Statistical Applications in Knowledge Translation Research Implemented Through the Information Assessment Method

Jonathan L. Moscovici

Master Of Science

Department of Mathematics and Statistics

McGill University

Montreal,Quebec

2012-11-01

A thesis submitted to McGill University in partial fulfilment of the requirements of the degree of Master of Science

DEDICATION

This document is dedicated to the loving memories of Solomon Moscovici, and Ştefania & Constantin Mitran.

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my supervisor, Dr. Alain C. Vandal, for several reasons. These include, but are not limited to, allowing me the rare opportunity to pursue something I enjoy, even through following the beat of a different drum, and for perceiving me as an individual, not a number (an uncommon event in a math department, pun intended). Success that I may encounter will be your reward for allowing me the opportuniy. Secondly, I extend many, many thanks to the ITPCRG research team (notably my co-supervisor Dr. Roland Grad) for exposing me to such interesting research topics ultimately leading to this thesis. I find your discipline, productivity and good nature to be inspiring. Next, I give thanks to all the professors in the department who have imparted wisdom upon me (make no mistakes, it was all useful). These include Dr. David Wolfson, Dr. Russell Steele, Dr. Jose Correa, Dr. Robert Vermes and Dr. David Stephens, with special mention to Rafaella Bruno, Spencer Keys-Schatia and Gregory LeBaron (for helping me out and putting up with me). Lastly, I thank my parents, grandmother, brother, girlfriend, and friends for providing support and laughter without which this would have been improbable.

ABSTRACT

Of interest are two knowledge translation [27] research projects conducted by and with the ITPCRG (Information Technology Primary Care Research Group) during the period 2010-2012, as well as their underlying statistical analyses. For physicians, continuing medical education (CME) is a critical activity that helps them acquire new knowledge and keep their practice up to date. In Canada, popular CME programs are structured around the reading of short synopses or summaries of important clinical research on e-mail. After reading one synopsis, the physician completes a short reflective exercise, using the Information Assessment Method (IAM). IAM is a brief questionnaire that asks physicians to reflect on the following: -The relevance of the information? -The impact of the information e.g. did you learn something new? -If they intend to use the information for a specific patient? -Whether they expect to see health benefits for that patient as a result? This type of CME is very popular. Since September 2006, about 4,500 members of the Canadian Medical Association have submitted more than one million IAM questionnaires linked to e-mailed synopses. Previous work suggests the response format of the IAM questionnaire can impact the willingness of physicians to participate, and that information use for a specific patient might be linked to certain factors measurable by IAM. Therefore, the objectives were to improve CME programs that use the IAM questionnaire by determining which response formats optimize physician participation and their reflective learning, and explore the determinants of information use. These were accomplished by implementing a survival analysis framework, as well as mixed logistic regression models.

ABRÉGÉ

Ce mémoire porte sur deux projets de mise en pratique des connaissances menés par et avec le ITPCRG (Information Technology Primary Care Research Group) de 2010 à 2012, ainsi que l'analyse statistique qui s'en est issue. La formation médicale continue est une activité essentielle qui aide l'acquisition de nouvelles connaissances et la mise à jour des pratiques pour les médecins. Au Canada, des programmes populaires utilisent la lecture de courts synopsis ou de sommaires de recherches cliniques importantes transmis par courriel. Après la lecture du synopsis, le médecin complète un bref exercice de réflexion en utilisant le Information Assessment Method (IAM). IAM est un petit questionnaire qui demande aux médecins de réfléchir aux sujets qui suivent: -La pertinence de l'information? - L'impacte de cette information ex : avez-vous appris quelque chose? -L'intention d'utiliser cette information pour un patient spécifique? - Anticipent-ils observer des bénéfices de santé pour ce patient grâce à cette information? Ce type de formation continue médicale est très populaire. Depuis septembre 2006, près de 4500 membres de l'Association médicale canadienne ont soumis plus d'un million de questionnaires IAM reliés aux synopsis reçus par courriel. Les recherches précédentes suggèrent que le format de réponse des questionnaires IAM peut influencer la participation des médecins et que l'utilisation de l'information pour un patient spécifique peut être liée à certains facteurs mesurables par IAM. Les mêmes recherches indiquent que certains formats peuvent stimuler des réponses plus réfléchies. Aucune recherche n'a étudié l'effet de ce genre de formation continue sur la santé de patients spécifiques. Les objectifs étaient donc d'améliorer les programmes d'éducation continue médicale qui utilisent les questionnaires

IAM en déterminant les formats de réponse qui optimisent la participation des médecins ainsi que l'apprentissage réflectif, et d'explorer les facteurs reliés à l'utilisation de l'information. Ceux-ci ont été accomplis en exécutant une analyse de la survie, ainsi que des modèles de régression logistique mixtes.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
Study 1: Physician Participation by IAM Response Format

1.1 Previously Published Material

Excerpts of the introduction to this section have been taken from a related grant proposal [14].

1.2 Introduction

The "Acquisition-Cognition-Application" model from information science inspired the development of the Information Assessment Method (IAM) [1] in order to better understand the way physicians use clinical information in their decision-making.

This report deals with a prospective observational component of a mixed-methods study whose objective was to explore determinants of information use by examining the delivery of synopses of research-based clinical information to Canadian physicians.

A longstanding challenge for knowledge translation (KT) research involves the testing of methods to accelerate the use of clinical research in practice through continuing medical education (CME). In Canada, one specific type of accredited CME is very popular: the reading of e-mailed synopses of selected peer-reviewed clinical research, followed by the completion of a short questionnaire called the Information Assessment Method (IAM). Since September 2006, about 4,500 CMA members have submitted more than one million IAM questionnaires linked to e-mailed synopses. However, no research has studied how these types of CME programs lead to greater

reflection on this type of research-based information, its uptake in clinical practice, or observable improvements in patient health.

This study's objective is to refine the Information Assessment Method (IAM) linked to synopses of research articles delivered as email alerts. This is aligned with the needs of knowledge user applicants who seek to optimally promote reflective learning as the basis for continuing medical education. The term "reflection" is intended to indicate a conscious and deliberate reinvestment of mental energy aimed at exploring and elaborating one's understanding of the problem one has faced (or is facing).

While family physicians rarely have time to read journal articles in their entirety, continuing education programs that involve rating synopses of research articles delivered as email are popular. Based on a theoretical model from information science, IAM is a brief self-administered questionnaire developed with knowledge user partners at the College of Family Physicians of Canada and validated in CIHR funded research. When linked to a synopsis of clinical research, completing one IAM questionnaire encourages reflection on that clinical information while capturing its value for the health professional. IAM conceptualizes the value of clinical information in four constructs: (1) its clinical relevance for a specific patient, (2) its cognitive impact (10 item response categories), (3) any use of this information for the patient (four item response categories), and (4) if used, any expected health benefits (five item response categories) [1]. The construct of cognitive impact is relevant to KT insofar as clinical information that has a positive cognitive impact is strongly associated with the use of that information for a specific patient [15].

From September 2006 to the end of 2010, about 4,400 CMA members used IAM to submit 895,820 ratings of synopses called InfoPOEMs® (surely an unprecedented level of participation in a Canadian CME programme). For clinicians, synopses like InfoPOEMs® are a tailored product within the knowledge creation funnel of the Knowledge to Action process. The popularity of IAM linked to synopses has been replicated in other work in a project involving the Canadian Pharmacists Association and the CFPC. In 2010, 5,346 family physicians used IAM to evaluate emailed content from "Therapeutic Choices", a reference book. In addition to InfoPOEMs®, this project is ongoing, yet each project uses a different IAM response format. This difference arose because it was felt that altering the response format could more strongly guide reflective learning (See Appendix A2 for screenshots).

Of interest here is the fact that two different IAM response formats were used in 2008. This resulted in the following three data series:
Series 1 (S1) from September 8, 2006 to February 17, 2008: IAM was used by 1,324 practising family physicians to submit 62,928 synopsis ratings. Physicians used a response format that asked them to "Check all that apply". For S1, the use of IAM, as well as the content and construct validity of IAM have been documented. It has also been reported how IAM helped to document reflective learning among a subgroup of physicians who were interviewed about their synopsis ratings. Physicians reported an average of 1.4 items or types of cognitive impact (range 1-10) per synopsis.
Series 2 (S2) from February 18 to April 2, 2008: IAM was used by 965 practising family physicians to submit 10,316 synopsis ratings. To encourage respondents to consider and come to a judgement about each item of cognitive impact, a forced "Yes" or "No" response format was used. However,

response rates (defined as submission of completed IAM questionnaires) dropped.

Series 3 (S3) from April 3, 2008 to present: Given a drop in S2 response rates, the principal knowledge user and CMA decision-maker recommended that IAM revert back to a "Check all that apply" response format. After reverting back to "Check all that apply" in April 2008, IAM use continued to grow in S3. Presently, about 5,000 synopsis ratings are submitted each week. However, little is really known about what happened in S2 (e.g. what was the true magnitude of the drop in response rate?) To begin to address the question of what happened in S2, a pilot study was conducted of archived data to descriptively summarize IAM ratings. This revealed a striking increase in the number of checked items of cognitive impact ("Yes" responses) per synopsis. Thus, when physicians used a forced choice "Yes-No" response format, "Yes" responses for all items or types of cognitive impact increased in frequency compared to a "Check all that apply" format. At the same time, a noticeable drop in response rate during S2 occurred under the forced choice format, presumably due to the extra time and work required with no extra reward in the form of education credits. This raises the possibility that increased reflection occurred at the cost of participation by some physicians. Given the possibility of greater reflection and less participation when using a response format with greater cognitive burden, it is crucial for knowledge users to know which IAM response format optimally balances these issues. The proposed research will refine IAM, in terms of understanding which response format optimizes reflection and physician participation. The integrated KT approach will facilitate use of this new knowledge, given our knowledge user applicants have helped to define the

research questions, and have a proven record of translating the research findings into national policy and programs.

We are interested in measuring and comparing incidence of POEM rating between three time periods, the outer two of which are to be considered homogeneous in terms of rating formats, and one containing the additional burden of a "check all that apply" framework as opposed to a "default yes" one.

1.3 Analysis

1.3.1 The Data

The collected data set contains each recorded POEM rating as an observation with the variables Format (1 for S1 and S3, 2 for S2), MDID (MD ID number), Poem ID (POEM ID number), Poem Date (date of push or sending of POEM) and Answer Date (date of rating). 390,529 ratings were collected among 2,615 MDs and 473 different POEMs. Fig 1–1 shows a month-by-month comparison between ratings received and number of different POEMs being rated. There appears to be a dip in S2 that might be suspected to be attributable to the format change.

1.3.2 The Problem

Our goal is to investigate the possibility that rating incidence in S2 was hindered by the format change. A naive attempt to accomplish this would be to simply calculate a hereafter named "naive incidence ratio" for each period letting

$$H_i = \frac{\#Ratings\ Received\ in\ period\ Si}{\#PossiblePOEM/MD\ pairs\ pushed\ in\ period\ Si}$$

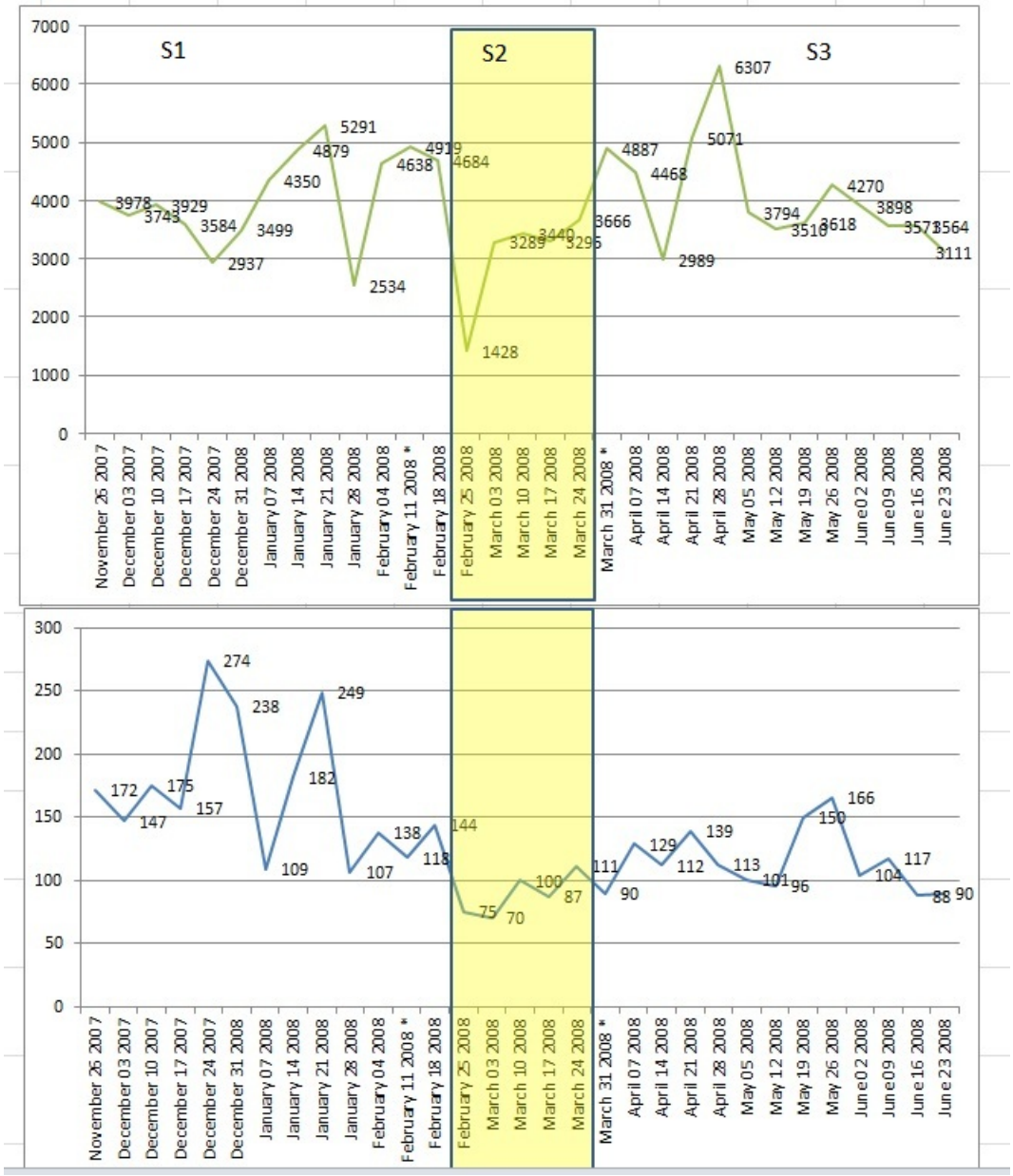


Figure 1–1: Comparison of Ratings Received (Top) to Number of POEMs Rated (Bottom)

In this way, we would compute:

$$H_1 = \frac{318,301}{695,550} = 45.76\%, H_2 = \frac{17,044}{66,832} = 25.50\%,$$

$$H_3 = \frac{46,904}{176,988} = 26.50\%$$

Table 1–1: Summary of "Naive Incidence Ratios" for All Periods

H1	H2	H3
45.76%	25.50%	26.50%

Although this does seem to indicate a drop in participation in S2, it does not demonstrate any redemptive qualities in S3. However, there are reasons that prevent this reasoning from holding water. These reasons will be presented in order of their discovery.

Namely, the "naive method" ignores the possibility that a POEM may be pushed in during one period, but rated during another. For example, if a POEM is pushed to a particular rater in S1 who then only rates it in S2, there will be an artificial deflation and inflation of H_1 and H_2 respectively, since the denominators will not be affected. Fundamentally, a proper analysis should take these considerations into account.

Before any course of remedy can be suggested, there is a fundamental flaw in the data collected. Namely, only rating events were recorded. The case where a POEM is pushed and never rated is entirely possible (and in fact, a frequent enough occurrence) but was not explicitly recorded in the data set collected. This issue did not affect the "naive method" very much, since the denominators were ascertained globally by counting the POEMs pushed during each period, and not from the data set itself. Nevertheless, ignoring this possibility can cause dramatic differences in analysis results, and this data must be (and was) obtained.

1.3.3 A Solution

One viewpoint is that a suitable technique should be able compare risks of an event (rating) between periods (rating formats), while dealing with a possible period transition during the time between POEM push and rating.

Indeed, by defining an event as a rating of a pushed POEM, it is possible to calculate hazard ratios [21] for each rating format using a Cox regression model from survival analysis using rating format as a time-dependent covariate. We discuss this method in some detail in the next section.

1.3.4 Cox Regression with Time-Dependent Covariates

Cox regression was first developed in 1972 by Sir David Cox in his paper "Regression Models and Life Tables" [9], which became one the most popular journal articles in all of statistics (it is also the first appearance of the famous proportional hazards model as well as partial likelihood estimation). Unlike parametric methods, Cox's method does not require the choice of a particular probability distribution for survival times, allowing it to be more robust. In this way, it is deemed *semi*parametric. Furthermore, Cox proposed a relatively easy process to include so-called "time-dependent covariates", that may change in value over time, in Cox regression models. This topic was further discussed and improved upon by Kalbfleisch [20]. A powerful and convenient implementation of these methods can be found in SAS's PROC PHREG. A very popular and comprehensive description of this is found in [2] and is summarized and reviewed here.

The basic proportional hazards model without time-dependent covariates is usually written as

$$h_i(t) = \lambda_0(t) \exp\left(\beta_1 x_{i1} + ... + \beta_k x_{ik}\right) \tag{1.1}$$

Meaning, in this context, that the hazard for individual i at time t is a product of an unspecified non-negative function $\lambda_0(t)$ (which can be viewed as a baseline hazard function representing an individual with all covariates at 0), and an exponentiated linear function of a set of k fixed covariates.

Immediately, it is worth noting the relations between (1.1) and a few other models. Taking logarithms, we see that (1.1) is equivalent to

$$\log h_i(t) = \alpha(t) + \beta_1 x_{i1} + ... + \beta_k x_{ik} \tag{1.2}$$

where $\alpha(t) = \log \lambda_0(t)$. Now, if we happen to choose $\lambda_0(t)$ such that $\alpha(t) = \alpha$, a constant, then we have an exponential model. In the same way, picking so that $\alpha(t) = \alpha \log(t)$ or $\alpha(t) = \alpha \log(t)$ yields the Gompertz and Weibull [18] models respectively. Of course, the punchline here is that choice of $\lambda(t)$ is unnecessary for Cox regression which in a sense makes it a generalization of these other methods.

Equation (1.1) is called the proportional hazards model since it embodies the idea that the hazard for any individual is constantly proportional to the hazard for all other individuals. This is easily demonstrated by taking the ratio of hazards for two individuals i and j using (1.1):

$$\frac{h_i(t)}{h_j(t)} = \exp[\beta_1(x_{i1} - x_j) + ... + \beta_k(x_{ik} - x_{jk})] \tag{1.3}$$

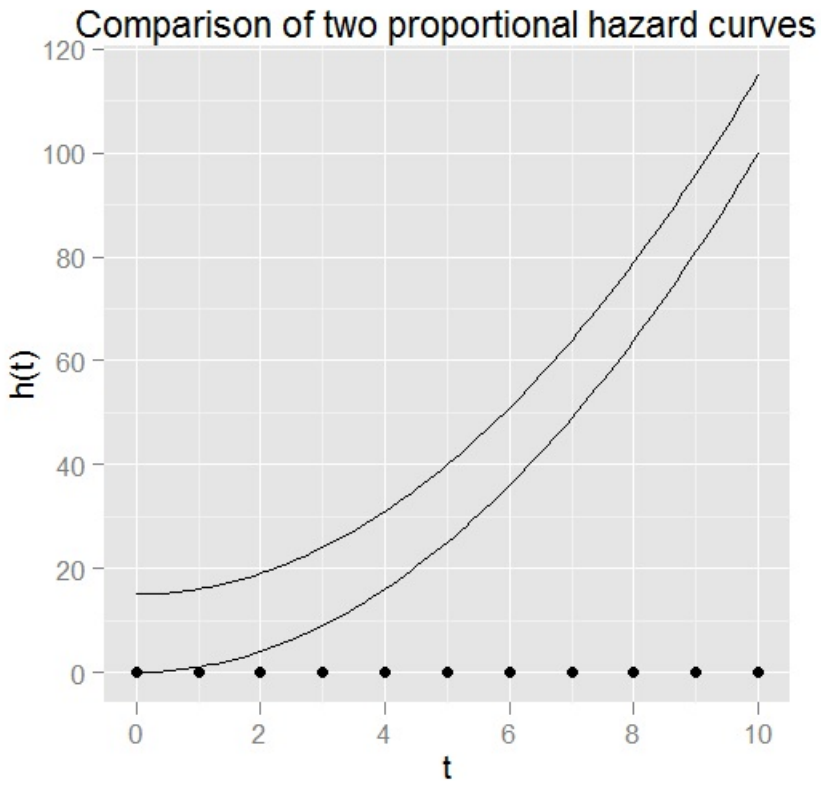


Figure 1–2: Comparison of two proportional hazard curves

We see that (1.3) is a constant over time, since $\lambda_0(t)$ is cancelled out. Thus, plotting hazards for two individuals over time should reveal parallel functions, as demonstrated in Fig 1–2, similar to the one by Allison [2].

1.3.5 Partial Likelihood

Although often overlooked, it is worth discussing the partial likelihood estimation method, as it differs from maximum likelihood. In this case, it allows for the estimation of $\boldsymbol{\beta}$ coefficients of the proportional hazards model without specification of the baseline hazard function $\lambda_0(t)$. Here, we will consider some general properties of the partial likelihood method, as well as underlying mathematics, as first formulated by Cox [8] and popularly summarized by Allison [2].

As will be shown, the likelihood function for the proportional hazards model in (1.1) can be factored into two parts. The first depends on $\lambda_0(t)$ and $\boldsymbol{\beta}$. The second depends solely on $\boldsymbol{\beta}$.

The idea behind partial likelihood is to disregard the first part that depends on $\lambda_0(t)$ and treat the second one (named the partial likelihood function) as a regular likelihood function by maximizing it with respect to $\boldsymbol{\beta}$. Since some information about $\boldsymbol{\beta}$ is discarded, there is loss of efficiency in the resulting estimates. However, these estimates are robust to the shape of the baseline hazard function, which is a great advantage. Just like maximum likelihood estimates, partial likelihood estimates are consistent and asymptotically normal.

Partial likelihood estimates depend only on the order of event times, not the numerical values themselves. In other words, monotonic transformations on the event times do not affect partial likelihood estimates.

We will now look at a few details about the method itself.

Assuming we have n independent individuals ($i = 1, ..., n$) and for each individual i we have three bits of data: t_i, σ_i and $\boldsymbol{x}_i$, where t_i is the time of event or censoring, σ_i is an indicator variable for whether or not t_i was censored, and $\boldsymbol{x}_i$ is a vector of the corresponding k covariate values.

While the regular likelihood function is written as the product of all the likelihoods of the individuals in the sample, a partial likelihood is the product of the likelihoods for only the observed events. Thus, if we have M events, the partial likelihood P can be written as

$$P = \prod_{j=1}^{M} L_j \tag{1.4}$$

where L_j is the jth event's likelihood, which is now described. The idea behind the likelihood construction for an event j is very well described in [2] through a question: Given that an event has occurred at time t, what is the probability that it occurred to individual i instead of another? The derivation of the intuitively appealing answer to this is too technical for this description, but the solution is to take the ratio of the hazard for individual $i(j)$ at time t_j to the sum of the hazards for all individuals who were also at risk at that time. So, the likelihood for event j would be

$$L_j = \frac{h_{i(j)}(t_j)}{\sum_{k \in R(t_j)} h_k(t_j)} \tag{1.5}$$

where $R(t_j)$ is the risk set of individuals (set of individuals who are currently at risk of experiencing an event) at time of event j, t_j. Assuming for now that there are no ties (events occurring at the same time value), it is then possible to compute the likelihoods for all events. For convenience, it is a good idea to sort all events increasingly by survival time (time to

event/censoring). This will make evaluating risk sets for each individual easier, since they will form a series of subsets. Conventionally, in case of a tie between event and censoring time, the censored observation is considered to be at risk at that time. Thus, substituting (1.1) into (1.5) we get

$$L_j = \frac{\lambda_0(t_j)exp(\boldsymbol{\beta x_{i(j)}})}{\sum_{k \in R(t_j)} \lambda_0(t_j)exp(\boldsymbol{\beta x_k})} \tag{1.6}$$

where $i(j)$ is the individual associated with event j. We notice that $\lambda_0(t_j)$ is common to all terms in the numerator and denominator, so we may cancel it out (demonstrating the lack of need for specification of the baseline hazard function in order to estimate $\boldsymbol{\beta}$). Thus, we are left with

$$L_j = \frac{exp(\boldsymbol{\beta x_{i(j)}})}{\sum_{k \in R(t_j)} exp(\boldsymbol{\beta x_k})} \tag{1.7}$$

It is also possible now to explain the earlier claim that partial likelihood estimation depends only on the order of event times, rather than the numerical values themselves. This can be seen by noticing that L_j only depends on t through the risk set $R(t_j)$. For example, supposing that the event indices $j = 1, ..., M$ have been ordered increasingly by event time, it is easy to see that the expression for L_1 would be unchanged $\forall t_1 < t_2$, since the risk set $R(t_1)$ would not be affected. Now, should t_1 be set so that $t_1 > t_2$, then we would have that individual $i_2 \notin R(t_1)$, since the event for individual i_2 would have occurred sooner and would thus no longer be at risk. In this way, all risk sets (and thus partial likelihoods and $\boldsymbol{\beta}$ estimates) are invariant under monotonic transformations of event times, since order would be preserved.

Now, putting all the partial likelihood pieces together we obtain a general expression for the partial likelihood for data with no ties and no time-dependent covariates yet:

$$P = \prod_{i=1}^{n} \left\{ \frac{exp\{\boldsymbol{\beta x_i}\}}{\sum_{j=1}^{n} Y_{ij} exp\{\boldsymbol{\beta x_j}\}} \right\}^{\sigma_i} \tag{1.8}$$

where Y_{ij} is an indicator variable for whether individual i experiences an event after individual j ($Y_{ij} = 1$ for $t_j \geq t_i$, and $Y_{ij} = 0$ otherwise).
It should be noted that the product in (1.8) is taken over all individuals instead of over all events as previously seen in (1.4). This is made possible by the indicator variable σ_i, which excludes all product terms in (1.8) corresponding to censored observations.

Now that the partial likelihood has been constructed, maximization with respect to $\boldsymbol{\beta}$ can be achieved as with a regular likelihood function, using a Newton-Raphson algorithm on the log-partial-likelihood:

$$\log(P) = \sum_{i=1}^{n} \delta_i \{\boldsymbol{\beta x_i} - \log \sum_{j=1}^{n} Y_{ij} \exp\{\boldsymbol{\beta x_j}\}\} \tag{1.9}$$

1.3.6 Ties

Up until this point, we have assumed that all events were distinct from one another (no ties), although identical censoring and event times could occur without violating any previous assumptions. Naturally, most data contain ties for event times and a proper treatment of them should be able to handle this contingency. Several methods exist to this end and may substantially change parameter estimates. In this sense, it is worth considering three of the candidates implemented in PROC PHREG in SAS. By default, SAS uses a method called Breslow's approximation, a speedy method that handles most situations well as long as the number of ties are proportionally small

to the size of the risk sets at all time points. In situations of heavy ties [17], Breslow's method has been known to underestimate certain parameter values. The approximation's speed is a result of simplifications applied to the likelihood equation denominator. Another approximation is from Efron, which also attempts to simplify the denominator of the likelihood equation, thereby increasing computational speed. It performs much better than Breslow's approximation, although computationally more intensive. In this way, Efron's method serves as a sound compromise between the Breslow and Exact methods. However, the Exact method is naturally preferable where resources permit [17].

1.3.7 The Exact method

The first and most plausible method for dealing with ties assumes that while events are recorded as having occurred simultaneously, this is simply a measurement error or simplification and the events did actually have a specific order in time. The main idea behind the exact method is that if it is not possible to determine the original ordering of tied events, it is then necessary to consider all of the possible orderings. For instance, let d be the number of tied events at a particular time measure. Then, there would be $d!$ different possible orderings for the actual event times. Labelling each of these orderings E_i, for $i = 1, ..., d!$, we are interested in $\Pr(E_1 \cup E_2 \cup ... \cup E_{d!})$. Now, since all of these events are mutually exclusive, we are able to write this as the sum of probabilities of each event so that

$$L_{t_i} = \sum_{i=1}^{d!} \Pr(E_i) \tag{1.10}$$

where t_i is the tied event time. Each term in (1.10) is a partial likelihood, as previously seen, for the particular ordering of the corresponding events.

It is not hard to see that for 10 tied events, the number of terms in 1.10 surpasses three million. Thus, computational concerns are inherent. However, this method has been re-expressed since its conception to allow for easier computation. Although PROC PHREG is able to perform these easier computations, it is still very intensive. Two approximations to this methods are commonly employed: Breslow's [6] (default in PROC PHREG) or Efron's [11]. In general, these are believed to work very well, but are known to potentially bias parameter estimates if the proportion of tied events to the size of the risk set at any time gets too large [12]. There is a contrasting method developed by Cox [9] called the "Discrete" method that treats time as discrete, so that several events can conceivably occur simultaneously without any particular ordering, but this technique will not be used here.

1.3.8 Time-Dependent Covariates

Mentioned by Cox in his original paper and further discussed by Kalbfleisch [20] in the context of Cox regression, time-dependent covariates are a kind that can change in value over the observation period, such as whether or not a patient had ever received a heart transplant, or whether someone was currently employed. To include time-dependent covariates, we modify (1.2) to become

$$\log h_i(t) = \alpha(t) + \beta_1 x_{i1} + ... + \beta_k x_{ik}(t) \tag{1.11}$$

where, in this instance, x_{ik} is the only time-dependent covariate, allowed to vary in time using any information about the individual before time t. Although several methods are used to handle time-dependent covariates, we will be concerned only with the *programming statements* method in PROC PHREG, where input data contains one observation per individual and the

time-dependent covariates are defined at time of program run using the original data variables. There are no differences in terms of results between this method and other ones, but it is deemed simplest to implement.

1.3.9 Partial Likelihood with Time-Dependent Covariates

The partial likelihood function $P(\boldsymbol{\beta}; \boldsymbol{Y}, \boldsymbol{t})$ with time-dependent covariates has the same form as equations (1.4) and (1.8), however covariates are now indexed by time.

$$P(\boldsymbol{\beta}; \boldsymbol{Y}, \boldsymbol{t}) = \prod_{i=1}^{n} \left\{ \frac{exp\{\boldsymbol{\beta x_i}(t_i)\}}{\sum_{j=1}^{n} Y_{ij} exp\{\boldsymbol{\beta x_j}(t_j)\}} \right\}^{\sigma_i} \tag{1.12}$$

The time-dependent covariates for each term in (1.8) must be individually computed, since they may change for each event time. When clustering is present, a sandwich covariance estimator by Lin & Wei [22] can be used by Proc PHREG to account for the dependency that, if ignored, can dramatically skew variance estimates, as later illustrated.

1.3.10 Application of Cox Regression with Time-Dependent Covariates

In the previous section, key concepts about Cox Regression and time-dependent covariates were described. The precise way these methods can be applied to the data at hand will now be outlined. Popular examples of similar applications of these models have been documented by Suriyasathaporn [28] and Crowley [10], and are of great reference. Remembering that the objective is to determine if there are significant differences between the risks of rating a POEM for different rating formats, we may define the rating of a POEM as an event and view the POEM's rating format as a time-dependent covariate.

Analysis of Maximum Likelihood Estimates								
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
tds2	1	-0.58385	0.00750	6056.9908	<.0001	0.558	0.550	0.566
tds3	1	-0.04025	0.00854	22.2094	<.0001	0.961	0.945	0.977

Figure 1–3: Cox regression of POEM ratings with rating period as a time-dependent covariate, using EFRON's approximation for ties without accounting for clustering

In this way, Cox Regression may be applied to the hazard of POEM rating (event) using rating format as a single time-dependent covariate. This covariate has three levels; one for each rating period (S1, S2, S3).

The results are shown in Fig 1–4 and describe a highly significant difference (p-value<0.0001) between the hazard of POEM rating in the reference period (S1) and in S2 (variable name *tds2*). In fact, the hazard ratio estimate for S2/S1 is 0.56 (95 % C.I: [0.442, 0.703]), meaning that there is an estimated 44 % (C.I: [29.7, 55.8]) decrease in the hazard of rating in S2 versus S1. No significant difference was found between the hazard of rating in S1 and S3, which is in line with the fact that the rating formats for those periods were similar. By comparing these results with an "unclustered" analysis without using a sandwich variance estimator in Fig 1–3, we see that although the parameter estimates have not changed, the confidence limits have become wider, since by taking into account the dependency inherent in the data, we are acknowledging a lack of precision compared to the "unclustered" counterpart where dependency in the data is not accounted for. An exponential model was fitted to the S1 data, to obtain a constant hazard estimate for this period. Using the hazard ratio estimates from the previous Cox models, Fig 1–5 shows the absolute hazard estimates by period, under the piecewise constant hazard assumption conditional on the

Analysis of Maximum Likelihood Estimates									
Parameter	DF	Parameter Estimate	Standard Error	StdErr Ratio	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
tds2	1	-0.58385	0.11824	15.761	24.3827	<.0001	0.558	0.442	0.703
tds3	1	-0.04025	0.02373	2.779	2.8757	0.0899	0.961	0.917	1.006

Figure 1–4: Cox Regression of POEM rating with Rating Period as a Time-Dependent Covariate, using EFRON's approximation for ties and a sandwich covariance matrix estimator

S1 estimate, with 95% confidence intervals. In the interest of comparison, a piecewise linear continuous hazard model was also fitted, using only two variables that each attributed a fixed risk value for all ratings that were the same number of days past the beginnings of S2 and S3 respectively. Fig 1–6 shows the absolute hazard estimates by period (delimited by the curve "elbows") with 95% confidence intervals, using a sandwich covariance estimator. While both covariates were found to be significant, the S3 hazard is underestimated by this model (using the data as a reference point). To mitigate this, another piecewise linear hazard model was fitted with the variable *tds3* added from earlier. This was done to account for the possibility that aside from the gradual return of raters to the study after to the format reversion, several raters might have started rating again as soon as the change from S2 to S3 took place. Fig 1–7 illustrates the results. Although all covariates were found to be significant, S3 still appears underestimated. Thus, the constant hazard approach seems to perform best in this situation, which could be confirmed by a partial likelihood comparison (a discussion beyond the scope of the illustrative intentions here).

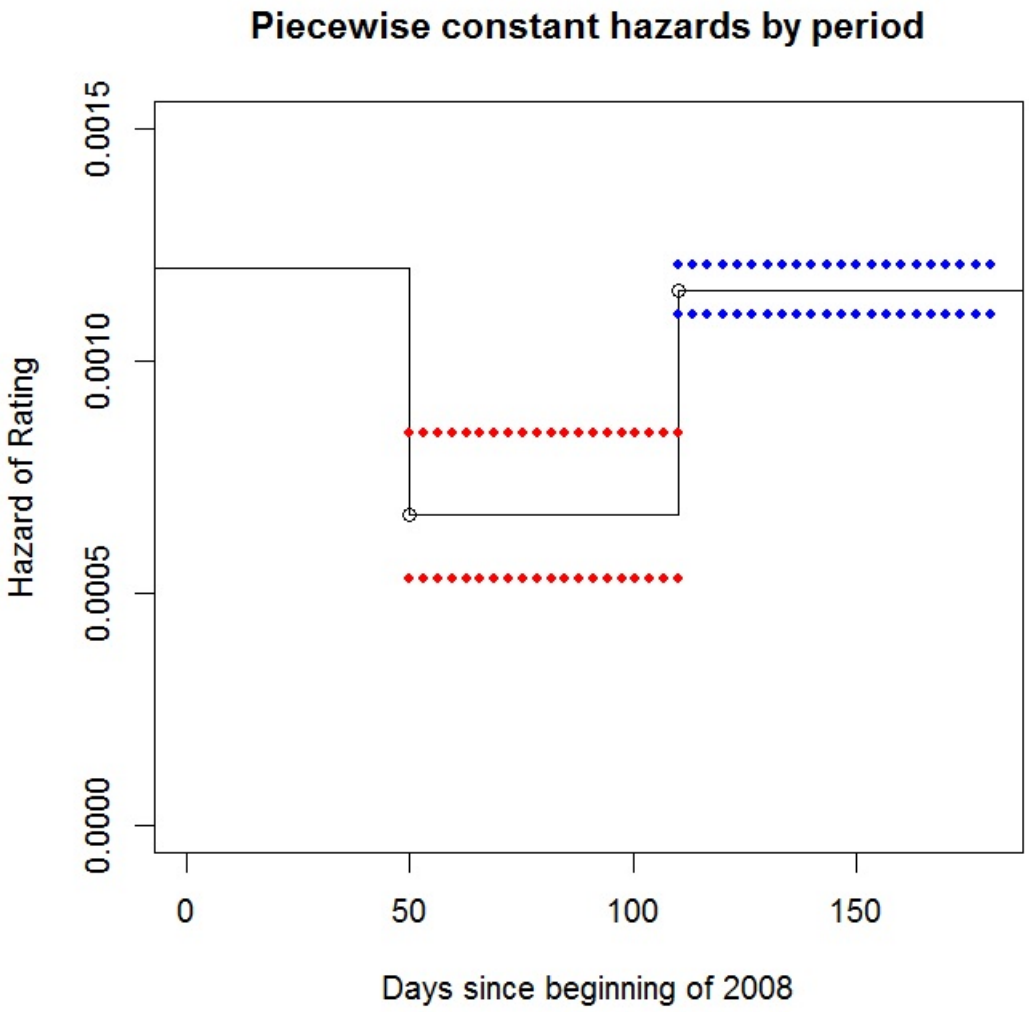


Figure 1–5: Piecewise constant hazards by period with 95% C.I., conditional on S1 exponential model hazard estimate

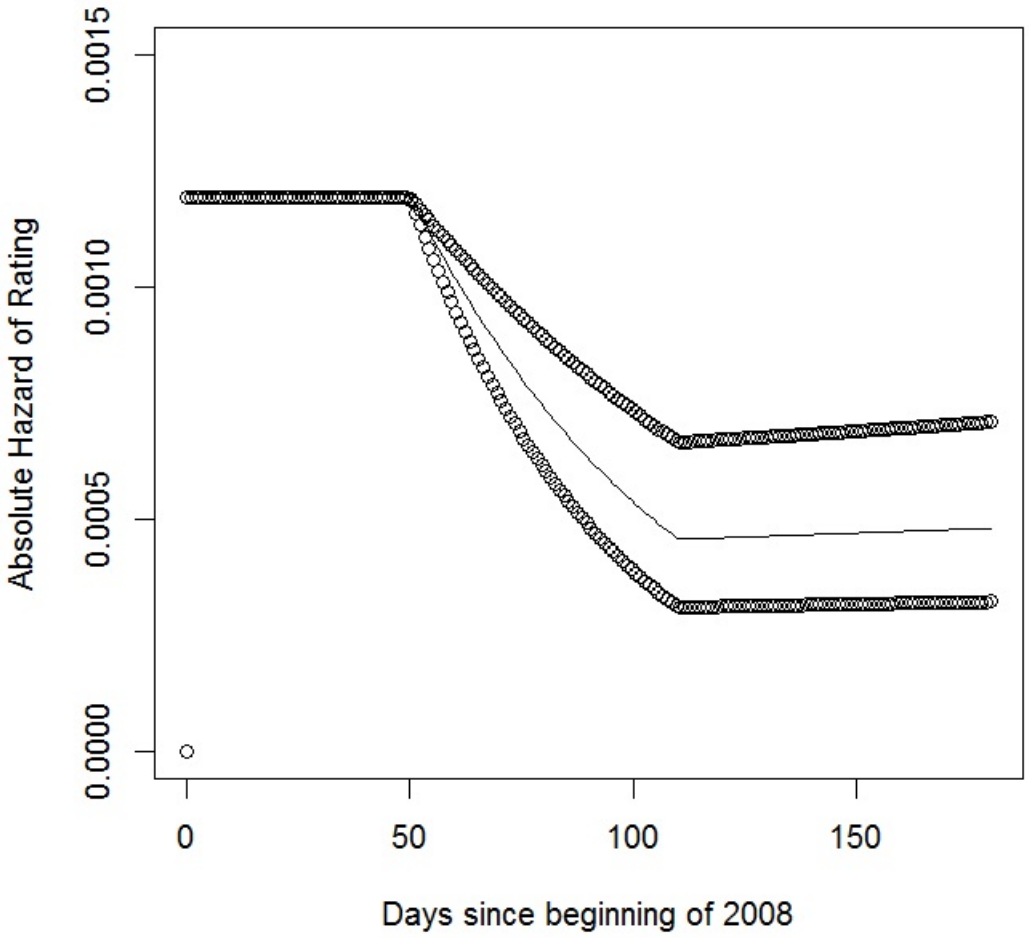


Figure 1–6: Piecewise linear continuous hazards model by period, with 95% C.I., conditional on the S1 exponential model hazard estimate

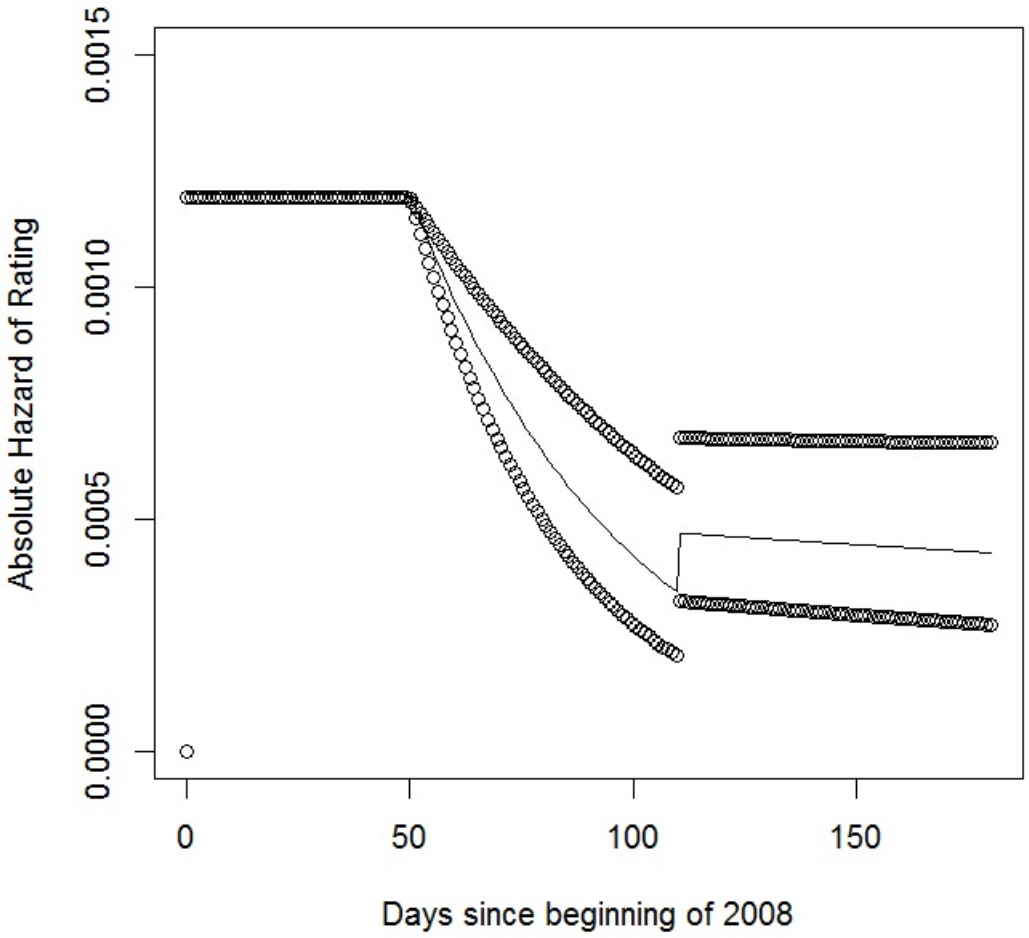


Figure 1–7: Piecewise linear hazards model by period with *tds3* added, with 95% C.I., conditional on the S1 exponential model hazard estimate

1.4 Discussion

Although we have shown that there is evidence supporting the idea that format S2 has a negative effect on user participation, it is possible that the higher reflective learning that it induces offsets this. Using this work as a pilot study, a grant proposal was drafted and submitted [14], but ultimately not funded. The study proposed could offer greater insight into this phenomenon, and is left for further investigation.

CHAPTER 2
Study 2: Assessing Determinants of Information Use

2.1 Previously Published Material

Excerpts of this section's introduction and results have been taken from [15], which is the published article that the following study has led to.

2.2 Introduction

Here, we are concerned with a prospective observational component of a mixed-methods study whose objective was to explore determinants of information use by examining searches for clinical information conducted by family physicians (FPs).

In information studies, multiple models have conceptualized information behaviour; however, no single model has dominated the research landscape, in part because models focus on different elements of information behaviour. The acquisition cognition application (ACA) model is unique in that it describes sequential phases involved in the assessment of the value of information, whereby the value of information is ultimately exhibited by its application or use. Originally, the ACA model was illustrated through a scenario whereby a scholar comes to a library to consult books or articles to be better informed about the state of knowledge in a particular field (acquisition). During his or her reading, cognition takes place. In the application phase, choices are made about which information is used to create his or her paper. In this sense, the ACA model is particularly suited to study information use in sequence, complementary to models which

illustrate information seeking and information search behaviour. In this study, we show how a naturalistic and longitudinal study of searches for clinical information in primary health care provides empirical data to support the applicability of the ACA model, as operationalized through the Information Assessment Method(IAM) [1]. IAM is a research tool that operationalizes the ACA model to study the value of objects of clinical information as perceived by the health professional in practice. This is conceptually different from the general utility of electronic resources at the point of care, which has been well studied. In accordance with the ACA model, health professionals (a) search for information to fulfill an objective, and retrieve objects of information such as a synopsis of clinical research (acquisition); (b) they absorb, understand, and integrate that synopsis (cognition); and then (c) they may use this newly understood and cognitively processed synopsis (application). In the context of primary care practice, when the family doctor rates an information object such as a synopsis of original clinical research, for example, IAM 2008 (see Fig 2–1) conceptualizes its value in three constructs: situational relevance, cognitive impact, and use or application of clinical information for a specific patient.

2.2.1 Acquisition

The construct of situational relevance is defined by acquiring information that achieved a search objective. In this construct, we seek to understand whether a search objective is met. Therefore, the IAM questionnaire asks the clinician to evaluate the situational relevance of the retrieved information. In information science, and particularly information retrieval,relevance can be seen from two perspectives: system relevance and user relevance.

While the system perspective is concerned with the relevance of retrieved information with respect to an explicit query, situational relevance is a manifestation of the user perspective [24]. Situational relevance refers to the relationships between retrieved information objects and a specific task or problem, as experienced by the clinician. Relevance is determined by how well the retrieved information contributes to the resolution of this problem. The IAM construct of situational relevance is largely driven by the fact that physicians are frequently attempting to solve explicit patient-related problems. In a literature review previously conducted by the team, physicians' search objectives were examined [23]. The findings were operationalized in the IAM questionnaire as seven reasons or objectives for a search. These seven reasons capture the reasons why physicians search for information. While information technology evolves rapidly, the basic information needs that arise from clinical practice are relatively stable. For example, in a study that predates the widespread use of electronic resources, answering clinical questions about specific patients was the main driver of information need [7].

2.2.2 Cognitive Impact

In this construct, we seek to understand the types of cognitive impact that result when health professionals reflect on one object of retrieved information. IAM operationalizes the construct of cognitive impact through nine items that are a mix of both positive (e.g., I learned something new) and negative (e.g., I disagree with this information). The user may check one or more than one type of cognitive impact, and as such, a complex range of possibilities can be observed.

2.2.3 Application

In this construct, we simply seek to document whether there was an intention to use the retrieved information with a specific patient, operationalized as a "yes or no" question. In line with the ACA model, the application of retrieved information depends on (a) successfully acquiring information that is relevant to a search objective (i.e., situational relevance) and (b) a positive cognitive impact of that information on the professional (i.e., cognition). Consequently, two levels of analysis have emerged in the work: Level 1 is an evaluation of the search objective(s), and Level 2 is an evaluation of the cognitive impact of information hits. Thus, IAM is a multilevel questionnaire for the evaluation of retrieved clinical information. IAM is the product of publicly funded research, and both content and construct validity are presented elsewhere [5] and summarized at http://iam2009.pbworks.com (soon to be moved to http://iam.mcgill.ca).

2.2.4 Design and Participants

A prospective longitudinal study was conducted involving a cohort of physicians to whom research-based information was provided in a searchable knowledge resource on a handheld computer (shown in Fig 2–1). From 9 of 10 provinces, 41 family physicians (FPs) consented to participate, 36 of whom were certified by the College of Family Physicians of Canada. There were 24 men and 17 women, all in active practice, ranging in age from 28 to 70 (median = 44) years. Twenty-eight (68%) had a connection through teaching or research to a faculty of medicine. Participants entered the study between November 2007 and May 2008. Each participant had a unique start date defined by the date of their first rated search. Data collection ended March 2009.

2.2.5 Intervention/Instruments

Within IAM 2008 (see Fig 2–1), search objectives were operationalized as a checklist. This checklist of search objectives comprised seven reasons such as "to address a clinical question" or "to look up something I forgot". The construct of situational relevance was defined by acquiring information that achieved the search objective. The construct of cognitive impact was operationalized in a checklist of brief statements, such as "This information confirmed I did the right thing" (positive cognitive impact) or "There was a problem with this information" (negative cognitive impact). The use or application of clinical information for a specific patient was documented as a "yes or no" response. "Yes" responses to the question on application were pursued through semi-structured interviews. In these interviews, IAM ratings linked to a specific search were used by the interviewer to stimulate the participant's memory of that event. In psychological research, studies have examined real-time data collection using this technique, called Computerized Ecological Momentary Assessment. These studies have demonstrated this technique can enhance memory of events, such as searches for clinical information applied to a specific patient [26]. Each participant received a handheld computer (personal digital assistant) or Smartphone with IAM and Essential Evidence Plus® software providing access to the following resources: clinical decision or prediction rules, diagnostic calculators, abstracts of Cochrane Reviews, POEMs®, (see Appendix A1 and Fig 2–1) and EBM(Evidence Based Medicine) Guidelines.

Initial software installation was performed so the device was ready to go on delivery. Participants were trained to use IAM and Essential Evidence Plus®, and to transfer their rated searches to a server. As a single-user

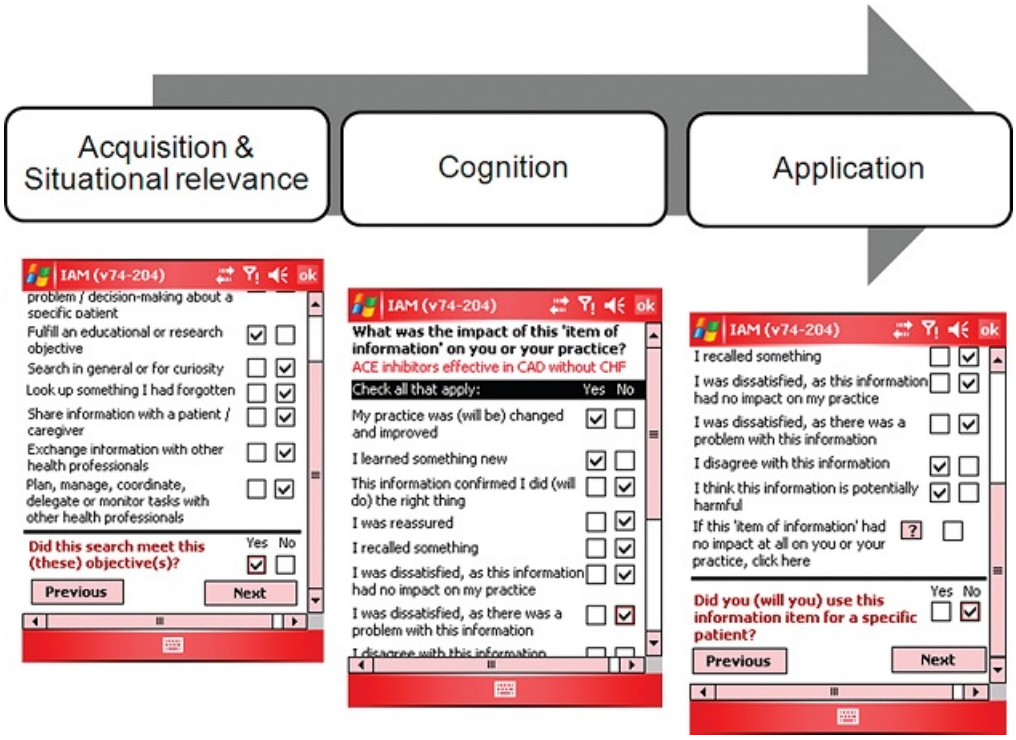


Figure 2–1: IAM 2008 on the handheld computer

device, IAM on the PDA documented the date and time of all information hits attributed to each participant. Searches contained one or more than one information hits, which were pages opened in resources within Essential Evidence Plus®. While PDAs had Wi-Fi enabled through the Windows Mobile 6 operating system, no data plan was provided. As such, PDA software was used offline. On each PDA, IAM copied the tracking of information hits from Essential Evidence Plus®. This allowed each information hit to be IAM-rated by participants, who earned continuing education credits for this activity. Rating a search required the participant to open IAM, and participants were reminded to rate their searches at device startup.

2.2.6 Data Analysis

For each IAM question, descriptive summaries of the ratings of information hits were produced. In bivariate analyses, we cross-tabulated the cognitive

impact of clinical information with its situational relevance and with its use for a specific patient. We used mixed logistic regression models to examine associations between information use (i.e., the outcome) and covariates: search objectives, achieving these search objectives (i.e., situational relevance), and the type of cognitive impact arising from the retrieved clinical information. The models accounted for the clustering of hits within searches and of searches within physicians.

2.3 Analysis

2.3.1 The Data

The data set contains ratings of information hits within searches (one rating= one observation). Some of the recorded variables were UserID (MD ID number), SearchID (unique ID given to each information search), hit ID (hit ID number), Search Date, Answer Date (date of rating), Age (age of rater), and binary outcomes for each of the IAM questionnaire items. After data cleaning (removal of data resulting from factors such as improper functioning of PDA devices, etc), 2,131 rated searches and 3,300 rated hits were collected among 40 MDs.

2.3.2 Random Effects Models for Clustered Binary Data

Here we present some background for generalized linear mixed models, in particular for binary outcomes in a clustered context, using work from Guo [16], West [16], Kachman [19] and Goldstein [13].

2.3.3 Generalized Linear Mixed Models

Suppose $\mathbf{Y}$ is a $(n \times 1)$ vector of observations and γ is a $(r \times 1)$ vector of random effects. A generalized linear mixed model (GLMM) can be regarded

(by Guo [16], for instance) as taking the form

$$\mathbb{E}[\mathbf{Y}|\gamma] = g^{-1}(\mathbf{X}\beta + \mathbf{Z}\gamma) \tag{2.1}$$

where g is a differentiable monotonic link function. We assume that $\mathbf{X}$ is a $(n \times p)$ matrix of rank k, and $\mathbf{Z}$ is a $(n \times r)$ design matrix for the random effects. Here, random effects are assumed to be normally distributed with mean $\mathbf{0}$ and variance matrix $\mathbf{G}$. We note that (2.1) contains a linear mixed model as an argument for g^{-1}. This is named the linear predictor component of the generalized linear mixed model:

$$\eta = \mathbf{X}\beta + \mathbf{Z}\gamma$$

Now, the variance of the data, conditional on the random effects, can be expressed as

$$\text{Var}[\mathbf{Y}|\gamma] = \mathbf{A}^{1/2}\mathbf{R}\mathbf{A}^{1/2}$$

where $\mathbf{A}$ is a diagonal matrix containing the variance functions of the model. A variance function is an expression of the variance of a response as a function of the mean. The matrix $\mathbf{R}$ is a variance matrix that, for models calling for so-called "G-side" random effects only (i.e. all random effects in the model are elements of γ) as opposed to "R-side" random effects, allows for additional scale parameters to be included in the conditional distribution of the data through the inclusion of a scale parameter ϕ ($\phi > 0$) in

$$\mathbf{R} = \phi\mathbf{I}$$

Guo [16] mentions that an important distinction between these type of models is that a model containing no G-side random effects is known as a marginal or "population-averaged" model.

DIST=	Distribution	Variance function $a(\mu)$	$\phi \equiv 1$
BETA	beta	$\mu(1-\mu)$	No
BINARY	binary	$\mu(1-\mu)$	Yes
BINOMIAL | BIN | B	binomial	$\mu(1-\mu)/n$	Yes
EXPONENTIAL | EXPO	exponential	μ^2	Yes
GAMMA | GAM	gamma	μ^2	No
GAUSSIAN |G | NORMAL | N	normal	1	No
GEOMETRIC | GEOM	geometric	$\mu+\mu^2$	Yes
INVGAUSS | IGAUSSIAN | IG	inverse Gaussian	μ^3	No
LOGNORMAL |LOGN	log-normal	1	No
NEGBINOMIAL | NEGBIN | NB	negative binomial	$\mu+k\mu^2$	Yes
POISSON | POI | P	Poisson	μ	Yes
TCENTRAL | TDIST | T	t	1	No

Figure 2–2: Variance Functions in PROC GLIMMIX, found in the SAS User's Guide [25]

Fig 2–2 contains some examples of values for variance functions and whether or not the corresponding scale parameter is 1. Specifically, these are the ones included in PROC GLIMMIX in SAS.

2.3.4 Multilevel Data

Quite often, observational data collected in the sciences contain a clustered or hierarchical structure. For example, Goldstein [13] draws attention to the biological sciences where inheritance is a common vehicle for many traits of interest, a natural hierarchy exists between subjects of the same family. It is natural, for instance, to suggest that the progeny of the same family may very well share more similarities in their characteristics (measurable or otherwise) than subjects chosen randomly from the total population. However, certain experimental designs also include such hierarchies in the collected data. Clinical trials are a common occurrence of this, where patients can be chosen from several randomly chosen care centres. However, there is no reason to limit this structure to a depth of only one step.

Care centres may also be randomly chosen from different cities, different countries, and so on.

As explained by Goldstein [13], hierarchy is looked at as *units* being grouped at different *levels*. Thus, patients could be level-1 units in a 3-level structure where level-2 are the care centres, and level-3 are the cities.

Goldstein also shows that a good way of understanding why the presence of such hierarchies is not trivial can be found through examples of social science applications through the following example. In some instances, selective schools or colleges may contain students with similar goals or skills. In others, young children are divided into different elementary schools, and patients are assigned to different clinics. The idea is that in the first instances, the grouping (selective schools) could be results of the similarities between those individuals (students), but in the latter instances, the grouping (elementary schools and clinics) could potentially just be a random allocation that will tend to eventually differentiate the groups from each other. In either case, taking group effects into account can be a very important part of the analysis. A failure to do so can lead (and has led) to invalid analysis. This occurred in an influential study involving elementary school children in the 1970's [4]. This study concluded that formal (as defined by them) styles of teaching reading produced better progress than other styles. In this study, data analysis was performed using multiple regression in the traditional way, considering children as the units of analysis without regard to possible groupings by teachers or by classes. The results of the analysis were statistically significant. A few years later [3], it was shown that if we account for the grouping by classes properly, the significant results vanish, thereby invalidating the inference that the difference in

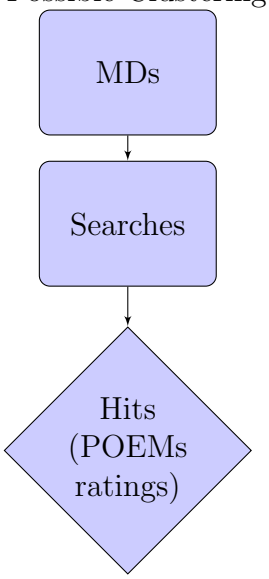
reading progress between the formal styles and the others. The literature cites this case as the first important instance of a multilevel analysis of social science data. The phenomenon taking place in this instance was that children within the same classroom tended to perform similarly, since they were taught together. Consequently, the students being studied are less informative than if the same number of students had been taught each separately by different teachers. In this way, one can argue that the unit of comparison should likely have been the teacher instead of the student. It is interesting to note that in this case, the students act an estimate of each teacher's effectiveness. Thus, including more students per teacher in the study would increase precision of the estimates, but the number of teachers being compared would remain the same. To improve the precision of the comparisons themselves, one would need to increase the number of teachers being compared (with the same or possibly somewhat smaller number of students per teacher).

In light of the aforementioned, it is not unreasonable to look for possible groupings in the IAM data set. Namely, the research team had initially suggested that ratings completed among a simply MD may contain similarities amongst each other and later on, that rating completed among a single search by an MD may also contain certain similarities amongst each other. In other words, a 3-level hierarchical structure was hypothesized and set to be investigated. It is illustrated in Fig 2–3.

2.3.5 The 2-level Linear Mixed Model

In the interest of eventually getting to a 3-level model, it is instructive to first describe the 2-level model, as many notions carry over and aid in understanding the framework as a whole. Since there are few conceptual

Figure 2–3: Diagram of Possible Clustering Hierarchy in IAM data



differences between the binary outcome and continuous outcomes cases, we will first deal with the continuous case and generalize later on.

For simplicity, we may first consider a 2-level model with one single explanatory variable, as viewed by West [29] and Goldstein [13],

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_j + e_{ij} \tag{2.2}$$

where y_{ih} is the outcome variables for the ith unit for level one and the jth unit for level two. β_0 is the intercept, x_{ij} is the explanatory variable, β_1 is the associated effect, u_j is the random effect describing the random variation in level two, and e_{ij} is the level one random effect.

We assume that

$$\mathbb{E}[u_j] = \mathbb{E}[e_{ij}] = 0$$

$$\text{Var}(u_j) = \sigma_u^2$$

$$\text{Var}(e_{ij}) = \sigma_e^2$$

$$\text{Cov}(u_j, e_{ij}) = 0$$

$$\text{Cov}(u_j, u_{j'}) = 0 \text{ for } j \neq j'$$

Within-cluster (or "intraclass") correlation, after controlling for the explanatory variable), can be described by

$$\rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

Goldstein mentions that in this setup, (2.2) can be considered a random effect model for panel data (such as longitudinal) or a growth curve model. In either case, i and j would be time point and individual indices respectively, and x_{ij} would be a time-dependent covariate.

Now, a more general notation can be achieved from (2.2) by replacing β_j by β_{1j} and letting

$$\beta_{0j} = \beta_0 + u_{0j}$$

$$\beta_{1j} = \beta_1 + u_{ij}$$

where u_{0j} and u_{1j} are random variables also with $\text{Var}(u_{0j}) = \sigma_{u0}^2$, $\text{Var}(u_{1j}) = \sigma_{u1}^2$, $\text{Cov}(u_{0j}, u_{1j}) = \sigma_{u01}$ Thus, we can now write the model equation as

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + (u_{0j} + u_{1j} x_{ij} + e_{0ij}) \tag{2.3}$$

$$\text{Var}(e_{0ij}) = \sigma_{e0}^2$$

We have added an extra index to the level-1 residual term, which proves useful for complex variance structure models (longitudinal, etc). Now, in this way, the response y_{ij} is expressed as the sum of both fixed terms and random terms (in parentheses). In matrix form, we would have

$$\mathbb{E}(Y) = X\beta \tag{2.4}$$

where $Y = \{y_{ij}\}$, so that

$$\mathbb{E}(y_{ij}) = X_{ij}\beta = (X\beta)_{ij}$$

where $X = \{X_{ij}\}$ is the design matrix for the explanatory variables and X_{ij} is the ijth row of X. Thus, for our model with one explanatory variable described by 2.3, we have $X = \{1, x_{ij}\}$. The random variables in the model are usually referred to as the "residuals". As a sanity check, we may notice that if we have only a 1-level model, then e_{0ij} becomes the typical linear model residual term. In the interest of symmetry, we may also define an explanatory variable x_{0ij} (which simply takes the value of 1) to be associated with β_0 and its residual u_{0j}.

2.3.6 Variance Components Model Parameter Estimation

The model described in (2.3) suggests that we must estimate six parameters. Namely, the fixed coefficients β_0 and β_1, as well as the variances σ^2_{u0}, σ^2_{u01} and σ^2_{e0} of the three random effects and the covariance of u_{0j} and u_{1j} that was named σ_{u01}. Those last four are named the *random parameters*, since they describe the random effects portion of the model. Now, in its simplest form, a 2-level model includes only σ^2_{u0} and σ^2_{e0} as random parameters (e.g. no random coefficients). Such a model is called a variance components model since given the fixed part of the model (referred to as the *fixed predictor*), we compute

$$\mathrm{Var}(y_{ij}|\boldsymbol{\beta_0}, \boldsymbol{\beta_1}, x_{ij}) = \mathrm{Var}(u_0 + e_{0ij}) = \sigma^2_{u0} + \sigma^2_{e0} \tag{2.5}$$

Since u_0 and e_{0ij} are independent random variables, we get that the response variance is effectively the sum of a level 1 and of a level 2 variance. This is to say that if in the current IAM context, if we were only to consider

the levels of hits and MD's (POEM ratings being clustered by MD's only), then the total variance for each rating is a constant and any two ratings by the same MD have computed covariance

$$cov(u_{0j} + e_{0i_1j}, u_{0j} + e_{0i_2j} = cov(u_{0j}, u_{0j}) = \sigma^2_{u0}, \tag{2.6}$$

since the e_{ij}'s have been assumed independent. In this way, the correlation between two clustered-by-MD POEM ratings can be expressed as

$$\rho = \frac{\sigma^2_{u0}}{(\sigma^2_{u0} + \sigma^2_{e0})} \tag{2.7}$$

This is essentially a measure of the proportion of the total variance that is found between MDs (intra-MD). Now, so long as there is more than one residual in the model, this correlation will be non-zero, thereby rendering OLS estimation inefficient in this context. In this way, from (2.5) and (2.6) we may construct the covariance matrix of, for instance, 4 ratings made by a single MD:

$$\begin{pmatrix} \sigma^2_{u0} + \sigma^2_{e0} & \sigma^2_{u0} & \sigma^2_{u0} & \sigma^2_{u0} \\ \sigma^2_{u0} & \sigma^2_{u0} + \sigma^2_{e0} & \sigma^2_{u0} & \sigma^2_{u0} \\ \sigma^2_{u0} & \sigma^2_{u0} & \sigma^2_{u0} + \sigma^2_{e0} & \sigma^2_{u0} \\ \sigma^2_{u0} & \sigma^2_{u0} & \sigma^2_{u0} & \sigma^2_{u0} + \sigma^2_{e0} \end{pmatrix} \tag{2.8}$$

Furthermore, the assumption that covariance between level-1 units in different level-2 clusters (ratings made by different MDs, for example), the covariance matrix for ratings from two different MDs (4 ratings from MD-1 and 2 ratings from MD-2) would have the following block-diagonal structure:

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \tag{2.9}$$

where

$$A_1 = \begin{pmatrix} \sigma^2_{u0} + \sigma^2_{e0} & \sigma^2_{u0} & \sigma^2_{u0} & \sigma^2_{u0} \\ \sigma^2_{u0} & \sigma^2_{u0} + \sigma^2_{e0} & \sigma^2_{u0} & \sigma^2_{u0} \\ \sigma^2_{u0} & \sigma^2_{u0} & \sigma^2_{u0} + \sigma^2_{e0} & \sigma^2_{u0} \\ \sigma^2_{u0} & \sigma^2_{u0} & \sigma^2_{u0} & \sigma^2_{u0} + \sigma^2_{e0} \end{pmatrix}, A_2 = \begin{pmatrix} \sigma^2_{u0} + \sigma^2_{e0} & \sigma^2_{u0} \\ \sigma^2_{u0} & \sigma^2_{u0} + \sigma^2_{e0} \end{pmatrix} \tag{2.10}$$

as Goldstein [13] succinctly presents it,

$$V = \begin{pmatrix} \sigma^2_{u0} J_{(4)} + \sigma^2_{e0} I_{(4)} & 0 \\ 0 & \sigma^2_{u0} J_{(2)} + \sigma^2_{e0} I_{(2)} \end{pmatrix} \tag{2.11}$$

where $I_{(n)}$ is the $n \times n$ identity matrix and $J_{(n)}$ is the $n \times n$ matrix of ones. Although this is for the 2-level case, we are reminded that this model is collapsible to the 1-level case by setting $\sigma^2_{u0} = 0$.

2.3.7 The 3-level Linear Mixed Model

Now, in order to extend model (2.3) to a three-level case, we may write

$$y_{ijk} = \beta_0 + \beta_1 x_{ijk} + v_{0k} + u_{0jk} + e_{0ijk} \tag{2.12}$$

where no random coefficients are present, k has been added as a level-3 index, v_{0k} and u_{0jk} are level-3 and level-2 random intercepts respectively, and x_{ijk} is still an observed explanatory variable. We also assume

$$\mathbb{E}[u_{0jk}] = \mathbb{E}[v_{0k}] = \mathbb{E}[e_{0ijk}] = 0$$

$$\text{Var}(u_{0jk}) = \sigma^2_{u0}$$

$$\text{Var}(v_{0k}) = \sigma^2_{v0}$$

$$\text{Var}(e_{0ijk}) = \sigma^2_{e0}$$

together with the restriction that random effects across clusters for the same level, as well as random effects across different levels are uncorrelated.

Borrowing from West [29], we can write the marginal variance-covariance matrix for an MD, following the hierarchy described by Fig 2–3. Analogously to (2.7), the MD-level correlation between two clustered-by-search-and-by-MD POEM ratings can be expressed as

$$\rho_{MD} = \frac{\sigma^2_{v0}}{(\sigma^2_{v0} + \sigma^2_{u0} + \sigma^2_{e0})} \tag{2.13}$$

And the search-level correlation would be:

$$\rho_{search} = \frac{\sigma^2_{v0} + \sigma^2_{u0}}{(\sigma^2_{v0} + \sigma^2_{u0} + \sigma^2_{e0})} \tag{2.14}$$

For illustrative purposes, we'll assume that this MD performed two searches, obtaining two hits in the first search and three in the second. Thus, the first two rows and columns correspond to observations from both hits from the first search, and the other three rows/columns are for the three hits from the second search:

$$A_1 = \begin{pmatrix} \sigma^2_{v0} + \sigma^2_{u0} + \sigma^2_{e0} & \sigma^2_{v0} + \sigma^2_{u0} & \sigma^2_{v0} & \sigma^2_{v0} & \sigma^2_{v0} \\ \sigma^2_{u0} + \sigma^2_{u0} & \sigma^2_{v0} + \sigma^2_{u0} + \sigma^2_{e0} & \sigma^2_{v0} & \sigma^2_{v0} & \sigma^2_{v0} \\ \sigma^2_{v0} & \sigma^2_{v0} & \sigma^2_{v0} + \sigma^2_{u0} + \sigma^2_{e0} & \sigma^2_{v0} + \sigma^2_{u0} & \sigma^2_{v0} + \sigma^2_{u0} \\ \sigma^2_{v0} & \sigma^2_{v0} & \sigma^2_{v0} + \sigma^2_{u0} & \sigma^2_{v0} + \sigma^2_{u0} + \sigma^2_{e0} & \sigma^2_{v0} + \sigma^2_{u0} \\ \sigma^2_{v0} & \sigma^2_{v0} & \sigma^2_{v0} + \sigma^2_{u0} & \sigma^2_{v0} + \sigma^2_{u0} & \sigma^2_{v0} + \sigma^2_{u0} + \sigma^2_{e0} \end{pmatrix} \tag{2.15}$$

2.3.8 Binary Data

Up until this point we have assumed continuous data. We now specify the framework explained by Guo [16] for dealing with binary outcomes. In terms

of the initial description in (2.1), we may set

$$g(x)^{-1} = \text{expit}(x) = \frac{e^x}{1+e^x}$$

In this way, then that

$$g(x) = \text{logit}(x) = \log\left(\frac{x}{1-x}\right)$$

would be the link function for a simple two-level model that can be expressed as

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_0 + \beta_1 x_{ij} + u_j$$

where $p_{ij} = \Pr(y_{ij} = 1)$ (assuming a Bernoulli distribution for the y_{ij}) and u_j is a level-2 random effect. Furthermore, independence between observations is assumed, conditional on the random effects, which are assumed to follow

$$u_j \sim N(0, \sigma_u^2)$$

just as for linear multilevel models.

Now, since this implies

$$\Pr(Y = 1|\boldsymbol{x}_j, u_j) = \frac{exp(\beta_0 + \beta_1 x_{ij} + u_j)}{1 + exp(\beta_0 + \beta_1 x_{ij} + u_j)}$$

$$\Pr(Y = 0|\boldsymbol{x}_j, u_j) = \frac{1}{1 + exp(\beta_0 + \beta_1 x_{ij} + u_j)}$$

which is analogous to the traditional logistic regression case. Thus for this model, the conditional density function for the jth cluster can be neatly expressed as:

$$f(\boldsymbol{y}_j|\boldsymbol{x}_j, u_j) = \prod_{i=1}^{n_j} \Pr(Y = y_{ij}|\boldsymbol{x}_j, u_j) = \prod_{i=1}^{n_j} \frac{exp[y_{ij}(\beta_0 + \beta_1 x_{ij} + u_j)]}{1 + exp(\beta_0 + \beta_1 x_{ij} + u_j)} \quad (2.16)$$

After which, model parameter estimation boils down to integrating over the random effect (although other techniques do exist):

$$f(\boldsymbol{y}_j|\boldsymbol{x}_j) = \int f(\boldsymbol{y}_j|\boldsymbol{x}_j, u_j)h(u_j)\,du_j \tag{2.17}$$

where h() is the normal density function. Although the quantity in (2.17) is defined conceptually, if we were to approach it from a maximum likelihood point of view (which is not always the case), one of various approximation methods must be implemented.

2.3.9 IAM Application

Now, with this setup, the type of model we can to fit the IAM data would be a three-level with random intercepts and m explanatory variables:

$$\log\left(\frac{p_{ijk}}{1-p_{ijk}}\right) = \beta_0 + \sum_{l=1}^{m}\beta_l x_{ijk}^{(l)} + v_{0k} + u_{0jk} \tag{2.18}$$

where i, j and k are indices for levels 1, 2, and 3, and v_{0k} and u_{0jk} are the level-3 (MDs) and level-2 (searches) random intercepts.

2.4 Results

2.4.1 Acquisition of Clinical Information

Over an average of 320 days, 2,131 searches for clinical information were conducted by 40 family physicians. (One participant provided no data.) This frequency of searches averages to roughly one search per physician per week. Prior to the study, 34 (83%)physicians reported using online practice guidelines or journals. During the study, no attempt was made to influence the use of electronic knowledge resources. In terms of computer self-efficacy, 8 (20%) physicians rated their level of skill as advanced, 32 (78%) physicians

Meeting the search objective versus type of cognitive impact.

Type of cognitive impact	Objective met ($n = 2{,}482$; 75.2%)	Objective not met ($n = 818$; 24.8%)
Positive Cognitive Impact		
My practice was (will be) changed and improved	899 (36.2%)	64 (7.8%)
I learned something new	1,104 (44.5%)	142 (17.4%)
This information confirmed I did (will do) the right thing	1,378 (55.5%)	138 (4.2%)
I was reassured	1,351 (54.4%)	117 (14.3%)
I recalled something	1,021 (41.1%)	115 (14.1%)
Negative Cognitive Impact		
I am dissatisfied, as this information has no impact on my practice	10 (0.4%)	69 (8.4%)
I am dissatisfied, as there is a problem with this information	21 (0.9%)	46 (5.6%)
I disagree with this information.	6 (0.2%)	1 (0.1%)
I think this information is potentially harmful	11 (0.4%)	2 (0.2%)
This item of information had no impact at all on me or my practice	288 (11.6%)	492 (60.2%)

Figure 2–4: Meeting the search objective, by type of cognitive impact [15]

as intermediate, and 1 physician as beginner. Of these 2,131 searches, 83% were IAM-rated (n = 1,768). Each physician rated on average 44 searches (range=6−148). Seventy-five percent of rated searches were done with more than one objective in mind; the most frequently reported objective was to address a clinical question. In terms of situational relevance, at least one search objective was successfully met in 1,336 rated searches (76%). A tabulation of achieving the search objective by type of cognitive impact is found in Fig 2–4, and a distribution of search reasons can be found in Fig 2–5.

2.4.2 Cognition (Cognitive Impact of Clinical Information)

As more than one type of cognitive impact could be reported per information hit, 7,275 cognitive impacts were linked to 3,300 rated hits.

Reason	
Address a clinical question/problem/decision about a specific patient	1,310 (74%)
Look up something I had forgotten	672 (38%)
Share information with a patient/caregiver	624 (35%)
Exchange information with other health professionals	520 (29%)
Search in general or for curiosity	496 (28%)
Fulfill an educational or research objective	434 (25%)
Plan, manage, coordinate, delegate, or monitor tasks with other health professionals	197 (11%)

Figure 2–5: Reasons for searching [15]

2.4.3 Application of Clinical Information

Fifty-two percent of rated information hits (n = 1,708) were used for a specific patient. A summary of the types of reported cognitive impacts are found in Fig 2–6.

Types of reported cognitive impact		
This information confirmed I did (will do) the right thing	1,516	46%
I was reassured	1,468	45%
I learned something new	1,246	38%
I recalled something	1,136	34%
My practice was (will be) changed and improved	963	29%
No impact	780	24%
Negative impact (all four types combined)	166	5%

Figure 2–6: Types of reported cognitive impact [15]

Information use for a specific patient versus type of cognitive impact.

Type of cognitive impact	Used for a specific patient ($n = 1{,}708$; 51.8%)	Not used for a specific patient ($n = 1{,}592$; 48.2%)
Positive Cognitive Impact		
My practice was (will be) changed and improved	709 (41.5%)	254 (16.0%)
I learned something new	756 (44.3%)	490 (30.8%)
This information confirmed I did (will do) the right thing	1,081 (63.3%)	435 (27.3%)
I was reassured	1,027 (60.1%)	441 (27.7%)
I recalled something	803 (47.0%)	333 (20.9%)
Negative Cognitive Impact		
I am dissatisfied, as this information has no impact on my practice	23 (1.4%)	56 (3.5%)
I am dissatisfied, as there is a problem with this information	22 (1.3%)	45 (2.8%)
I disagree with this information	2 (0.1%)	5 (0.3%)
I think this information is potentially harmful	5 (0.3%)	8 (0.5%)
This item of information had no impact at all on me or my practice	106 (6.2%)	674 (42.3%)

Figure 2–7: Information use for a specific patient by type of cognitive impact [15]

2.4.4 Association Between Acquisition and Cognitive Impact (A–C)

A relationship between situational relevance and cognitive impact was suggested in so far as positive cognitive impact was more likely when the search objective was met. Failing to meet search objectives was seen more commonly with negative cognitive impact.

2.4.5 Associations Between Cognitive Impact and Information Use for a Specific Patient (C–A)

Clinical information that had a positive cognitive impact was more likely to be used for a specific patient. This suggests an effect of cognitive impact on the use of clinical information. A summary of information use for a specific patient versus the type of cognitive impact can be found in Fig 2–7. In a mixed logistic regression model that included all nine types of cognitive impact, we found significant associations between specific types of cognitive impact and information use for a specific patient. Three types of cognitive

Solutions for Fixed Effects					
Effect	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept	-1.9717	0.3499	1497	-5.63	<.0001
Improved	1.2136	0.1346	1161	9.02	<.0001
Learned	-0.4256	0.1335	1161	-3.19	0.0015
Confirmed	0.8119	0.1220	1161	6.66	<.0001
Recalled	0.3028	0.1202	1161	2.52	0.0119

Figure 2–8: Proc GLIMMIX output for C-A Model

Associations between types of cognitive impact and the use of information for a specific patient.

Mixed logistic regression model	Use of information ~ Cognitive impact	
Type of cognitive impact	Estimated odds ratio	95% confidence interval
My practice was (will be) changed and improved	3.4	2.6–4.4
This information confirmed I did (will do) the right thing	2.3	1.8–2.9
I recalled something	1.4	1.1–1.7
I learned something new	0.7	0.5–0.9

Figure 2–9: Estimated Odds Ratios for C-A Model

impact were positively associated while one type of cognitive impact was negatively associated with the use of information. The SAS output and model summary are found in Fig 2–8 and Fig 2–9 and For example, the odds that clinical information was used for a specific patient increased by an estimate of 3.4-fold when the physician reported "My practice was (will be) changed and improved" as a result of this information. In contrast, reports of "I learned something new" (by itself) were negatively associated with the use of information for a specific patient.

2.4.6 Associations Between Acquisition and Information Use for a Specific Patient (A−A)

The results of a mixed logistic regression model including six search objectives along with a situational relevance variable (search objective met) are found in Fig 2–10 and Fig 2–11. In addition to searches done to address a clinical question, clinical information perceived as relevant to the situation (when a search objective was met) was positively associated with the use of that information for a specific patient. The odds that clinical information was used for a specific patient increased by an estimated 13.4-fold when the physician reported "To address a clinical question" as a search objective. Also, meeting the search objective increased the odds of information use for a specific patient 10.3-fold. However, searches done out of curiosity were negatively associated with the use of that clinical information. The model estimates a 70% drop in odds of in information use when searching for curiosity. No significant interaction was found between searches done to address a clinical question and the achievement of search objectives. Covariance parameter estimates were computed in Tables 2–1 and 2–2. According to equations (2.13) and (2.14), correlations of 0.15 and 0.07 were estimated between ratings collected from a same MD, and 0.56 and 0.43 among ratings acquired from a same MD within a same search, for the C-A and A-A models respectively. These estimates provide some insight into the clustered nature of the data and on the reasons why special consideration was indeed necessary. Conditional residuals for both C-A and A-A models were plotted as rudimentary diagnostics in figures 2–12 and 2–13.

Solutions for Fixed Effects					
Effect	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept	-7.2006	0.6608	1491	-10.90	<.0001
Address	2.5974	0.1795	1491	14.47	<.0001
Fulfill	-0.3591	0.1634	1491	-2.20	0.0281
Curiosity	-1.2161	0.1566	1491	-7.77	<.0001
Forgotten	0.3064	0.1481	1491	2.07	0.0388
Share	0.6047	0.1618	1491	3.74	0.0002
ObjectiveMet	2.3341	0.1964	1491	11.88	<.0001

Figure 2–10: Proc GLIMMIX output for A-A Model

Associations between search objectives, relevant clinical information (objective met), and the use of information for a specific patient.

Mixed logistic regression model	Use of information ~ Situational relevance + Search objectives	
Search objectives	Estimated odds ratio	95% confidence interval
To address a clinical question	13.4	9.4–19.1
Search in general or for curiosity	0.3	0.2–0.4
Search objective met	10.3	7–15.2

Figure 2–11: Estimated Odds Ratios for A-A Model

Table 2–1: Covariance Parameter Estimates for C-A Model

Cov Parm	Est	Sd
$\hat{\sigma}^2_{v0}$	1.44	0.18
$\hat{\sigma}^2_{u0}$	0.62	0.22

Table 2–2: Covariance Parameter Estimates for A-A Model

Cov Parm	Est	Sd
$\hat{\sigma}^2_{v0}$	1.76	0.18
$\hat{\sigma}^2_{u0}$	1.08	0.33

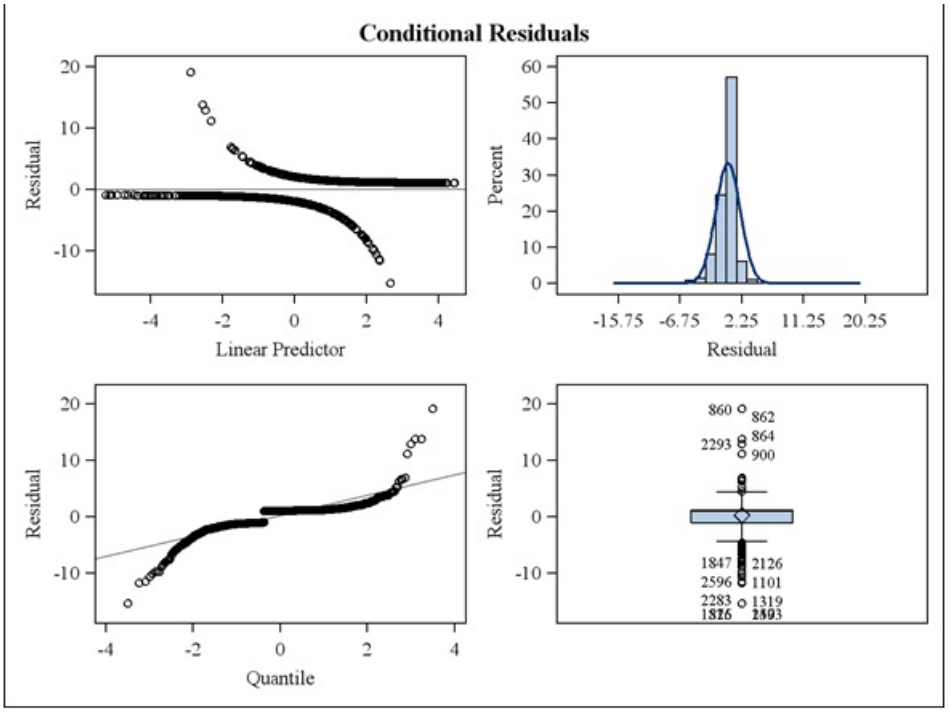


Figure 2–12: Conditional Residuals for C-A Model

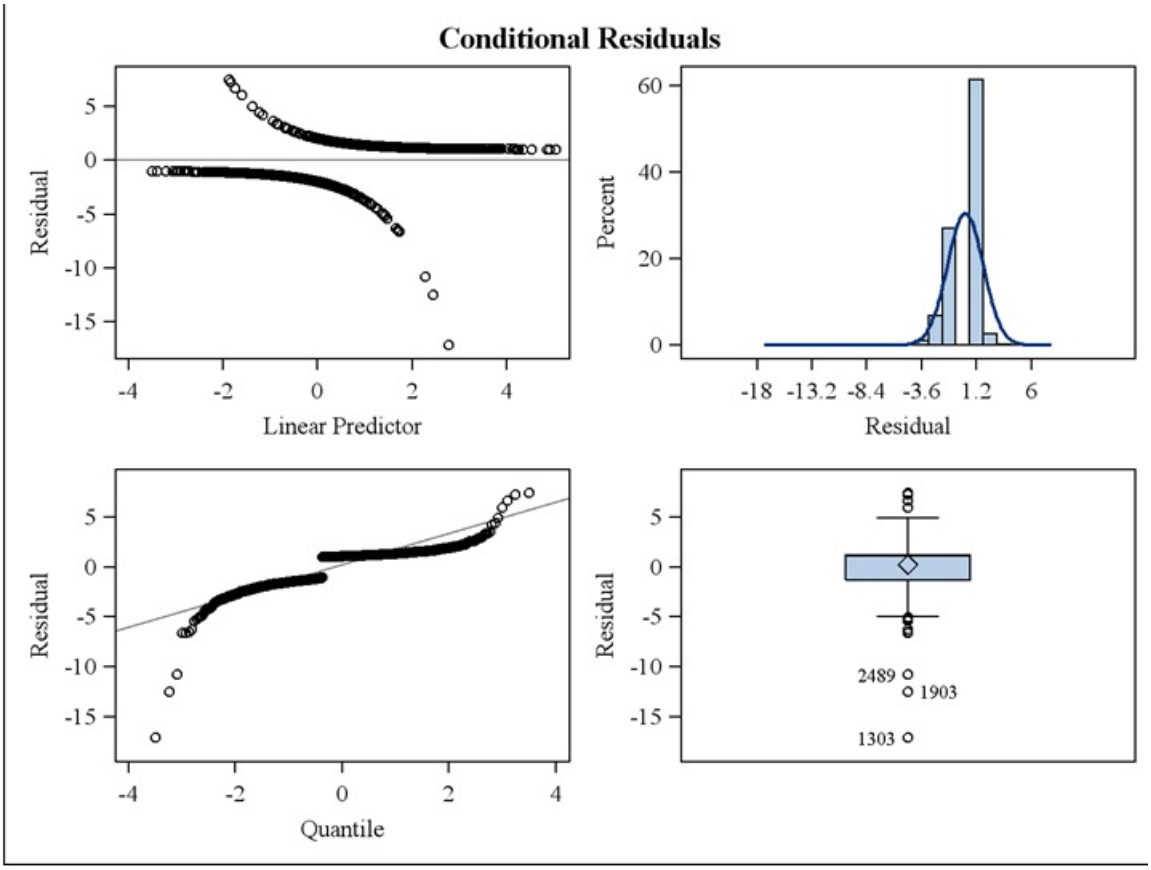
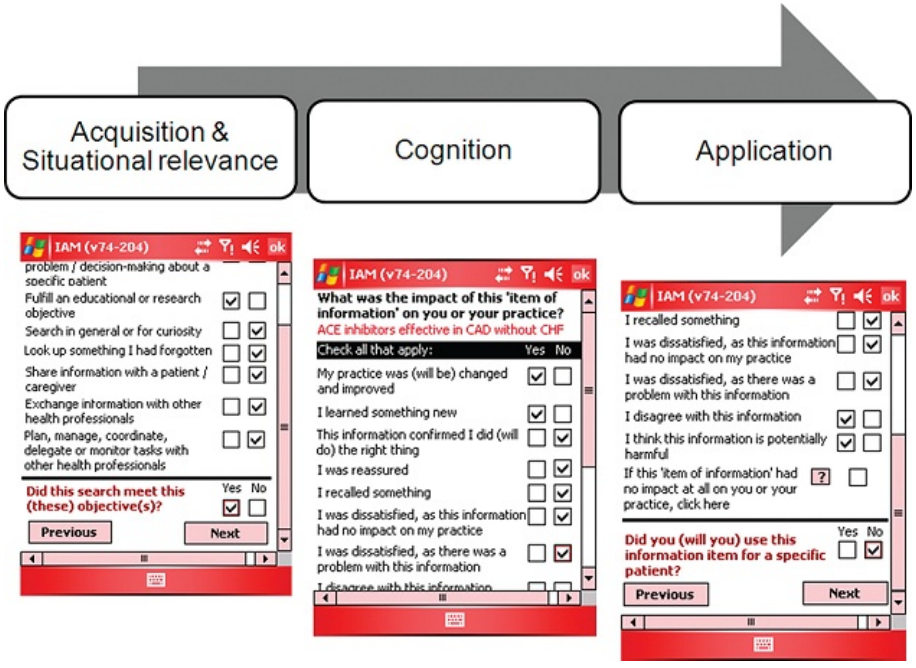


Figure 2–13: Conditional Residuals for A-A Model

2.5 Conclusion

Knowledge Translation is a developing field in the medical sciences that benefits greatly from statistical support, as electronic resources lend themselves very easily to data collection. In these studies, we provided evidence that a subtle change in an IAM response format can have a significant impact on user participation, and that several factors can influence the use of information for a specific patient, such as situational relevance, search objectives and cognitive impact. Although due to the individual-specific nature of the intention behind the studies and so a random effects approach was taken in Study 2, a treatment of the data using GEEs would make for an interesting comparison and is left for future research to investigate.

Appendix A1



IAM 2008 on the handheld computer

IAM 2008

1) Search objective: *Why did you do this search?*

   I. Address a clinical question/problem/decision-making about a specific patient
   II. Fulfill an educational or research objective
   III. Search in general or curiosity
   IV. Look up something I had forgotten
   V. Exchange information with other health professionals
   VI. Share information with patient or caregiver
   VII. Plan, manage, co ordinate, delegate, or monitor tasks with other health professionals

2) Situational relevance: *Did this search meet this (these) search objective(s)?*

   I. Yes/No

3) Cognitive Impact: *What was the impact of this item of information on you or your practice?*

   I. My practice will be changed and improved.
   II. I learned something new.
   III. This information confirmed I did (will do) the rig thing.
   IV. I was reassured.
   V. I recalled something.
   VI. I was dissatisfied, as this information had impact on my practice.
   VII. I was dissatisfied, as there was a problem wi this information.
   VIII. I disagree with this information.
   IX. I think this information is potentially harmful.

4) Use: Did you/will you use this information for a specif patient?

   I. Yes/No

IAM 2008

High dose statin reduces cardiac events in pts with high CRP (JUPITER)

Clinical question
In patients with normal LDL cholesterol but elevated C-reactive protein, is a high dose statin effective for primary cardiovascular prevention?

Bottom line
In this study of patients with normal LDL and elevated CRP, use of a high dose statin reduced the risk of death over a 2 year period (NNT = 180). A about $216,000. This study raises many questions. What is the long term safety of lowering LDL cholesterol to 55 mg/dl in otherwise healthy pers benefit of this drug? Can less expensive statin drugs, perhaps at lower doses, provide a similar benefit with less risk? (LOE = 1a)

Reference
Ridker PM, Danielson E, Fonseca FAH, et al. Rosuvastatin to Prevent Vascular Events in Men and Women with Elevated C-Reactive Protein. N En

Study design: Randomized controlled trial (double-blinded)

Funding: Industry

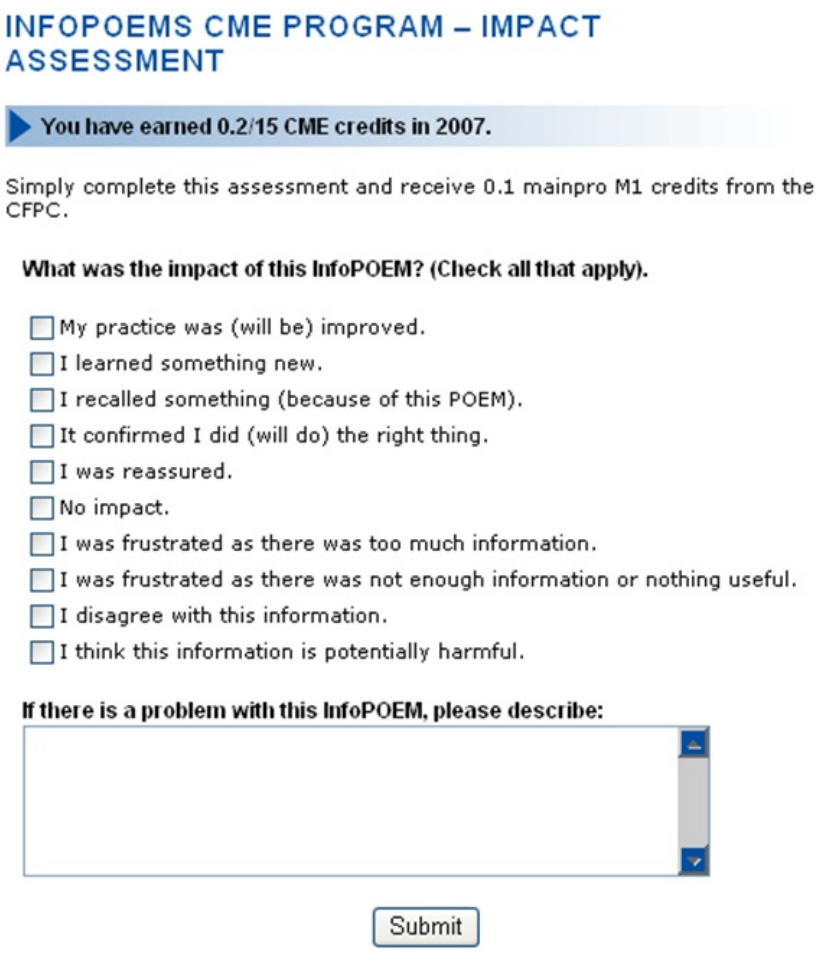
Allocation: Concealed

Setting: Outpatient (any)

Synopsis
The Air Force/Texas Coronary Atherosclerosis Prevention Study found that statins may be effective in patients with normal cholesterol but elevated identified adults with LDL cholesterol < 130 mg/dl and C-reactive protein > 2.0 mg/L. Nearly 90,000 men over age 50 years and women over age 6 excluded due to an elevated LDL (37,611), low CRP (25,993), withdrawal of consent (3948), diabetes (957), hypothyroidism (349), or other reason hormone replacement therapy were ineligible, as were patients with elevated creatine kinase, creatinine, or hepatic transaminases at baseline. T those taking less than 80% of the study drug were excluded. This of course has the effect of making the study drug look more effective than it is in white, mean age 66 years) were randomized to rosuvastatin (Crestor) 20 mg once daily or matching placebo. At each of the annual follow-up visits group (55 vs 110 mg/dl) and the CRP was also significantly lower (~2.0 vs 3.5 mg/L). The study was terminated early after 1.9 years of median foll 1.25 per 100 patient years, p = 0.02). There was a consistent pattern of fewer cardiovascular events for patients taking rosuvastatin, including fewe strokes (0.18 vs 0.34 per 100 patient years, p = 0.002). Patients taking rosuvastatin were more likely to be diagnosed with diabetes mellitus, thou rhabdomyolysis, which occurred in a patient taking rosuvastatin.

Example of a POEM

Appendix A2



IAM Format During S1

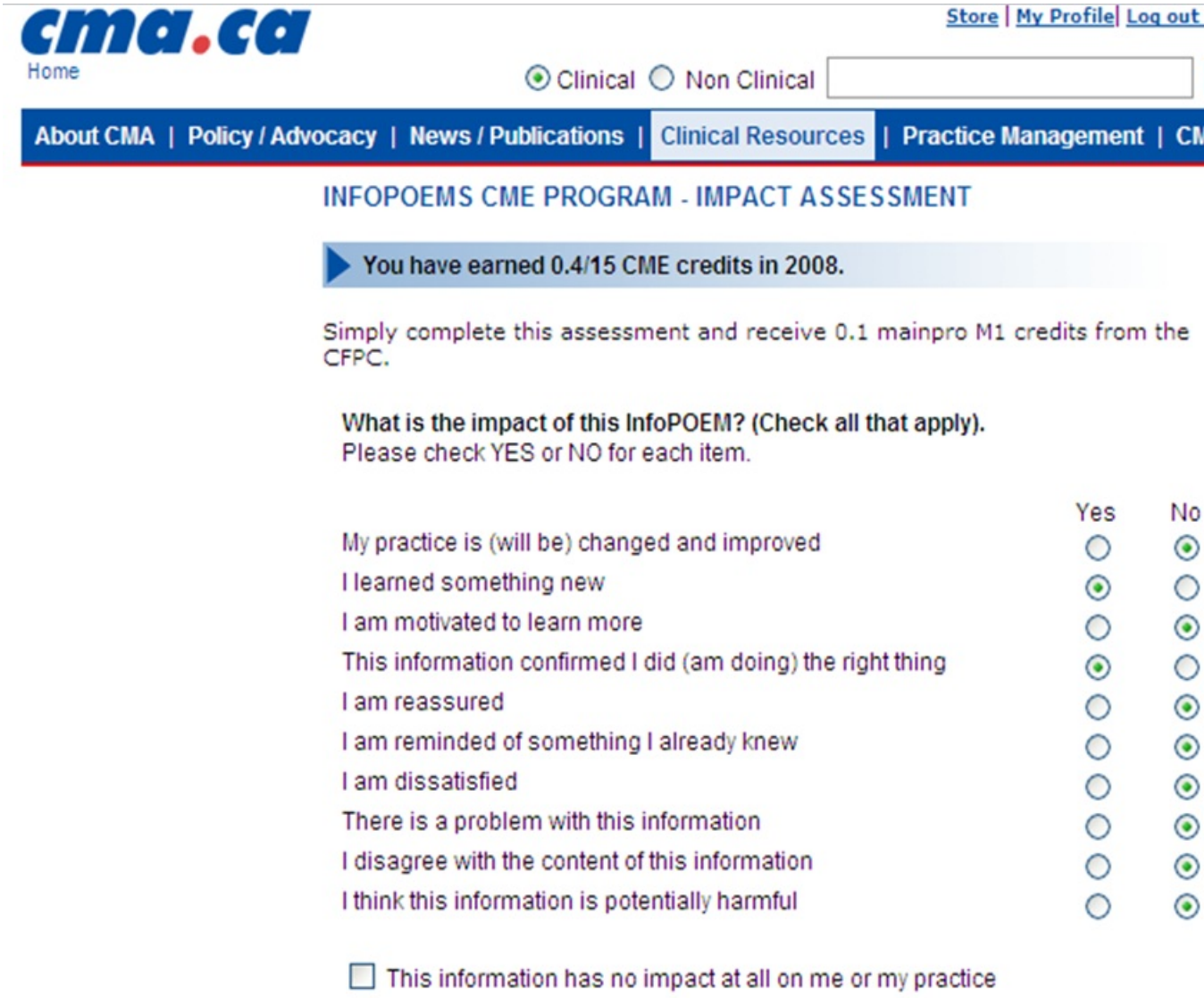


Figure 2–14: IAM Format During S2

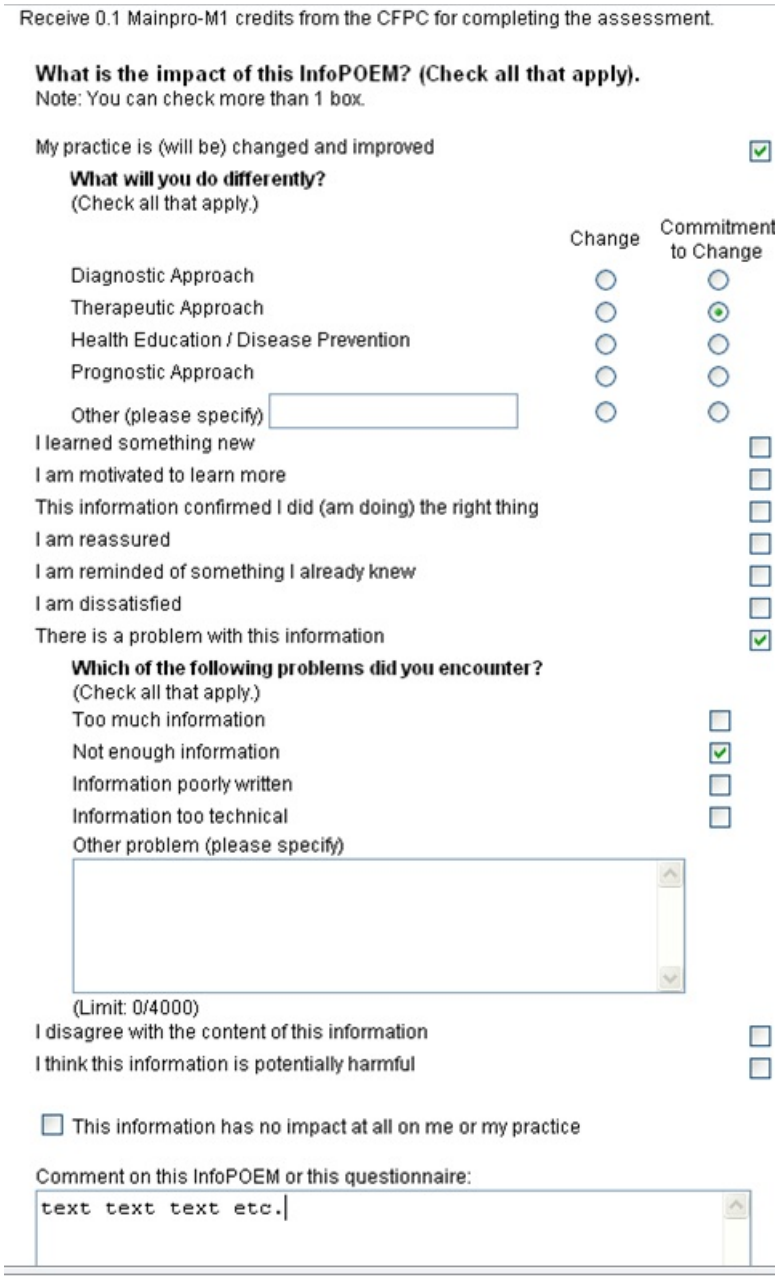


Figure 2–15: IAM Format During S3

References

[1] *Handbook of Research on Information Technology Management and Clinical Data Administration in Healthcare*, chapter 33. IGI Global, 2009.

[2] *Survival Analysis Using SAS: A Practical Guide*, chapter 5. SAS Publishing, 2010.

[3] M. AITKIN, S. N. BENNETT, and JANE HESKETH. Teaching styles and pupil progress: A re-analysis. *British Journal of Educational Psychology*, 51(2):170–186, 1981.

[4] N Bennett. *Teaching Styles and Pupil Progress*. Open Books, London [u.a.], 1975.

[5] S. Bindiganavile Sridhar, P. Pluye, and R.M. Grad. In pursuit of a valid information assessment method. Master's thesis, McGill University, 2011.

[6] N. Breslow. Covariance analysis of censored survival data. *Biometrics*, 30(1):89–99, 1974.

[7] D. G. Covell, G. C. Uman, and P. R. Manning. Information needs in office practice: are they being met? *Ann. Intern. Med.*, 103(4):596–599, Oct 1985.

[8] D. R. COX. Partial likelihood. *Biometrika*, 62(2):269–276, 1975.

[9] David R. Cox. Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society*, B(34):187–220, 1972.

[10] John Crowley and Marie Hu. Covariance analysis of heart transplant survival data. *Journal of the American Statistical Association*, 72(357):pp. 27–36, 1977.

[11] B. Efron. The efficiency of cox's likelihood function for censored data. *Journal of the American Statistical Association*, 72(359):557–565, 1977.

[12] V. T. Farewell and R. L. Prentice. The approximation of partial likelihood with emphasis on case-control studies. *Biometrika*, 67(2):pp. 273–278, 1980.

[13] Harvey Goldstein. *Multilevel statistical models*. Number 3 in Kendall's library of statistics. Arnold [u.a.], London [u.a.], 2. ed edition, 1995.

[14] R.M. Grad, P. Pluye, A.C. Vandal, B. Marlow, S.E.D. Shortt, and L. Hill. Continuing medical education outcomes associated with e-mailed synopses of research-based clinical information. Grant proposal reviewed and not funded in the September 2011 CIHR Open Operating Grants Competition. To be posted at McGill Family Medicine Studies Online http://mcgill-fammedstudies-recherchemedfam.pbworks.com.

[15] Roland Grad, Pierre Pluye, Vera Granikov, Janique Johnson-Lafleur, Michael Shulha, Soumya Bindiganavile Sridhar, Jonathan L. Moscovici, Gillian Bartlett, Alain C. Vandal, Bernard Marlow, and Lorie Kloda. Physicians' assessment of the value of clinical information: Operationalization of a theoretical model. *Journal of the American Society for Information Science and Technology*, 62(10):1884–1891, 2011.

[16] Guang Guo and Hongxin Zhao. Multilevel modeling for binary data. *Annual Review of Sociology*, 26:pp. 441–462, 2000.

[17] Irva Hertz-Picciotto and Beverly Rockhill. Validity and efficiency of approximation methods for tied survival times in Cox regression. *Biometrics*, 53(3):1151–1156, 1997.

[18] Norman Lloyd Johnson, Samuel Kotz, and Narayanaswamy Balakrishnan. *Continuous univariate distributions*. Distributions in statistics. Wiley, New York, NY [u.a.], 2. ed. edition, 1995.

[19] S. D. Kachman. An introduction to generalized linear mixed models. In *Proc. of a Symposium at the Organizational Meeting for a NCR Coordinating Committee on "Implementation Strategies for National Beef Cattle Evaluation*, pages 59–73, 2000.

[20] J. D. Kalbfleisch and R. L. Prentice. Marginal likelihoods based on cox's regression and life model. *Biometrika*, 60(2):pp. 267–278, 1973.

[21] Richard Kay. An explanation of the hazard ratio. *Pharmaceutical Statistics*, 3(4):295–297, 2004.

[22] D. Y. Lin and L. J. Wei. The robust inference for the cox proportional hazards model. *Journal of the American Statistical Association*, 84(408):pp. 1074–1078, 1989.

[23] P. Pluye, R. M. Grad, M. Dawes, and J. C. Bartlett. Seven reasons why health professionals search clinical information-retrieval technology (CIRT): toward an organizational model. *J Eval Clin Pract*, 13(1):39–49, Feb 2007.

[24] T. Saracevic and K.B. Kantor. Studying the value of library and information services: Part i. establishing a theoretical framework. *Journal of the American Society for Information Science*, 48:527–542, 1997.

[25] SAS Institute Inc. *SAS 9.2 User's Guide*, second edition.

[26] S. Shiffman and A. A. Stone. Ecological momentary assessment: A new tool for behavioral medicine research. *Technology and methods in behavioral medicine*, pages 117–131, 1998.

[27] S. E. Straus, J. Tetroe, and I. Graham. Defining knowledge translation. *CMAJ*, 181(3-4):165–168, Aug 2009.

[28] W. Suriyasathaporn, M. Nielen, S. J. Dieleman, A. Brand, E. N. Noordhuizen-Stassen, and Y. H. Schukken. A cox proportional-hazards model with time-dependent covariates to evaluate the relationship between body-condition score and the risks of first insemination and pregnancy in a high-producing dairy herd. *Preventive Veterinary Medicine*, 37(1–4):159–172, 1998.

[29] B. West, K. Welch, and A. Galecki. *Linear Mixed Models: A Practical Guide Using Statistical Software*. Chapman Hall / CRC Press, first edition, 2006. ISBN 1584884800.