Medical Decision Making

http://mdm.sagepub.com

Time-Tradeoff Utilities for Identifying and Evaluating a Minimum Data Set for Time-Critical Biosurveillance

Jason N. Doctor, Janet G. Baseman, William B. Lober, Jac Davies, John Kobayashi, Bryant T. Karras and Sherrilynne Fuller

Med Decis Making 2008; 28; 351 originally published online May 13, 2008;
DOI: 10.1177/0272989X08317011

The online version of this article can be found at:
http://mdm.sagepub.com/cgi/content/abstract/28/3/351

Published by:
SAGE
http://www.sagepublications.com

On behalf of:

Society for Medical Decision Making

Additional services and information for *Medical Decision Making* can be found at:

Email Alerts: http://mdm.sagepub.com/cgi/alerts

Subscriptions: http://mdm.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations http://mdm.sagepub.com/cgi/content/refs/28/3/351

Time-Tradeoff Utilities for Identifying and Evaluating a Minimum Data Set for Time-Critical Biosurveillance

Jason N. Doctor, PhD, Janet G. Baseman, PhD, William B. Lober, MD, Jac Davies, MS, MPH, John Kobayashi, MD, MPH, Bryant T. Karras, MD, Sherrilynne Fuller, PhD

Background. *Researchers and policy makers are interested in identifying, implementing, and evaluating a national minimum data set for biosurveillance. However, work remains to be done to establish methods for measuring the value of such data.* ***Purpose.*** *The purpose of this article is to establish and evaluate a method for measuring the utility of biosurveillance data.* ***Method.*** *The authors derive an expected utility model in which the value of data may be determined by trading data relevance for time delay in receiving data. In a sample of 23 disease surveillance practitioners, the authors test if such tradeoffs are sensitive to the types of data elements involved (chief complaint v. emergency department [ED] log of visit) and proportional changes to the time horizon needed for receiving data (24 v. 48 h). In addition, they evaluate the logical error rate: the proportion of responses that scored less relevant data as having higher utility.* ***Results.*** *Utilities of chief complaints were significantly higher than ED log of visit,* $F(1, 21) = 5.60$, $P < 0.05$, *suggesting the method is sensitive. Further utilities did not depend on time horizon used in the exercise,* $F(1, 21) = 0.00$, $P = ns$. *Of 92 time tradeoffs elicited, there were 5 logical errors (i.e., 5% logical error rate).* ***Conclusions.*** *In this article, the authors establish a time-tradeoff exercise for valuing biosurveillance data. Empirically, the method shows initial promise for evaluating a minimum data set for biosurveillance. Future applications of this approach may prove useful in disease surveillance planning and evaluation.* ***Key words:*** *utility theory; public health; bioterrorism; epidemiology; surveillance.* ***(Med Decis Making 2008;28:351–358)***

The purpose of a national biosurveillance program is to implement real-time, nationwide public health event monitoring to support early detection, situation awareness, and rapid response management across public health and care delivery communities and other authorized government agencies.[1] Multiple federal agencies have been exploring the feasibility of developing a minimum data set for biosurveillance and the American Health Information Community (AHIC), an advisory body to the US Department of Health and Human Services, has recommended identifying, implementing, and evaluating a national minimum data set.[2] Table 1 shows several clinical items under consideration for inclusion in the minimum data set. However, work has yet to be completed to establish formal methods for identification and evaluation. Such work is important because a minimum set that does not meet fundamental criteria may not serve its stated purpose well. In this article, we develop a multiattribute utility model for identification and evaluation of a minimum data set for event

Received 2 July 2007 from the Department of Clinical Pharmacy and Pharmaceutical Economics and Policy, University of Southern California, Los Angeles (JND); Biomedical and Health Informatics, School of Medicine, University of Washington, Seattle (JGB, WBL, JD, JK, BTK, SF); Inland Northwest Health Services, Spokane Washington (JD); and Center for Public Health Informatics, School of Public Health & Community Medicine, Seattle, Washington (JND, JGB, WBL, JD, JK, BTK, SF). This work was supported by a grant from the Centers for Disease Control and Prevention, "Surveillance Integration and Decision Support," P01-CD000261-01: Centers of Excellence in Public Health Informatics, Centers for Disease Control and Prevention (Oberle and Fuller, principal investigators). We thank Mark Oberle and John Miyamoto for their helpful comments. We would also like to thank Lynne Steele of the Centers for Disease Control and Prevention for her comments related to this work. Finally, we would like to thank Ning Yan Gu for her assistance in preparation of this article. Revision accepted for publication 8 January 2008.

Address correspondence to Jason N. Doctor, PhD, Clinical Pharmacy and Pharmaceutical Economics and Policy, University of Southern California, School of Pharmacy, 1540 East Alcazar Street, CHP-140, Los Angeles, CA 90089-9004; e-mail: jdoctor@usc.edu.

DOI: 10.1177/0272989X08317011

Table 1 Sample Clinical Items Proposed by the AHIC Biosurveillance Working Group for Use in Biosurveillance

Data Element	Description	Notes
Symptom/illness onset date/time	Documented date/time of symptom/illness onset by triage or clinician	General: Symptom onset typically recorded in free text without any coded value Paper-dominated process at present, but evolving electronic applications make data capture more feasible in the future May require significant reformatting of onset date/time (e.g., 2 weeks ago to actual date)
Chief complaint	Short description, recorded during triage, for seeking care	General: May be text string or coded (e.g., ICD-9 CM[a]) values
Temperature	Recorded temperature during triage	General: (HL7[b] & LOINC[c]) LOINC code for body temperature Where used: OBX-3[d] Feasibility: Temperature routinely collected; for current surveillance system, only 1 of 67 hospitals store electronically
Pulse oximetry	Record pulse oximetry value during triage	General: (HL7 & LOINC) LOINC code for pulse oximetry Where used: OBX-3 Feasibility: Pulse oximetry routinely collected; for current surveillance system, only 1 of 67 hospitals store electronically

a. The International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9 CM).
b. Health Level 7 (HL7).
c. The Logical Observation Identifier Names and Codes (LOINC).
d. In the HL7 nomenclature, OBX-3 is the field that carries the observation identifier.

detection in time-critical biosurveillance. We focus on data elements for event detection and rapid situational awareness valued against competing objectives. The need for a cardinal means to value data is clear for 2 reasons. First, a rank order of data elements with respect to preference is not useful in determining overall value when rankings are combined with empirical data on delay in receiving information and the expected number of usable cases per unit time. In addition, the value of data are being weighed against the cost of collecting data, thus making the analysis of incremental cost per unit value important to decision makers choosing a biosurveillance data set.

The organization of the article is as follows. In the second section of this article, we describe a multiattribute characterization of biosurveillance data. In the third section, we present a notation and structural assumptions behind our model. In the fourth section, we give the main theoretical result of the article. In the fifth section, we present the result of an experiment that demonstrates the characteristics of the model proposed in the fourth section. Finally, in the sixth section, we offer concluding remarks.

CHARACTERISTICS OF DATUM FOR BIOSURVEILLANCE

There are 2 fundamental attributes that affect the value of a single piece of information (datum) for biosurveillance. First, data must be relevant to be of use in a surveillance implementation. For example, data from a pharmacy benefit manager on whether a prescription or claim is for an antibiotic, antiviral, or pain medication together with antiviral may be of considerable relevance, whereas the dispensing cost, average wholesale price, or preauthorization status of a prescription drug is of little or no direct relevance to event detection. Second, the length of the delay in getting data is crucial in biosurveillance, partially because it may prevent or minimize the spread of disease.[3,4] All other attributes being equal, data that arrive quickly are more valuable than data that arrive more slowly. A shorter delay in receiving information about new cases encourages rapid identification of clinical cases and epidemiologically linked case clusters. With respect to operational capacity of health facilities, shorter delays enhance decision-making capabilities regarding the routing of resources during an outbreak.

Two other properties of relevance and delay are worth noting. First, when data are not relevant to surveillance, preferences over varying delays are equal. For example, data on the average wholesale price of prescription orders filled arriving in 12 h are of the same (zero) value if they arrive in 72 h; that is, both data are worthless to biosurveillance because the average wholesale price is not relevant to biosurveillance. Second, after some delay, M, receiving data is of zero value. This means that at some point, data become too old for event detection or rapid situation awareness. Such may be the case when a situation is changing rapidly or when a disease is undergoing a rapid expansion period.

The aforementioned characteristics of biosurveillance data suggest that the value of data may be determined by trading relevance for delay in receiving data. Tradeoff exercises are a common approach to valuing outcomes in cost-effectiveness analysis of health programs[5] and have application in many policy analyses both in and outside of health care.[6] Within the tradeoff exercises, we ask respondents to form equivalences between data of a certain type that arrive in x hours and data that are more relevant to biosurveillance that arrive in more than x hours. Under standard assumptions, we may deduce a utility for each type of data being evaluated.

The approach we propose may be useful not only for identifying and evaluating a biosurveillance minimum data set but also for relating utility to cost of a national biosurveillance program. The AHIC biosurveillance data steering group suggests that a national biosurveillance system (as currently proposed) will cost \$0.04 to \$0.05 per US citizen per year to operate.[2] But without an analysis of the utility of the data set, the cost utility of collecting these data or revising the minimum data set is unknown.

Structural Assumptions

Formally, a pair (r, y) characterizes a piece of datum in the context of time-critical disease surveillance, where r denotes *relevance* to the surveillance task and y denotes *delay* in receiving data. We characterize delay as time losses in hours, so, for example, if y represents a 16-h delay, then $y = -16$ h. The set of all delays is given by $[M, 0]$, where $|M|$ is some maximal time after which data are of no value for outbreak detection or rapid situation awareness. A delay of zero affords immediate data. Data elements do not offer certainty with respect to delay in arrival time and relevance. Delay in receiving information may depend on network traffic or laboratory medicine back logs. The relevance of a data element may depend on the accuracy of that element to distinguish cases by syndrome. Therefore, we may characterize the choices as probability distributions over attribute levels (i.e., gambles). A gamble is a probability distribution over pairs, specifying the probability, p_i, with which different combinations of levels of the attributes obtain. Probabilities, $p_1, \ldots, p_i, \ldots, p_m$ are nonnegative and sum to 1. We assume the decision maker has preference over all pairs of gambles. We assume that preferences over gambles satisfy expected utility theory.[7] Under expected utility theory, a real-valued, cardinal utility function, $u(r, y)$, exists, the expected value of which represents preferences over gambles involving data elements that yield varying delays and relevance to biosurveillance.

THE MAIN RESULT

In this section, assuming expected utility theory and indifference between receiving a data element with relevance A in y hours and a data element with relevance B in x hours if data are of no use after M hours and where $M < x \leq y \leq 0$, we identify preference conditions that are necessary and sufficient for

$$\frac{v(A)}{v(B)} = \frac{x - M}{y - M}, \tag{1}$$

where v is a utility function for the relevance of data elements to biosurveillance. Equation 1 is a means of evaluating the utility of data elements through delay tradeoff ratios. Equation 1 requires that $u(A, y) = v(A)[y - M]$ for delays in $[M, 0]$. This model is characterized by increasing (linear) utility with decreasing delay and imposes that all utilities are equal (worthless) when delays are sufficiently large.

A common assumption employed to simplify measurement requirements in tradeoff exercises involving time is to assume risk neutrality. This yields utility as a linear function of time. For example, this is a key assumption of quality-adjusted life years.[8] Time-tradeoff utilities for calculating quality-adjusted life years are valid under a risk neutrality assumption. Such an assumption is useful here as well. Risk neutrality requires that if relevance of the data element is held fixed, then the decision maker is indifferent between a gamble over delays and the expected delay of that gamble. Formally, risk neutrality imposes that, for any particular level of relevance r, the decision maker is indifferent between the following choices:

1. a probability p of having an x-hour delay of data with relevance r and a probability $(1-p)$ of having a y-hour delay of data with relevance r and
2. $px+(1-p)y$ hour delay of data with relevance r for certain.

Because under this assumption, utility is a linear function of delay, it follows that $u(A,y)=v(A)y+w(A)$, where $v(A)$ is positive and $w(A)$ is a constant that does not depend on delay y. In addition, to risk neutrality, preferences over delays in receiving data must satisfy a condition that characterizes data (no matter how relevant) as being worthless if they arrive too late. For example, the data may be too outdated. Or, an outbreak may have run its course over a period of rapid expansion of disease and is no longer of immediate concern. M is the greatest lower bound because all losses $N \le M$ yield data that are of no use. We call this the M condition.

Under expected utility theory, risk neutrality and the M condition easily follow from equation 1. To derive equation 1 from these conditions, assume the respondent is indifferent between receiving a data element with relevance A in y hours and a data element with relevance B in x hours and that data are of no use after M hours and where $M < x \le y \le 0$, as shown in Figure 1. By this assumption and expected utility, $u(A,y)=u(B,x)$. By the independence axiom of expected utility and because $u(A,y)=u(B,x)$, it follows that a gamble that offers a p chance of (A, y) otherwise (B, M) has the same expected utility as a gamble that offers a p chance of (B, x) otherwise (B, M).[9] By the M condition, we know that for all A and B, $u(A,M)=u(B,M)$, so we may substitute $u(A,M)$ for $u(B,M)$ as we see fit. Therefore, setting $p=1/2$ and by substitution, we may write the following:

$$1/2u(A,y)+1/2u(B,M)=1/2u(B,x)+1/2u(A,M). \quad (2)$$

Multiplying both sides of equation 2 by 2 and rearranging the equation, we see that

$$u(A,y)-u(A,M)=u(B,x)-u(B,M). \quad (3)$$

Because risk neutrality imposes that $u(A,y)$, $u(B,x)$, $u(A,M)$, and $u(B,M)$ equal $v(A)y+w(A)$, $v(B)x+w(B)$, $v(A)M+w(A)$, and $v(B)M+w(B)$, respectively, we may substitute these accordingly in

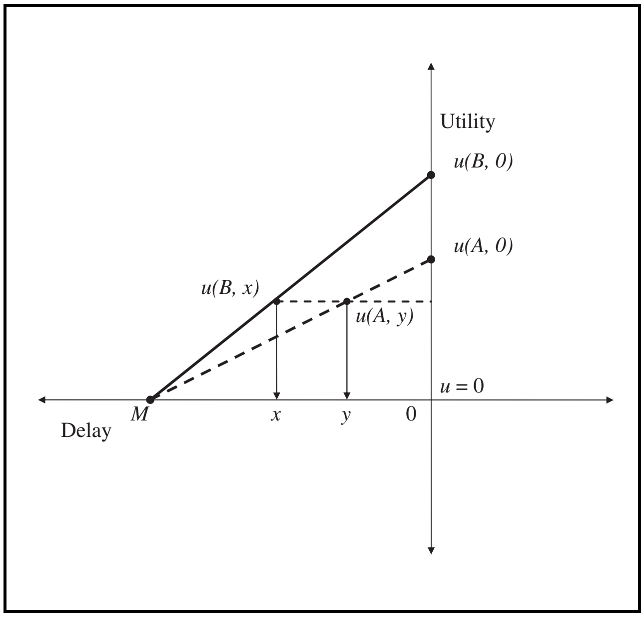

Figure 1 Utility of biosurveillance data elements A *and* B *as a function of delay in receiving the data. The dashed horizontal line represents a tradeoff whereby receiving element* B *at delay* x *is equal in preference to receiving element* A *at delay* y.

equation 3, whereby, through algebraic manipulation, $w(A)$ and $w(B)$ cancel and $v(A)$ and $v(B)$ distribute, yielding,

$$v(A)[y-M]=v(B)[x-M]. \quad (4)$$

Dividing equation 4 by $v(B)$ and also by $[y-M]$ gives us equation 1. Notice also that if v^* is any other function that satisfies equation 1, then $v^*=\alpha v$ for some positive real α. To see this, we note that if v^* satisfies equation 1, then $\frac{v(A)}{v(B)}=\frac{x-M}{y-M}=\frac{v^*(A)}{v^*(B)}$. Let B be fixed. Then, for every element A, $v^*(A)=\alpha v(A)$, where $\alpha=\frac{v^*(B)}{v(B)}$ and is a constant. Therefore, we have established the following.

Theorem 1: Under expected utility, if preferences over data elements arriving at different delays satisfy risk neutrality and the M condition, then there exists a utility function v from data elements to the positive real numbers such that for every relevance level A and B and every delay x and y, where $M < x \le y \le 0$, indifference between receiving an element of relevance A in y hours and an element of relevance B in x hours holds if and only if equation 1 holds. Furthermore, if v^* is any other function that satisfies equation 1, then $v^*=\alpha v$ for some positive real number α.

Theorem 1 establishes preference conditions that make equation 1 a valid formula for representing preferences for data elements. One important feature of equation 1 is that proportional changes in delays should not affect the utility of data element A with respect to data element B. This constant proportional time-tradeoff condition requires that if x is the indifference with respect to M and y in a scenario comparing data elements of relevance A and B, then αx is the indifference for αM and αy in another scenario comparing data elements of relevance A and B. It is easy to see that when M, x, and y are replaced by αM, αx, and αy, respectively, in equation 1, then α cancels because it is a common factor of the numerator and denominator. Later in this article we report the results of an empirical test of this constant proportional tradeoff condition.

A Worked Example

It is worthwhile to consider how utilities may be elicited using the model developed in this article. Consider the following scenario:

> *Scenario 1:* You are responsible for conducting local infectious disease surveillance and outbreak investigations for your health department. You have access to different types of electronic information from the Emergency Department at your local hospital. Data types have an equal number of missing or miscoded values. You have real-time (immediate) access to Chief Complaint (*CC*) data; however, data on Discharge Diagnosis (*DD*) data are delayed in getting to you. Assume that if data arrive after 72 h then they are of no use to you. Please choose a number of hours (less than 72 h) that you would be willing to wait for Discharge Diagnosis data to arrive so that you are indifferent to choosing between these data types.

Suppose the respondent indicates that he would be indifferent between waiting 24 h for the discharge diagnosis data and receiving chief complaint data immediately. By equation 1, $v(CC) = v(A)$ and $v(DD) = v(B)$, $M = -72$ h and $y = 0$. The value $x = -24$ h was elicited from the respondent. By the interval scale properties of expected utility functions, we are free to set $v(DD) = 1$; this gives the utility of chief complaint data on a 0 to 1 utility scale as

$$v(CC) = \frac{-24hrs - (-72hrs)}{0hrs - (-72hrs)} = \frac{48hrs}{72hrs} = 0.67. \qquad (5)$$

Application of equation 1 in equation 5 thus provides means for establishing the value of data to biosurveillance.

EXPERIMENT

Participants

Participants were 23 Washington State and County public health professionals recruited by e-mail to take an online experiment. The average subject was male (65%) and 51.2 y of age ($\pm$9.9). Participants described their position as an administrator, manager, or director (56.5%); epidemiologist or biostatistician (21.7%); public health nurse (8.7%); bioterrorism coordinator (4.4%); infection control (4.4%); and other (4.4%). Twenty-two of the 23 subjects were county employees. One worked for the state of Washington.

Research Design

The experiment described here is part of a larger study on the value of public health data. The experiment is a 2×2 within-subject factorial design. Factor 1, Relevance, was designed to test the hypothesis that the time-tradeoff method is sensitive to detecting utility differences among data elements of varying relevance. This factor consists of 2 levels: 1) chief complaint data, which are relevant to surveillance, and 2) emergency department (ED) log of visits, which are somewhat relevant to surveillance. Both chief complaint and ED log of visits indicate a patient was seen, but chief complaint data indicate that the patient was both seen and also reports the primary presenting problem. Factor 2, Constant Proportional Time Tradeoff, was designed to test the hypothesis that under equation 1, proportional changes in delays should not result in changes in the value of the data elements. Factor 2 consists of 2 levels: 1) data are of no use in 24 h and 2) data are of no use in 48 h. Participants completed 4 tradeoff exercises against a more highly valued data element, discharge diagnosis, in a scenario of the form given in A Worked Example subheading of the Main Result section. Table 2 illustrates the independent and dependent variables within the 4 time-tradeoff questions.

Procedure

Recruitment procedure. E-mails including survey links were sent to a master list of health officers or directors of health at all local health departments in Washington State from our Center for Public Health Informatics. The list was supplemented with a list of local public health practitioners in Washington working in the areas of infectious disease surveillance and/or reporting, who also received e-mails that provided

Table 2 Independent and Dependent Variables for the 4 Time-Tradeoff Questions[a] Presented for Each Receiving Chief Complaint and Emergency Department (ED) Log of Visit 8 Data over 24- and 48-h Time Horizons

	Independent Variables		Dependent Variable
Time-Tradeoff Question	**Data are of no use after. . .**	**I am indifferent to _____ data arriving immediately and. . .**	**Discharge diagnosis data arriving in. . . .**
1	24 h	Chief complaint	x hours
2	24 h	ED log visit	x hours
3	48 h	Chief complaint	x hours
4	48 h	ED log visit	x hours

a. Actual time-tradeoff questions were worded as in scenario 1 reported in the text.

passwords and included survey hyperlinks. E-mails included a request that the messages be forwarded to anyone working in the local health department who worked in the areas of infectious disease surveillance and/or reporting. Follow-up e-mails were sent approximately 1 wk and 3 wk from the date of the initial e-mail.

Elicitation procedure. Participants completed 4 tradeoff exercises against a more highly valued data element, discharge diagnosis, in a scenario of the form given in A Worked Example subheading of the Main Result section. These exercises were part of an online survey written in PHP, using open-source software, PHPsurveyor. The exercises were 1) over a 24-h time horizon (chief complaint, 0-h delay, v. discharge diagnosis, x-hour delay), 2) over a 24-h time horizon (ED log of visits, 0-h delay, v. discharge diagnosis, x-hour delay), 3) over a 48-h time horizon (chief complaint, 0-h delay, v. discharge diagnosis, x-hour delay), and 4) over a 48-h time horizon (ED log of visits, 0-hr delay v. discharge diagnosis, x-hour delay), where x is the dependent variable, or tradeoff, elicited from the respondents.

Statistical Analysis

Statistical analyses of our main hypothesis were completed using the GLM procedure in SAS, version 9.1 (SAS Institute, Cary, NC). We examine results under both the univariate and multivariate approaches to analysis of variance (ANOVA) with 2 within-subjects factors.[10] Also using SAS, we report descriptive statistics on demographics and means and standard deviations of the time tradeoffs. Finally, we examine the logical error rate.[11] As defined by Lenert and others,[11] in this context, a logical error is time-tradeoff scoring of a data element ostensively having lower relevance as being more desirable than an element ostensively having higher relevance (i.e., a logical error occurs when the ED log of visits receives a higher time-tradeoff utility than chief complaint data).

Results

All 23 subjects completed the exercise. The mean ($\pm s$) utilities for the 4 conditions are presented in Table 3. An analysis of logical errors revealed that of the 23 respondents, 4 made a logical error on at least 1 comparison between chief complaint and ED log visits (17% logical error rate). In other words, 4 respondents rated the utility of ED log visits data as more desirable than chief complaint data for at least 1 time horizon (24 or 48 h). However, only 1 of the 23 respondents made this error for both time horizons (logical error rate = 4%). Thus, in total, there were $4 \times 23 = 92$ time tradeoffs given and 5 logical errors observed (logical errors as a percentage of time tradeoffs = 5%). Because previous research has shown that individuals with illogically ordered ratings may bias estimates of mean utilities,[11] the 1 subject who made logical errors on all responses was removed from further analysis.

Analysis of variance results revealed that utilities for chief complaint were significantly higher than ED log of visits, $F(1, 21) = 5.60$, $P < 0.05$, thus supporting that time tradeoffs are sensitive to data element relevance. Further utilities were not significantly different when time horizon was varied; that is, there was no main effect for the time, $F(1, 21) = 0.00$, $P = \text{ns}$, thus supporting that constant proportional time tradeoff is satisfied as predicted by equation 1. There was also no significant interaction between time and relevance factors, $F(1, 21) = 0.40$, $P = \text{ns}$. Multivariate analysis of these data produced commensurate results with the univariate analysis.

Table 3 Mean and Standard Deviation by Time-Tradeoff Condition ($N = 22$)

	Data Element	
Time Horizon	**Chief Complaint**	**ED Log of Visit**
24 h		
$\bar{x}$	0.59	0.55
s	0.19	0.27
48 h		
$\bar{x}$	0.61	0.53
s	0.24	0.31

A possible explanation for our observation of no statistically significant effect for time horizon is low statistical power. Therefore, we analyzed (post hoc) the size of a true difference in utility between time conditions necessary to achieve a statistical power of 0.80. Our analysis of variance consisted of 22 subjects, each of whom provided 2 observations per time point; utilities at the 2 time horizon conditions were correlated 0.68; and we observed a pooled standard deviation of 0.25 across time. Assuming also an alpha level of 0.05, an ANOVA repeated-measures effect could be detected with power 0.80 provided the difference in mean utility between the time conditions was equal to one-eighth a utility unit (or 0.125 utils). Low statistical power is a plausible explanation for our finding if one assumes a true difference in mean utility between the time conditions exists and is less than one eighth a utility unit.

CONCLUDING REMARKS

In this article, we have established theoretically that the utility of the relevance of data elements may be elicited via time-tradeoff exercises. Empirically, we have shown that the method has promise for use in evaluating a minimum data set for biosurveillance. The method we propose was sensitive to change in data element relevance and not sensitive to arbitrary changes in time horizon. This supports the idea that hypothetical scenarios will yield sensitive and valid hypothetical utilities. In addition, the overall logical error rate was low among the respondents in our sample. This low logical error rate of total ratings (5%) is much lower than that reported in the health utilities literature (17%).[11] This may be because public health practitioners typically have a high level of education, with most having advanced professional or graduate degrees and detailed domain knowledge. In contrast, patient and general population samples used in health utility studies come from varying educational levels and may have, in some cases, limited health knowledge.

There are several limitations to our proposed model that deserve mention. First, the scope of how we define value is limited. Elements may differ in their degree of informativeness, that is, the ability to reduce uncertainty in classifying syndromic (or diagnostic) cases. For example, although both a white cell count and a blood culture reduce uncertainty in a disease state of an individual, a blood culture may do so to a greater extent. We have not explicitly modeled informativeness simply because informativeness depends heavily on the event in question. Furthermore, relevance captures a subjective aspect of informativeness that may go into the characterization of a data element. For example, we found that ED discharge diagnosis was more relevant than chief complaint, which was more relevant than ED visit counts. Clearly, ED discharge diagnosis is most relevant because it is most informative, and ED visit count is least relevant because of its purported least informativeness. Second, the validity of the utility estimates for relevance depends on the descriptive validity of the model. To the extent that respondents violate assumptions or predictions of the model, the model will not yield valid utilities. We tested 1 such prediction, constant proportional time tradeoff, which was satisfied. Although this result is encouraging, it does not guarantee the validity of the estimates. An alternative explanation for our finding that utilities do not differ across different time horizons is low statistical power. This explanation becomes more likely when one assumes that the differences in mean utility between the time conditions are small (less than one-eighth a utility unit). Finally, the model we propose is applicable only to time-critical surveillance of outbreaks with rapid case-doubling times or rapid initial infection rates (e.g., smallpox, severe acute respiratory syndrome [SARS], avian flu, and anthrax). Some troubling infectious disease, such as AIDS, have a much slower case-doubling time and would violate the M condition because even after very long delays, case information would be of some value.

In this article, we have conceptualized the act of selecting a minimum data set for time-critical public health biosurveillance systems as a form of decision making under risk. We address here the problem of selecting and evaluating data elements for surveillance. Expected utility theory provides a framework for characterizing the utility or value of a data element for time-critical surveillance purposes. We have shown that data element utilities may be elicited through tradeoff exercises between delays and relevance and that these tradeoffs are sensitive to relevance and may not be affected by the time horizon chosen in the exercises. Future applications of this approach may prove useful for disease surveillance.

REFERENCES

1. Office of the National Coordinator for Health Technology. Harmonized Use Case for Biosurveillance. Washington (DC): Department of Health & Human Services; 2006.

2. American Health Information Committee. Biosurveillance Working Group Archives, Data Steering Group recommendations letter, October 31, 2006. Available from: URL: http://www.hhs.gov/healthit/ahic/materials/meeting10/bio/BDSG_Rec_Letter.doc. Accessed 18 June 2007.

3. Galvani P, Lei X, Jewell NP. Severe acute respiratory syndrome: temporal stability and geographic variation in case-fatality rates and doubling times. Emerg Infect Dis. 2003;9:991–4.

4. Poutanen SM, Low DE, Henry B, et al. Identification of severe acute respiratory syndrome in Canada. N Engl J Med. 2003;348: 1995–2005.

5. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness Analysis in Health and Medicine. New York: Oxford University Press; 1996.

6. Keeney R, Raiffa H. Decisions with Multiple Objectives: Preferences and Value Tradeoffs. New York: John Wiley and Sons; 1976/1993.

7. von Neumann J, Morgenstern O. Theory of Games and Economic Behavior. Princeton (NJ): Princeton University Press; 1944/1947/1953.

8. Bleichrodt H, Wakker P, Johannesson M. Characterizing QALYs by risk neutrality. J Risk Uncertain. 1997;15:107–14.

9. Jensen NE. An introduction to Bernoullian utility theory. I, II. Swedish J Econ. 1967;69:163-83, 229–247.

10. Maxwell SE, Delaney HD. Designing Experiments and Analyzing Data. Belmont (CA): Wadsworth; 1990.

11. Lenert LA, Sturley A, Rupnow M. Toward improved methods for measurement of utility: automated repair of errors in elicitations. Med Decis Making. 2003;23:67–75.