parameters to estimate its variation would also be very low values. A simpler and more intuitive result would be reached by regressing the KUF performance residuals (difference), on the individual factors directly. This approach was used in our study.

Deriving Individualized Performance Factors

The indicators that were derived include, (a) previous KUF score ("PrevKUF"), (b) annual number of cases by provider ("ProvID_Cases"), (c) annual number of cases by department ("Dept_Cases"), (d) number of providers in department ("Dept_ProvIDs"), (e) whether the department is in a hospital ("Hosp_Setting"), (f) whether the provider is male or female ("ProvFemale"), (g) the title of provider Medicine degree ("MD") vs. Doctor in Osteopathic Medicine degree ("Title_DO"), (h) whether encounter occurred in the office ("Office_Visit"), rather than by phone, (i) annual number of encounters of all kind by provider ("All_Encs") and (j) annual number of specific cases by provider ("CasesPerProvID").

Desired information about training, other than "MD" vs. "DO", was unavailable as was information on the cognitive ability of the physician, age, and training test scores. These data points could be collected through interview or other survey instrumentation, but that was beyond the scope of this phase of the project.

RESULTS

Decay Curve Estimates

Primary results from the study consisted of plotting and correlating the knowledge decay curves for each clinical condition. These results are described in the following sections.

1) Obesity

The decay curve for obesity was estimated using data from the upper 5th percentile KUF scores, for each value of Δt, representing time in days since previous encounter. That resulted in 2053 observations. It produced a significant fit (P< 0.0001) with ψ = -2.99, ρ = 6.22 and γ = 0.00582. Visually one can see in Figure 2 that the curve is not as good a fit as one would expect, due to the high level of knowledge expression variation found in the EHR. The mean residual difference from the curve for all 20,264 encounters [Actual – Curve Estimate] was -3.19 with a minimum of -4.62 and a maximum of 2.49, with a standard deviation of 1.54.

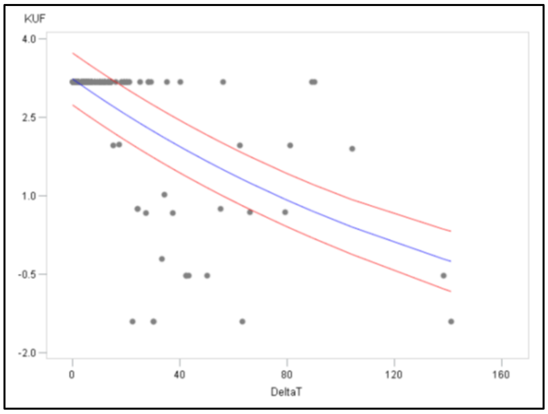

Figure 2 Nonlinear curve estimate for Obesity KUF with respect to Δt. Red lines indicate 95% forecast lines.

The larger cluster of points around Δt =0 indicates the high rate of patients arriving with BMI greater than the 95th percentile. We see that some physicians continued to raise the obesity education points with patients and parents, recording that they did so as late as three months after the previous encounter. Some failed to do so at even low values of Δt.

2) GERD

The decay curve for GERD was also estimated from the upper 5th percentile, for each value of Δt. That resulted in 432 observations. It produced a significant fit (P<0.0001) with ψ = -0.54, ρ = 4.09 and γ= 0.05. As seen in Figure 3, the curve visually presents a much better fit to the observed data. The mean residual difference from the curve for all 4,530 encounters [actual – curve estimate] was -3.06 with a minimum of -3.70 and a maximum of 7.87, with a standard deviation of 1.22.

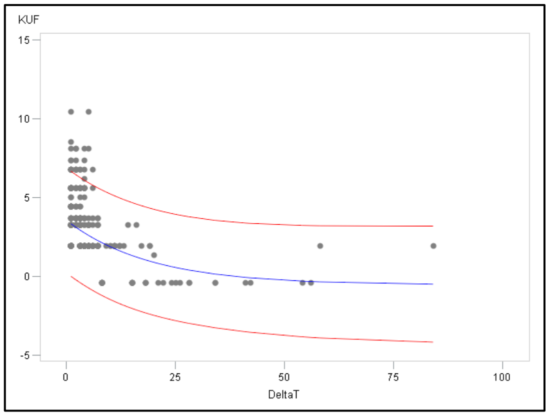

Figure 3. Nonlinear curve estimate for GERD KUF with respect to Δt. Red lines indicate 95% forecast lines.

Even among the upper five percent performers, we still see considerable range of performance at low values of Δt for GERD as was seen with obesity.

3) Supracondylar Fracture

The decay curve for supracondylar fracture was estimated from the upper 5th percentile, for each value of Δt. That resulted in 26 observations used for the curve estimate. It produced a significant equation (P<0.0001) as well, with ψ = -0.37, ρ = 0.70 and γ= 0.12. The curve appears to fit the data as shown in Figure 4. The average residual difference from the curve for all 192 encounters was -3.06 with a minimum of -0.58 and a maximum of 23.98, with a standard deviation of 1.77.

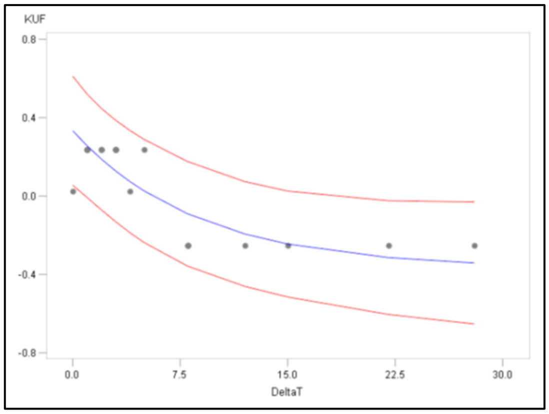

Figure 4. Nonlinear curve estimate for Supracondylar Fracture KUF with respect to Δt. Red lines indicate 95% forecast lines.

Individual Performance Characteristics

1) Obesity

A fit of the individual differences from the forecast curve resulted in a significant linear equation, see Table 1. The strongest indicator of knowledge expression at 0.22 with a positive relation to the performance difference was previous KUF. That is, the better one scored on the previous encounter, the better they scored on the present encounter.

Female providers tended to score better than males (.14) and a hospital setting tended to lower the score (-.46). That is, those encounters occurring during hospital visits tended to discuss other things than obesity.

Surprisingly, this analysis indicates a very low relationship between the number of annual encounters ("All_Enc") and the KUF score. This would indicate that when physicians do decide to discuss obesity with their patients, they are not seriously influenced by their workload.

Table 1. Linear regression analysis of factors influencing variation from the expected knowledge decay curve for obesity.

Parameter Estimates						
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standardized Estimate
Intercept	1	-3.67192	0.03734	-98.33	<.0001	0
PrevKUF	1	0.21916	0.00694	31.59	<.0001	0.21693
ProvID_Cases	1	-0.00047	9E-05	-5.2	<.0001	-0.06658
Dept_Cases	1	0.00022	2.62E-05	8.41	<.0001	0.13744
Dept_ProvIDs	1	0.00989	0.00129	7.65	<.0001	0.0888
Office_Visit	1	6.46E-05	1.92E-05	3.36	0.0008	0.05474
Hosp_Setting	1	-0.46303	0.07917	-5.85	<.0001	-0.04109
Title_DO	1	-0.19964	0.03939	-5.07	<.0001	-0.03759
ProvFemale	1	0.13903	0.02287	6.08	<.0001	0.04506
All_Encs	1	1.8E-05	6.15E-06	2.93	0.0034	0.04526
CasesPerProvID	1	-0.00062	0.000153	-4.07	<.0001	-0.07313

2) GERD

A fit of the individual differences from the forecast curve resulted in a significant equation (P<0.0001) as well, see Table 2. However of the individual parameter estimates only three were significant above the .05 level. These were Previous KUF, Number of department providers, and number of encounters of all types by the physician. The Previous KUF operated as expected, where larger previous values resulted in subsequent higher scores (0.04). The number of department providers increased performance by about 0.03. All of the GERD visits were office visits and did not occur in the hospital setting, so these factors were omitted from the equation.

Table 2. Linear regression analysis of factors influencing variation from the expected knowledge decay curve for GERD.

Parameter Estimates						
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standardized Estimate
Intercept	1	-2.642	0.10908	-24.22	<.0001	0
PrevKUF	1	0.04016	0.01892	2.12	0.0338	0.03602
ProvID_Cases	1	0.000283	0.000341	0.83	0.407	0.02232
Dept_Cases	1	-0.00021	0.000115	-1.82	0.0696	-0.06776
Dept_ProvIDs	1	0.03246	0.01197	2.71	0.0067	0.08233
Title_DO	1	0.08911	0.06226	1.43	0.1524	0.02644
ProvFemale	1	-0.07287	0.05395	-1.35	0.1769	-0.0297
All_Encs	1	-7.7E-05	9.31E-06	-8.3	<.0001	-0.22826
CasesPerProvID	1	0.00122	0.000751	1.63	0.1035	0.06901

3) Supracondylar Fracture

Likewise, for Supracondylar Fracture, the fit of the individual differences from the forecast curve resulted in a significant equation (P<0.0001), see Table 3. However, none of the indicator parameters were found to be significant in the equation. By itself, "Hosp_Setting" appeared to be as the simple regression estimate was 0.32 (P<0.0001). These encounters occurred exclusively in either urgent care facilities or emergency rooms. We see a tendency for emergency rooms to record more circulatory and neurological information in the EHR. Since data were taken from a later time period, no data for "All_Encs" was available.

Table 3. Linear regression analysis of factors influencing variation from the expected knowledge decay curve for supracondylar fracture.

Parameter Estimates						
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standardized Estimate
Intercept	1	-0.57225	0.12095	-4.73	<.0001	0
PrevKUF	1	0.07188	0.08467	0.85	0.3972	0.05425
ProvID_Cases	1	-0.00094915	0.00246	-0.39	0.6996	-0.03669
Dept_Cases	1	-0.00017467	0.00108	-0.16	0.8721	-0.12483
Dept_ProvIDs	1	0.00777	0.00709	1.1	0.2742	0.54408
CasesPerProvID	1	-0.00669	0.03792	-0.18	0.8602	-0.09488
Title_DO	1	-0.05012	0.03563	-1.41	0.1615	-0.09813
ProvFemale	1	-0.00483	0.03026	-0.16	0.8733	-0.01168
Hosp_Setting	1	0.16487	0.11942	1.38	0.1693	0.33572

SUMMARY OF RESULTS

In general, from these studies, we observe in two cases that past performance influences present performance. Furthermore, we also observed, except for the elbow fractures, that the number of providers in the department appeared to influence either the knowledge used in the encounters, or improves overall reporting on the EHR. This indicates the presence of a measureable peer influence effect in terms of the expression of declarative knowledge.

Gender differences were observed, but were not consistent across clinical conditions. The setting of the encounter would also appear to influence the outcome, at least for obesity. Encounters occurring in hospitals saw fewer references to obesity education. In contrast, since supracondylar fracture encounters were observed in urgent care and hospital settings, we noted that hospital settings appeared to improve reporting for this condition.

More generally, factors that did not significantly affect the residual differences from the expected decay curve were the case load indicators, “ProvID_Cases” and “Dept_Cases”. These indicated the total number of similar encounters with the condition over the year and this neither improved nor degraded the scores. However, when total workload data were available, including encounters of all clinical types, we did see a statistically discernable though very small relationship (obesity, 1.9E-05 and GERD, -7.7E-05). We are consequently unable to assume a relationship between workload and declarative knowledge expression in the recording of EHR information.

DISCUSSION

The objective of this study was to develop a method for forecasting individual knowledge decay on specific healthcare issues using the EHR. A major concern with regard to healthcare professionals was identified in that, as professionals, physicians vary substantially in how they conduct their clinical practice. Additionally, the EHR, especially in 2013, the data accessed for GERD and Obesity in this study, was a relatively new addition to the substantial task requirements faced by physicians. There may be some residual resistance to fully utilizing the system, as it may not be perceived as directly related to the provision of healthcare to their patients. The number of co-providers seemed to improve reporting somewhat, which explained some individual variation. The number of peers may provide greater access to assistance in reporting as well as a greater perception of the need to use the EHR to communicate with the other staff.

Results for supracondylar fracture, taken from 2015 data, were notable in that the facility setting, hospital or not, was an important indicator. The complex environment of the hospital may create a greater perception of the need to use the EHR to communicate patient clinical information.

This paper covers early phases of the research in a retrospective analysis of EHR data. It extends current research on knowledge retention through application of laboratory theory to a working institution. Continuing research includes a validation of the approach by providing interventions based on time since previous encounter and monitoring changes in declarative knowledge expression in the EHR. Using a field experiment design, various intervention approaches are being applied to volunteer care provider participants to mitigate the observed knowledge decay phenomena. These interventions include didactic and simulation based training as well as computerized clinical decision support tools within the EHR. This ongoing research further expands the list of individual performance characteristics as applied in the forecast model, by obtaining, from study participants, data on the participant that may be relevant to forecasting knowledge retention. These data were not available in the retrospective study and include information such as age of participant, years since fellowship and their internship. Initial results are promising and are undergoing further analysis.

The EHR is an increasing and permanent presence in the practice of medicine. The importance of complete representation in the EHR of the declarative knowledge used in decision making with regard to the patient is made particularly relevant when healthcare provision is recognized as a team effort, and the EHR is a communication tool. As it becomes more broadly shared in the healthcare community, the approach presented here offers value in a) identifying areas of concern for improving healthcare, and b) the construction of forecasting tools to direct scheduling of refreshment training or the presentation of job aids. A more affordable intervention schedule and higher quality of care are the anticipated outcomes.

ACKNOWLEDGEMENT

This research and development project was conducted by the Nemours Foundation and is made possible by a cooperative agreement that was awarded and administered by the U.S. Army Medical Research & Materiel Command, the Joint Program Committee 1, Medical Simulation and Information Sciences and the Congressionally Directed Medical Research Programs Office, at Fort Detrick, Md under award Number: W81XWH-13-1-0316.

The views, opinions and/or findings contained in this presentation are those of the author(s) and do not necessarily reflect the views of the Department of Defense and should not be construed as an official DoD/Army position, policy or decision unless so designated by other documentation. No official endorsement should be made.

REFERENCES

Abdul-Ghani M., T. D. (2006, May). Contributions of beta Cell Dysfunction and Insulin Resistance to the Pathogenesis of Impaired Glucose tolerance and Impaired Fasting Glucose. *Diabetes Care. Volume 29, Number 5, May Publication*.

Altman, E. M., & Gray, W. D. (2008). An Integrated Model of Cognitive Control ini Task Switching. *Psychological Review*, 602-639.

Barlow, S. (2007). Expert Committee recommendations regarding the prevention and assessment and treatment of child and adolescent overweight and obesity summary report. *Pediatrics 120(suppl4)* , 164-195.

Day, A. E., Arthur, J. W., Villado, A. J., Boatman, P. R., Kowollik, V., Bhupatkar, A., & Bennett, J. W. (2013). Relating individual differences in ability, personality, and motivation to the retention and transfer of skill on a complex comand-and-control simulaiton task. In W. Arthur, Jr, E. A. Day, W. Bennett, Jr, & A. M. Portrey, *Individual and Team Skill Decay* (pp. 282-301). New York: Routledge.

Ebbinghaus, H. (1999). *Memory. A Contribution to Experimental Psychology.* Bristol: Thoemmes Press.

Freedman, D. e. (2005). The relation of childhood BMI to adult adiposity: the Bogalusa Heart Study. *Pediatrics 115(1)*, 22-27.

Houshian, S., Mehdi, B., & Larsen, M. S. (2001). The epidemiology of elbow fracture in children: analysis of 355 fractures, with special reference to supracondylar humerus fractures. *Journal of Orthopedic Science*, 312-315.

Institute, N. H. (2012, October). Expert Panel on Integrated Guidelines for Cardiovascular Health and Risk Reduction in Children and Adolescents. *US Department of Health and Human Services, National Institutes of Health, NIH Publication NO.12-7486*.

Levy, B. J., Kuhl, B. A., & Wagner, A. D. (2010). The functional neuroimaging of forgetting. In S. D. Sala, *Forgetting* (pp. 135-163). Hove, UK: Psychology Press.

Lightdale, J. G. (2013). Gastroenterology, Hepatology, and Nutrition, Gastroesophageal Reflux: Management Guidance for the Pediatrician. *Pediatrics 131*, e1684–e1695.

Makary, M. A., & Daniel, M. (2016, May 3). *Medical error—the third leading cause of death in the US*. Retrieved from The BMJ: http://www.bmj.com/content/353/bmj.i2139

Radtke, P. H., & Shettel, H. H. (1985). *Training Program for Predicting Military Task Retention Evaluation Report.* Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Shashaj, B. e. (2014, August 11). Origin of Cardiovascular Risk in Overweight Preschool Children. A Cohort Study of Cardiometabolic Risk Factors at the Onset of Obesity. *JAMA Pediatrics*, p. 2014.

Spear B.A., B. S. (2007). Recommendations for treatment of child and adolescent overweight and obesity. *Pediatrics 120 (Suppl 4)*, 254-288.

Vandenplas Y., R. C. (2009). Pediatric Gastroesophageal Reflux Clinical Practice Guidelines: Joint Recommendations of the North American Society for Pediatric Gastroenterology, Hepatology, and Nutrition (NASPGHAN) and the ESPGHAN. *Journal of Pediatric Gastroenterology and Nutrition 49*, 498-547.

Wang W., L. E. (2011, February). Fasting Plasma Glucose and Hemoglobin A1c in Identifying and Predicting Diabetes. *Diabetes Care. Volume 34, February Publication*.