NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY
SUBMARINE BASE, GROTON, CONN.

REPORT NUMBER 1119

A SEVERITY OF ILLNESS INDEX FOR EVALUATION PRE-HOSPITAL CARE OF SUBMARINERS WITH ABDOMINAL PAIN

by

Morton SOLOMON, M.D., Kendall BRYANT, Ph.D.,
George MOELLER, Ph.D., Bernard RYACK, Ph.D.
David SOUTHERLAND, M.D., and Saul M. LURIA, Ph.D.

Naval Medical Research and Development Command
Research Work Unit 62233N - MM33C30.002-5004

Released by:

C. A. HARVEY, CAPT, MC, USN
Commanding Officer
Naval Submarine Medical Research Laboratory

15 August 1988

Approved for public release; distribution unlimited
A SEVERITY OF ILLNESS INDEX FOR EVALUATING PRE-HOSPITAL CARE OF SUBMARINERS WITH ABDOMINAL PAIN

by

Morton SOLOMON, M.D., Kendall BRYANT, Ph.D., George MOELLER, Ph.D., Bernard RYACK, Ph.D., David SOUTHERLAND, M.D., and Saul M. LURIA, Ph.D.

NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY

NSMRL REPORT NO. 1119

Naval Medical Research and Development Command
Research Work Unit 62233N-MM33C30.002-5004

Approved and Released by:

C. A. HARVEY, CAPT, MC, USN
Commanding Officer
NavSubMedRschLab

Approved for public release; distribution unlimited
SUMMARY PAGE

THE PROBLEM

To develop and validate a Severity of Illness Scale.

THE FINDINGS

Medical outcome measurements -- morbidity and mortality -- are not useful indicators of the severity of an illness. Mortality occurs too infrequently, and morbidity measures are not quantitative. Some other indicator is needed. A Severity of Illness Scale was developed which has two factors, named the Diagnostic and Treatment Factors. The scale scores were applied to submariners hospitalized with abdominal pain. The diagnosed severity of the illness was significantly related to the intensity of the treatment administered, indicating that the scale can be used as a general index of the severity of illness.

APPLICATION

This method of scaling severity of illness has potential value in clinical research and policy-making and can be used to evaluate the influence of environmental factors, such as type of ship, grade, type of therapy, on the outcome of an illness.
Abstract

It is difficult to compare the severity of different illnesses. Medical outcome measurements -- morbidity and mortality -- are not useful indicators. Another indicator is needed. A Severity of Illness Scale has been developed and validated on patients hospitalized with abdominal pain. A maximum likelihood factor analysis was carried out on the following variables: the number of bed days occupied by the patient, the number of diagnoses, procedures, preoperative, postoperative, and convalescent days, and the hospital admissions. This produced a two factor solution which reflected (i) the intensity of the treatment and (ii) the severity of the illness. There was a significant effect of the specific diagnosis on the treatment instituted, which demonstrated the validity of the Severity of Illness Scale. Other uses of the scale are discussed.
Introduction

The Independent Duty Hospital Corpsman on a submarine has far more responsibility than do mid-level practitioners in the general community. Considerable effort has, therefore, gone into developing or adapting computer-aided decision-making (CADM) systems for use on submarines. A CADM system evaluates solutions to a clinical problem using information about a specific patient's history, physical examination, and laboratory findings which are entered into a computer. Rogers, Ryack, and Moeller1 have reviewed 58 such programs. In general, the effectiveness of a diagnostic system depends on the algorithm, the database, the range of the diseases, and the number of signs, symptoms, and tests. Bayesian approaches have been successfully used in systems that address cardiac and gastroenterologic problems, and such an approach is the basis for an abdominal pain CADM that was developed by de Dombal.2

The purpose of such CADM systems is to enhance the ability of the mid-level practitioners to diagnose an illness, thus presumably resulting in better and earlier treatment and improving outcomes. The study of outcomes is the evaluation of end-results in terms of health and satisfaction; it is the only approach that directly measures how well the health care system has worked. In outcome studies, it is assumed that good results derive from appropriate structure and good care.

However, outcome measures of individuals and small groups are usually based on ill-defined opinions and expectations of health practitioners. For large populations, outcomes are defined on the basis of mortality, morbidity, attitude toward the health care system, and compliance with treatment regimens. Both mortality and morbidity are multidimensional. Mortality may be defined relative to all deaths or those specific to a particular hospital or condition. Morbidity may be defined in terms of incidence, prevalence, duration of impairment, health status index, or prognosis.

Outcome as the Appropriate Indicator

Adequate evidence that differences in an independent variable (e.g., a new method of diagnosing abdominal pain) produce significant improvements in health can be provided only by an effect on the outcome of the illness. Outcome is especially appropriate for the evaluation of pre-hospital care because it emphasizes concern for health and achievement of health care objectives with respect to a particular episode of illness. In addition, as an indicator of quality, outcome has the advantage of face validity: good and bad outcomes are culturally defined.3,4
Problems with Outcomes as an Indicator

Parameters such as number of postoperative days and number of hospital admissions for the same or similar problem measure aspects of severity and may be very useful in the evaluation of the effectiveness of a system of health delivery. There are, however, three major problems in using conventional outcomes to evaluate quality of care:

1. A major assumption in outcome assessment is that the outcome measures chosen are affected by medical care and that both favorable and unfavorable outcomes occur frequently enough to be detected and measured. However, the outcomes most frequently used, death or incidence of major complications, may be so uncommon that detections of significant differences between groups require very large samples of patients.

2. "Ultimate" or long-term outcomes such as death or restoration to normal function may occur so late after an illness or specific treatment that timely evaluation of the results is frequently impossible, and outcome assessment may require follow-up information that may be difficult to obtain or inordinately expensive.

3. Commonly used measures are heavily influenced by confounding factors, such as genetic makeup and the physical and social environment that cannot be controlled by medical care.

Nevertheless, any study that uses quality of care as a criterion for making policy decisions will require an outcome measurement, the traditional method of determining the efficacy of medical procedures, techniques, and therapies.

The Need for a Severity of Illness Scale

In evaluating a health delivery system, it is necessary to stratify patients into comparable groups. Severity of illness indices have been used to produce more homogeneous categories of patients with the same diagnosis. "Severity" in medicine relates to differences in the manifestations of diseases or disorders. The differences may be in the presenting clinical signs and symptoms, e.g., myocardial infarction with or without pulmonary edema, or in the outcomes, e.g., died or recovered. Thus, "severity" is used to distinguish between patients presenting with mild or severe manifestations of the same disease (sometimes called "staging") or to distinguish between diseases which are characterized by different outcomes (often in terms of morbidity or mortality). The intent of a severity of illness scale is to
improve the understanding of relationships between treatments and outcomes. It does this by being mathematically more precise in describing a patient or disease and by quantifying an illness. Without a severity scoring system, it is virtually impossible to compare the outcome of patients treated at different medical facilities with the same diagnosis and therapy.

Types of Severity Scales

Severity scales are based either on staging or on outcomes. Severity scales based on staging are described in terms of the specific clinical manifestations of the disease. For example, appendicitis can be scaled in terms of the amount of elevation of the white blood cell count, duration of pain, and presence or absence of rebound tenderness. Staging scales are useful in comparing occurrences of the same disease, but they have not been effective in comparing different diseases because the scales are defined in terms specific to a given disease.

Severity scales based on outcomes are described in general extrinsic terms, e.g., mortality, morbidity, satisfaction, and quality of life. These things are common to all diseases. The present study, therefore, sought to compare the outcome characteristics of dissimilar diseases, such as appendicitis and cholecystitis, on the same scale -- that is, in terms of mortality or of the hospital resources consumed.

Previous Severity Scales

Several scales based on symptoms and findings have previously been developed for illnesses such as neoplasms, burns, head trauma, and heart attack for the purpose of differentiating prognoses. The Johns Hopkins Severity of Illness Scale was developed as an alternative prospective payment system to the DRGs (Diagnosis Related Groups) "to account adequately for severity of illness and, by implication, for the costs of medical care". Horn et al claim, "within each DRG, substantial differences were found in the distribution of severity of illness in different hospitals". The Johns Hopkins scale assigns to each patient after discharge a severity score based on ratings of seven dimensions of the medical record. Using a staging scale of physiological measures, Pollock found a sixfold difference in mortality rates in the pediatric intensive care of nine teaching hospitals, but the severity of illness, not the hospital, was the most important predictor of the outcome.
The Present Study

In this study, a Severity of Illness Scale was developed. The present study followed the lead of previous studies in treating severity as an essential aspect of evaluating the effect of quality of care. In this study, the scores on the Severity of Illness Scale of patients who had been hospitalized with abdominal pain were related their diagnostic and treatment variables.

Method

Data from 7606 patients were used to derive the Severity of Illness Scale. The data were obtained from a pre-existing hospital reporting system developed by the Navy Medical Data Services Center (NMDSC), which maintains workload and morbidity statistics concerning the health of the U.S. Navy and Marine Corps for administrative purposes. NMDSC routinely collects coded reports from in-patient facilities including nonfixed naval and peripheral treatment facilities. The records contain the patient's name, social security number, age, sex, and race; the name and location of the treatment facility; the duty station, geographic location, and type of ship the patient left; the length of service and pay grade; the diagnoses; tests and surgical procedures performed; the pre-operative, post-operative, convalescent leave, and occupied bed days; and disposition type. The diagnoses are coded according to the International Classification of Diseases, 9th Revision, Clinical Modification (ICD9), and the procedures are coded following the International Classification of Procedures in Medicine published by the World Health Organization.

The study reported here was based on data compiled by NMDSC from in-patient facility records of active duty Navy and Marine Corps personnel admitted to Navy or Army hospitals between January 1983 and September 1985 with the diagnoses of peptic ulcer, acute appendicitis, cholelithiasis, calculus of kidney and ureter, pelvic inflammatory diseases, and symptoms of abdominal pain. Pelvic inflammatory disease was included in the data base to study other duty stations, although it is not a diagnosis that occurs on submarines. The specific ICD9 diagnostic codes selected were: 5311, 5321, 5322, 5400-5439, 560, 574, 575, 592, 614, and 789.

Patient Clinical Records

Clinical records, discharge abstracts, and face sheets are the source documents for most studies of medical care. Although their adequacy in describing a case is sometimes an open question, they remain the prime source of information by which to clinically evaluate outcomes of a health care system.
The approach, using hospital discharge records, is appropriate when studying a group of diseases in a large healthy population. The large number of cases and controls that are required to detect a significant difference, and the difficulties in collecting data from many sources (e.g., operational submarines) make hospital records a natural focal point. In these instances, the retrospective collection of patient data is probably more effective and less costly than the prospective collection of research data. It is a natural extension of the age-old practice of recording histories to aid in diagnosis and to increase the fund of medical knowledge.

Procedure

The Severity of Illness Scale was validated against two groups of submariners. They were assigned to a group on the basis of whether or not their submarine was reported to have a CADM program for abdominal pain. For reasons to be explained below, no further distinction will be made between these two groups.

Diagnosis

For most of the analyses, the discharge diagnosis at the first admission was used as the primary diagnosis. The cases in which appendicitis was the second to the eighth diagnosis on the record were considered to be secondary appendicitis and not appropriately grouped with primary appendicitis for the purpose of evaluating the severity scale.

Similarly, the discharge diagnosis at the first admission was the diagnosis that was evaluated even when the primary diagnosis changed in a subsequent admission. As in most retrospective studies, it was impossible to determine from the information available whether the diagnosis at the first admission was an error or whether the new diagnosis was actually a new disease. Thus, if the patient is listed as having cholecystitis after previously being admitted for nonspecific abdominal pain, it may not be possible even with the complete medical chart to distinguish an inaccurate original diagnosis from a new condition. Similarly, appendectomy per se cannot be used to evaluate the severity of a disease. A number of normal appendices are removed during the treatment of abdominal pain and a number of appendectomies are performed incidental to other abdominal surgery.

In cholelithiasis and cholecystitis, the risk of disease changes with age and there may be an apparent association of the disease with other age related factors in the Navy such as rank and pay grade.
RESULTS

Derivation of Severity Scales

In order to develop the most generalized index of severity, all of the outcome variables (preoperative days, postoperative days, convalescent days, occupancy bed days, number of admissions, number of procedures, and number of diagnoses) for the 7,606 patients in the data base were analyzed by means of a maximum likelihood factor analysis and reduced to two factors.

A two factor solution optimally fit the seven variables (chi square = 10.2, df = 8, p = .24), and the first and second factors accounted for 42.5% and 18.8%, respectively, of the variance (Table 1). Orthogonal and oblique rotations of the initial solution yielded similar results. The oblique rotation solution was selected as reflecting the patterns of hospital practice and was used in later analyses. The correlation between the factors was $r = .42$. The correlation for the submariners was $.38$.

Table 1. Correlation of Variables with Factor Scores

<table>
<thead>
<tr>
<th></th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postoperative Days</td>
<td>.94</td>
<td>.26</td>
</tr>
<tr>
<td>Number of Procedures</td>
<td>.70</td>
<td>.31</td>
</tr>
<tr>
<td>Preoperative Days</td>
<td>.56</td>
<td>.29</td>
</tr>
<tr>
<td>Convalescent Days</td>
<td>.55</td>
<td>.22</td>
</tr>
<tr>
<td>Occupancy Bed Days</td>
<td>.36</td>
<td>.23</td>
</tr>
<tr>
<td>Admissions</td>
<td>.36</td>
<td>.93</td>
</tr>
<tr>
<td>Number of Diagnoses</td>
<td>.38</td>
<td>.74</td>
</tr>
</tbody>
</table>

Correlation of the two factors was $r = .42$

This analysis resulted, therefore, in a Severity of Illness Scale with the component variables differentially weighted for the two factors. Table 2 gives the Scale. To obtain the score, the answers for each of the factors is multiplied by the appropriate weight and the results are summed. (Note, however, that the weight for the number of postoperative days for the Factor 2 is negative.) Higher values on the index indicate increasing severity of illness for both of the factors.
Factor 1 is related predominately to postoperative days (POSTOP), number of procedures (TNOS), preoperative days (PREOPD), convalescent days (CONVLD), and occupancy bed days (OBD) and reflects the intensity of treatment. Factor 2 is related to the number of diagnoses (TNOD) and number of admissions (ADMSNS) and reflects the degree of malignity that was diagnosed (see Table 3). Factor 2 has been tentatively named the Diagnostic Factor (DF) and Factor 1 the Treatment Factor (TF).

The total number of occupied bed days (OBD) has low commonality with other variables and loads only modestly on either of the two factor solutions. Thus, bed occupancy does not directly relate to the severity of illness and is not, in itself, a very useful outcome variable, confirming the concepts underlying the use of DRGs.

Table 2. The Severity of Illness Scale

<table>
<thead>
<tr>
<th>Variable</th>
<th>Treatment Factor</th>
<th>Diagnostic Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Preoperative Days</td>
<td>.04895</td>
<td>.06113</td>
</tr>
<tr>
<td>Number of Postoperative Days</td>
<td>.71259</td>
<td>-.11493</td>
</tr>
<tr>
<td>Number of Convalescent Days</td>
<td>.03677</td>
<td>.00730</td>
</tr>
<tr>
<td>Number of Days in Bed</td>
<td>.09113</td>
<td>.09546</td>
</tr>
<tr>
<td>Number of Admissions</td>
<td>.03517</td>
<td>.21638</td>
</tr>
<tr>
<td>Number of Procedures</td>
<td>.14256</td>
<td>.14008</td>
</tr>
<tr>
<td>Number of Diagnoses</td>
<td>.11058</td>
<td>.65512</td>
</tr>
</tbody>
</table>

Application of Scales to Submarine Patients

These two factors of the Severity of Illness Scale were then used to score the intensity of pre-hospital care for the two groups of patients. The Diagnostic Factor score and the Treatment Factor score were calculated for each hospitalized submariner using the weights in Table 2, and re shown in Table 3.

Table 3 shows that the Diagnosis Factor had a small range of variability. The values ranged from 1.12 to 1.65. The mean value for the Diagnosis Factor for all subjects was 1.28.
Table 3. Means and Standard Deviations of the Diagnoses for the Two Groups

<table>
<thead>
<tr>
<th>TREATMENT FACTOR</th>
<th>Group 1</th>
<th></th>
<th>Std. Dev.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appendicitis</td>
<td>2.030</td>
<td>.961</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>Renal Stones</td>
<td>2.389</td>
<td>.659</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Non-specific</td>
<td>1.360</td>
<td>1.157</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>.759</td>
<td>.957</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Total Sample</td>
<td>1.795</td>
<td>1.095</td>
<td>141</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DIAGNOSTIC FACTOR</th>
<th>Group 1</th>
<th></th>
<th>Std. Dev.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appendicitis</td>
<td>1.117</td>
<td>.346</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>Renal Stones</td>
<td>1.583</td>
<td>.719</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Non-specific</td>
<td>1.341</td>
<td>.600</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1.200</td>
<td>.387</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Total Sample</td>
<td>1.281</td>
<td>.593</td>
<td>141</td>
<td></td>
</tr>
</tbody>
</table>
The Treatment Factor, on the other hand, had a wide range of variability (.76 to 3.04) with a mean value for the overall treatment outcome for the diagnoses related to abdominal pain of 1.80. This indicates that there is a wide range of outcome severities, and that, surprisingly, it is more useful in determining the severity of illness.

Comparing the means for the two groups, those for the group 2 are higher for appendicitis, renal stones, NSAP, and other, but no statistically significant difference was found between the two groups when the initial diagnostic factor was controlled for (F(1,131) = 2.01, p = .16).

The Treatment Factor was then used as the dependent variable in an analysis of variance with the Diagnostic Factor as the covariant control. The diagnostic categories of appendicitis, gallbladder and bile duct diseases, renal stones, and nonspecific abdominal pain were analyzed. These categories addressed by the computer-based decision aid form a subset of those used in the derivation of the abdominal pain Severity of Illness Scale.

There is a highly significant statistical difference in the effect of the diagnosis on the Treatment Factor controlling for the Diagnosis Factor (F = 16.333, p < .001) (Tables 3 and 4). This indicates that different illnesses are given different treatments, and it suggests, therefore, that patients with similar medical problems probably get similar treatments. This result also shows that the Treatment Factor can be used as a predictor of the amount of treatment administered for a given diagnosis.

Table 4. Tests of Significance for Treatment Factor Controlling for Diagnostic Factor

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within cells</td>
<td>98.838</td>
<td>131</td>
<td>.754</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regression</td>
<td>16.722</td>
<td>1</td>
<td>16.722</td>
<td>22.164</td>
<td>.000</td>
</tr>
<tr>
<td>Constant</td>
<td>23.806</td>
<td>1</td>
<td>23.806</td>
<td>31.552</td>
<td>.000</td>
</tr>
<tr>
<td>Subgrp</td>
<td>1.523</td>
<td>1</td>
<td>1.523</td>
<td>2.018</td>
<td>.158</td>
</tr>
<tr>
<td>Diagnoses</td>
<td>49.292</td>
<td>4</td>
<td>12.323</td>
<td>16.333</td>
<td>.000</td>
</tr>
<tr>
<td>Subgrp by diagnoses</td>
<td>1.132</td>
<td>3</td>
<td>.377</td>
<td>.500</td>
<td>.683</td>
</tr>
</tbody>
</table>
DISCUSSION

The patients were originally divided into two groups based on whether or not their submarine was reported to have a CADM program for abdominal pain. This breakdown has not been analyzed in this paper for several reasons. Although certain submarines had the CADM, corpsmen could not be forced to use the CADM program; although the program was available, it was not always used. The training the corpsmen received in the use of the CADM was limited, and they were, thus, not very proficient with it. Since the amount of computer time available to the corpsman is somewhat limited, they often did not use it during emergencies simply because they were not proficient with it. On the other hand, there is reason to believe that some of the submarines which "officially" did not have a CADM program, had obtained pirated copies.

These problems have no bearing, however, on the validation of the Severity of Illness Scale. The outcome of a disease depends, to a considerable extent, on its severity, as may the speed and accuracy of diagnosis as well. However, it would not have been possible to compare the different and dissimilar diseases without some sort of severity scoring system. Such a system has been developed in this study.

The two factors of the Scale are potentially useful in quantifying the clinical characteristics of diseases. For example, those patients with renal stones have a high DF but a low TF, whereas those with acute uncomplicated appendicitis have a high TF and a low DF. That is, renal stone disease tends to lead to repeated admissions and complicating and associated problems but few procedures and a relatively "low level" of treatment; uncomplicated appendicitis seems to require few admissions but requires surgical procedures and a "high level" of treatment.

The development of scales measuring the severity of medical problems has been one of the major challenges facing medical science today. Such quantitative scales would be useful in evaluating diagnostic efficiency, refining prognosis, and determining therapeutic efficacy. They are necessary for comparing medical interventions among environments, physicians, hospitals, and health insurance programs.

In the present study, the diagnosed severity of the illness was significantly related to the intensity of the treatment. Since the scale was used with several dissimilar diseases, this demonstrated that the Severity of Illness Scale does serve as a general index of the severity of illness and that it is related to the outcome of an illness.
The Severity of Illness Scale has potential as a research tool. It may be possible to quantify the diagnostic and treatment aspects of many diseases by these factors, and their ratios may become medically meaningful.

DISCLAIMER

"Naval Medical Research and Development Command, Navy Department, Research Work Unit No. 62233N - MM33C30.002-5004. The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of the Defense, or the U.S. Government."
REFERENCES


A severity of illness index for evaluating pre-hospital care of submariners with abdominal pain

It is difficult to compare the severity of different illnesses. Medical outcome measurements — morbidity and mortality — are not useful indicators. Another indicator is needed. A Severity of Illness Scale has been developed and validated on patients hospitalized with abdominal pain. A maximum likelihood factor analysis was carried out on the following variables: the number of bed days occupied by the patient, the number of diagnoses, procedures, preoperative, postoperative, and convalescent days, and the hospital admissions. This produced a two factor solution which reflected (i) the intensity of the treatment and (ii) the severity of the illness. There was a significant effect of the specific diagnosis on the treatment instituted, which demonstrated the validity of the Severity of Illness Scale. Other uses of the scale are discussed.