An individual-based framework for the study of medical error

PETER J. VEAZIE

Department of Community and Preventive Medicine, University of Rochester

Abstract

Background. In the late 1990s, medical error came into focus as a problem to be explicitly acknowledged and addressed. Research on this topic is amassing in the epidemiology of medical error and the system and human factors that contribute to error. In addition, however, an understanding of medical errors in terms of the underlying decision process is needed.

Objective. To present an individual-based framework for the study of medical errors in the context of the decision maker.

Results. A framework is developed in terms of four state spaces: the decision environment, problem, goal, and action spaces. The role of information uncertainty is discussed. The framework is purposefully simple to provide flexibility and options for research-specific extensions, but sufficient structure is imposed to guide understanding and investigation.

Conclusion. Understanding medical error in terms of the proposed framework can guide research and subsequent interventions by illuminating where in the decision process such errors are generated.

Keywords: clinical guidelines, decision making, medical error

The 1999 Institute of Medicine’s report To Err is Human highlighted the prevalence and epidemiology of medical errors [1]. In the year 2000, the report to the president Doing What Counts for Patient Safety highlighted organizational factors as important causes of medical errors [2]. Subsequent interest in understanding and mitigating errors has continued to grow [3–5]. Research has focused on characterizing errors and understanding how organizational factors interact with human factors to facilitate and mitigate errors [6,7]. Such a focus is warranted [8], and frameworks for this analysis have been provided [e.g. 9–15]. Moreover, research from this perspective has been fruitful [16–24].

The human factors’ approach to error is based on understanding how human response moderates environmental and outcome covariation [25,26]. The purpose is to inform the design of task environments that minimize adverse consequences of human factors [27,28]. For example, color is symbolic for humans; to minimize errors associated with mistaken interpretations of color schemes, a task environment should be designed to adhere to color-coding standards. In the United States, humans expect valve handles to rotate counterclockwise to increase flow; hence, the design of such items should correspond to this expectation.

However, the explanation of error in terms of the interaction between environmental and human factors does not fully replace the explanation in terms of the decision maker. Understanding the role of the decision process in the generation of medical errors remains an important research effort, because in the context of health care there may remain a component of human judgement amenable to design constraints yet to be identified, and human judgement cannot be completely designed away. On the basis of the presented framework, research can locate medical errors in the decision-making process and identify additional human factors amenable to control via environmental design or develop interventions on the judgement process.

A framework for medical error

In this paper, the means by which a decision maker achieves an objective is framed in terms of relationships between four state spaces: (i) a decision environment, (ii) a problem space, (iii) a goal space, and (iv) an action space [29]. These are connected by three operators: (i) a problem-selection operator that links the environment to identified problems, (ii) a goal-setting operator that links identified problems to goals, and (iii) an action-selection operator that links goals to actions. Together, the state spaces and operators compose the dynamical system depicted in Figure 1.

The decision environment specifies the physical and social conditions (including the patient state) of a decision-maker, what Newell [30] calls the task environment. In a broad sense, the environment of an entity is the universe with the exception of the entity itself; hence, each entity’s environment is the only one that excludes itself and is therefore unique. Such an expansive definition is seldom useful; more specialized definitions are usually invoked. In the present context, the decision environment comprises a subset of the general environment that informs the decision task. Specifically, the decision environment comprises those aspects of the general environment that can logically vary to constrain or produce changes in the decision. A useful operational definition requires specifying...
the factors that can impact the decision process. This not only includes informative patient characteristics but constraints such as policies, formularies, or physical barriers as well.

The problem space is the set of statements that connote expected undesirable trajectories in the environment, partitioned by semantic equivalence. For example, the statement ‘The problem is that the patient’s glycated hemoglobin level is too high’ connotes that the physician believes the patient is on a trajectory whereby their glycated hemoglobin (A1c) levels indicate undesirable consequences. In a broad sense, the problem space is the set of meaningful substatements that can complete the phrase ‘The problem is...’ regardless of how appropriate or applicable the phrase is.

The goal space comprises the set of statements that index possible subsets of the environment, partitioned by semantic equivalence. Therefore, to satisfy a patient-oriented goal is to redirect a patient to a state in the environment space indexed by the goal statement. For example, a goal may be to ‘obtain patient A1c level information’ or ‘bring the patient’s A1c levels within the range of 6 to 7%’.

The action space comprises the set of statements describing possible actions that a decision maker may take, partitioned by semantic equivalence. For example, an action may be ‘Send patient to get A1c lab test’ or ‘Start the patient on metformin.’ Because it is possible that numerous statements have the same semantic content, the problem, goal, and action spaces are equipped with semantic equivalence classes to partition each space according to the relevant aspects of the framework. For each space, it is the semantic content of a statement that matters, not the specific phrasing.

Operationally, for most applications regarding error, the components of these spaces need not be explicitly defined. As detailed below, only subsets comprising the criteria for determining errors need be explicitly stated. Any statement of a problem, goal, or action need only be compared with the equivalence class associated with the criterion; if a statement has the same meaning as the criterion, then the corresponding problem, goal, or action is not an error, otherwise it is an error. The operational advantage is that only a small subset of each space need be identified: the statements representing appropriate problems, goals, and actions.

The problem-selection operator, which links the decision environment to the problem space, extracts information from the environment, constructs a mental representation, and identifies a corresponding problem. The goal-setting operator, which links the problem and goal spaces, identifies a preferred goal corresponding to a given problem. The action-selection operator, which links the goal and action spaces, identifies a preferred action. Each operator is influenced by the decision environment as well as the decision-maker’s experience and expertise; hence, each operator includes an interpretation component that integrates environment information into the operator’s task, which may include patient clinical information as well as constraining information such as guidelines, policies, and legalities.

**A definition of error**

Most definitions of error include those attributable to incorrect intentions and those attributable to unintended actions (e.g. [1,2]; see reason [8] for an explication of these types). Both are captured by the definition presented here. Formally, error is defined here as a relation on a set of elements relative to a criterion such that an identified problem, goal, or action is an error if it does not match the criterion (or is not an element of a criteria set). For the framework depicted in Figure 1, such errors can be defined for the sets of problems, goals, and actions. Errors exist for a judgement regarding a given patient if the identified problems for the patient do not match the problem criterion, the selected goals do not match the goal criterion, or the selected actions do not match the action criterion.

I assume there is an appropriate problem description (or set of descriptions) for each environment, an appropriate goal (or set of goals) for each problem description, and an appropriate action (or set of actions) for each goal, which can be many-to-many relations. Statements such as ‘$z$ is appropriate for $x$’ (e.g. ‘the goal $g$ is appropriate for the problem description $p$’) are interpreted to mean $z$ is a criterion in the set of possible $z$’s associated with the given $x$. Given this interpretation, it is evident that each criterion in our framework is derived from the preceding criterion except for the problem criterion, which is derived from the environment. Note that requiring the identification of appropriate problems, goal, and actions is a necessary constraint on the framework because the absence of identifiable appropriate or corresponding inappropriate elements of each space precludes the study of

---

**Figure 1** Patient Care Model. For a given state of the decision environment, the problem-selection operator identifies a corresponding problem, the goal-setting operator identifies an associated goal and the action-selection operator selects an action. The dashed lines indicate the influence of the decision environment on each operator, which includes explicit policy, legal, or system constraints as well as patient-specific information.
error. Moreover, the assumption, there exists appropriate sets of elements, is necessary for informed behavior; in the absence of this assumption, we are left with the uniform random selection of problems, goals, and actions. Consider the case where there are only two possible actions, in which one is bad for the patient and the second is even worse. We must either randomly or deliberately choose. Clearly, if the goal is to minimize harm, the first action is the appropriate one even though it is harmful.

We can define error functions for the operators as well. Any given operator partitions its input space into two sets: one comprising inputs for which the operator maps to an appropriate output and the other comprises inputs for which the operator maps to an inappropriate output (an error). Alternatively, any given input partitions the set of possible operators into error-generating and non-error-generating operators. That is, for a given input, each possible operator either identifies an appropriate output or identifies an inappropriate output (an error). These characteristics of the operators distinguishes the case of an appropriate operator that maps an error in its input to an error in its output from the case of an inappropriate operator that maps a correct input to an error in its outputs.

**Error criteria**

Criteria are arbitrary subsets of the space on which error is measured; their identification depends on the purpose of the error measurement. Criteria may be defined in terms of the same space or in terms of related spaces. For example, criteria for actions may be defined in terms of desired resultant outcomes; hence, an appropriate action is just that action that produces a specified result such as improved health outcome or reduced cost.

If we incorporate uncertainty regarding patient responses and consider a distribution of potential outcomes is associated with a given observed patient and a given action, then one plausible definition of the criterion on the action space is the action that maximizes the expected outcome given the patient state. Another definition incorporates the notion of satisfying [31]—that is, select any action that satisfies a set of constraints. Figure 2 depicts the expected outcomes of three actions (denoted as \(a_1, a_2, \) and \(a_3\)) on a continuously scaled patient state dimension. Criteria specified in terms of the maximum expected outcome are represented by the curve denoted \(a_j\) if the patient’s state is less than \(\alpha, a_2\) if the patient’s state is between \(\alpha\) and \(\beta\) or greater than \(\chi; a_3\) if the patient’s state is between \(\beta\) and \(\chi; \) and \(a_3\), again, if the patient’s state is greater than \(\chi\). If criterion was specified as those actions that produce satisfactory outcomes, then for a given patient any action with expected outcome in the satisfactory region would be appropriate (i.e. not an error). As shown in Figure 2, action \(a_1\) is appropriate if the patient state is less than \(\alpha\); action \(a_2\) if between \(\alpha\) and \(\beta\); either \(a_2\) or \(a_3\) if between \(\alpha\) and \(\chi\); and \(a_3\) if the patient state is greater than \(\chi\).

Criteria for the goal space may be defined by extending this logic backward from the action space. Assuming a distribution of potential actions is associated with each goal, the goal criterion may be defined as that goal for which the most likely action is the action criterion defined above. Similar logic can be used to define the criterion for the problem identification space.

**Uncertainty and criterion distributions**

In the preceding exposition, I assume an appropriate response exists for each environmental state. However, in practice, criteria are often based on information from a subspace of the environment; consequently, they may only approximate the appropriate response. This is evident if additional information would change the judgement of what is appropriate. For example, a researcher may define a criterion for action such that given specific patient characteristics and a goal of lowering A1c, the appropriate action is to administer drug \(A\). The researcher would consider as an error a physician’s move that does not include drug \(A\) under these circumstances. However, if the patient has a fatal allergy to drug \(A\), the error criterion does not apply.

Partial information and full information criteria do not necessarily correspond. Variation in appropriate responses is generated by variation in unobserved information and can produce mistaken judgements regarding error when using partial information criteria. This implies practice guidelines that are based on partial information should not be transformed into inviolable practice rules.

**Example**

Dhillon [32] discusses research in medication errors and describes eight types of error: prescribing error, unauthorized-drug error, incorrect-dose error, incorrect-dosage-form error, incorrect-administration-method error, incorrect-dug-preparation error, omission error, and incorrect-time error. Identifiable (in theory) with each error are corresponding responsible persons, whether they be physicians, nurses, pharmacists, or designers of medical technologies. Each type of error can be studied to determine contributing factors; for example, high nurse work loads may increase the instances of some errors. Numerous design elements have been developed to mitigate the consequence of contributing human factors,

**Figure 2** Expected outcomes for three actions \((a_1, a_2, \) and \(a_3\)) across potential patient health states.
for example, computerized physician order entry forms can improve readability of prescriptions and automated dispensing devices promise more accurate filling of prescriptions. Nonetheless, despite increased design efforts, personal judgement remains important in determining medication. For example, prescription error includes the wrong drug and dosage selection; these are decisions that require provider judgement. This paper provides a framework within which models of judgement and behavior can be implemented to better understand how errors can occur at the individual level in the presence of existing error reduction design constraints.

Implementation

One study design, as an initial stage of a research program to locate error within the individual-based framework, is as follows: provide patient case information to a sample of physicians; have the physicians identify corresponding problems, goals, and actions; and then have a panel of experts judge whether errors occur. In this design, the error criteria are implicit in the measurements derived from expert judgement; however, the presented framework is amenable to models and methods that include error specifications from clinical evidence and guidelines as well as the inclusion of specific expert systems and decision support algorithms. Indeed, error criteria are arbitrary as to their source and definition.

The experts judge whether the identified problem, goal, and actions are appropriate. From this information, we can infer an operator is not an error if its input and outputs are not errors, and we can infer the operator is an error if its input is not an error but its output is an error. Consider a case in which a physician is given patient information and selects a corresponding problem, goal, and action. The panel of experts then judge the identified problem is appropriate, but both the goal and action are not appropriate. From this, we can infer the operator connecting patient information to problem selection is not an error, and the operator connecting problem specification to goal selection is an error. We do not, however, know whether the operator connecting goal specification to action selection is an error, because if the input to an operator is an error, then the error status of the operator itself is ambiguous. In this case, we have the experts judge whether the output is an error given the counterfactual presumption that the input is not an error. This is to say the experts judge whether the selected action is appropriate for the selected goal regardless of the fact that the goal was actually inappropriate. For example, a physician may consider a diabetic patient’s problem is that her A1c level is too high; the physician selects a goal of assuring A1c level does not increase and takes the action of ‘no action’ by maintaining current medication levels. Experts may judge the problem is correctly identified; however, perhaps the goal should be to lower the A1c level to below 7% and the action should be to increase medication dosage. From this, we know that the problem identification operator was correct, but the goal selection operator was an error. To discern the status of the action selection operator, the experts judge whether keeping the medications at the current level for this patient would assure the A1c level does not increase: if yes, then the action selection operator is not an error and if no, then the operator is an error.

In this design, measuring the sequence of states in each state space associated with the decisions of a sample of providers is sufficient to locate operators that generate errors within the perception-action cycle (i.e. interpretation, goal-setting, or action-selection operators). A second phase of investigation can then focus on characterizing the specific operator as (i) preference-based, (ii) rule-based, or (iii) pattern matching. Preference based has a structure whereby selection is based on beliefs regarding the consequences. This model requires a high level of deliberation. Rule-based selection has a structure whereby selection is according to an algorithm. Deliberation is less than that in preference-based analysis. Pattern matching represents automated or intuitive responses to information: deliberation is minimal. These forms of cognitive operation are integrated into the problem solving hierarchy presented in Reason’s Generic Error Modeling System [8]. A third phase produces a detailed description of the operator and circumstances leading to error. These last two phases are amenable to process-oriented methods such as verbal protocols [33], in which physicians’ talk aloud while problem solving and the resulting transcripts are coded and analyzed to distinguish the types of operators. Finally, once the errors are identified, patient, physician, and system characteristics can be investigated as predictors of the location and type of error.

Discussion

This paper presents a framework for the study of patient care in terms of related sequences in the four state spaces of decision environment, problem space, goal space, and action space. Operators that identify problems, set goals, and select actions relate these spaces. The purpose is to provide a framework to situate medical errors in the operators underlying patient care decisions. For a given environment, if an identified problem is an error, then the problem identification operator generated the error. Consequently, errors in the problems, goals, or actions spaces are generated by errors in the operators. This suggests the medical error, in so far as they reflect decision behavior, is generated by the operators—the mechanisms by which problems, goals, and actions are identified.

At present, the organizational/system/human factors focused research, and analysis of medical errors often treats an individual-based approach in a pejorative manner, adopting language that implies moral overtones, and assumes such a project centers on blame and individual responsibility (see for example the introductions in [5,10,18]). The presented individual framework should not be viewed in this light. The proposed framework provides a means to better understand viable explanations of error, in terms of decision making and judgements, for the purpose of their mitigation. Indeed, the presented framework can be situated in the popular human factors models that draw on the latent versus active factors taxonomy of error-generating influences [8–10]. Panel A of
Panel A

Figure 3 shows a simplified schematic of the latent versus active factors taxonomy of medical error as a function of high-level and low-level actions. Panel A is an elaboration of the model using the individual-based framework, thereby including the problem and goal spaces. The simple environment, problem, goal, action structure presented in Figure 1 is evident here for both the latent and active factors.

Figure 3 shows a simplified schematic of the latent versus active factors in terms of high- and low-level actions. Latent factors ‘arise from decisions made by designers, builders, procedure writers,ulars, and top level management.’ [10] These are high-level actions that create workplace conditions facilitative of error under particular circumstances; they are usually not foreseen by the high-level decision makers. Active factors ‘are the unsafe acts committed by people who are in direct contact with the patient or system.’ [10] These are based on the decisions of the people most easily identified as proximally responsible for patient care.

As shown in Panel B of the Figure 3, the environment, problems, goals, actions cycle of the presented individual-based framework is an elaboration of the former model. This elaboration allows the researcher-analyst to ‘drill down’ and better understand the psychological genesis of either the high-level actions associated with latent factors or the low-level actions associated with active factors. Consider, for example, nurse work load as a latent factor from the high-level action of personnel scheduling. We may identify heavy work loads as contributing to error-prone environment. If we immediately address the issue by a policy dictate, we may generate unexpected consequences: lowering work loads requires providing less service or hiring more personnel, to afford more person-

References

A framework for the study of error


Accepted for publication 6 April 2006