

USING SAS TO ANALYZE HEALTH STATUS IN INSURANCE CLAIMS

Richard N. Lieberman
Health Data Systems, Inc.

ABSTRACT

Health care utilization and expenditure data extracted from insurance claims are increasingly being used to monitor the health care delivery system. Efforts to contain the incessant growth in health care expenditures and to enhance efficiency have taken many forms, including attempts at various forms of managed care, the development of utilization controls and implementation of provider profiling.

These trends have led to an increase in the desire to monitor the utilization experience of various patient or provider groups to identify inappropriate utilization of services or inefficient providers. In turn, this has led to a demand for relevant case-mix measures that can account for differences in health status between populations of patients. A good case-mix measurement tool can help to identify over- or under-utilization in populations that have the same health care needs. While approaches have been applied to adjust for case-mix differences in the provision of inpatient care, similar approaches have been not been applied on a wide scale in the ambulatory care environment. This paper describes the application of a case-mix tool, the Johns Hopkins Ambulatory Care Group (ACG) system, to analysis of health services utilization and costs. SAS is used to analyze health insurance claims to identify inappropriate utilization after controlling for the health status of the patients. The SAS techniques illustrated are especially useful because:

- No statistical routines are employed-- facilitating the dissemination of the technique to insurance company and managed care plan personnel who are not familiar or comfortable with multivariate regression modeling techniques.
- The system is designed to operate using only the data typically found in insurance claims files.
- The reporting, graphing, and statistical capabilities of SAS provide data analysts with the ability to examine case-mix adjusted utilization data with varying units of analysis.

INTRODUCTION

Health care reform during the last decade was dominated by government initiated changes in reimbursement (e.g, DRGs) and an employer stampede to self-insurance and alternative delivery systems. Responding to the loss of their traditional underwriting and risk management business on behalf of large groups, and recognizing the refusal of payers to shoulder continuously rising premiums, insurance companies have responded with various managed care strategies such as precertification, utilization management, case management, capitation, primary care gatekeepers, and selective contracting. These initiatives have transformed the insurance industry from its traditional role as a risk-taker and claims adjudicator to:

- Managers of health care rather than check-writers.
- Partners with providers of health care.
- Vendors of raw expertise-- i.e., marketing, administrative, and claims processing.

The modern insurance industry will be active in provider network selection and management, managing the care of individual subscribers, monitoring utilization patterns, maximizing quality and outcomes, evaluating the introduction and diffusion of new technology and services. These new functions demand the development of information systems that allow an insurer to monitor the behavior of individual patients and providers.

Traditional insurance company information systems functioned primarily to pay claims. Information from these systems was often limited to aggregate reporting necessary for overall financial management of the insurance company. It has become apparent that the ability to profile providers on utilization, cost and quality parameters is essential to the bottom line as well as to the industry's ability to compete in the managed care market.

Provider profiling is a computer-based management technique to monitor resource consumption and quality of care through an examination of patterns of health care delivery. Profiles can be constructed for individual practitioners, or groups of practitioners, defined by hospital, geographic area, specialty, type of practice or other characteristics.

Given the recent changes in the health care environment, today, a successful insurer must apply claims data to profiling activities. It is becoming apparent that an insurer's ability to profile providers on utilization, cost and quality parameters is essential if the company is to compete in the managed care marketplace. This requirement will be further underscored if US health care reform follows its present trajectory.

Using provider profiles to make management or financial decisions is complicated by the fact that one must adjust for differences in the underlying health status or case-mix of the patient populations of interest. Almost universally, when a difference in patterns of resource use across two or more patient or provider groups is documented, the issue of the comparability of the populations' case-mix is raised. For example, without taking health status into account, a provider profile of two managed care gatekeepers indicating a twenty percent difference in overall resource consumption, cannot conclude that this difference results from inappropriate provider behavior. It is possible that the patient population comprising the higher use group is markedly less healthy than the other population, and requires more intensive health care resources.

It is apparent that provider profiles should ideally be "normalized" or adjusted to account for differences in the clinical case-mix of the patient population served by different providers. This normalization process will eliminate (or at least reduce) the confounding effect of the patient group's underlying health status. Today, outside of the inpatient setting, few managed care plans normalize their profiles for any factors beyond the age and gender of the members.

THE JOHNS HOPKINS ACG CASE-MIX SYSTEM

The Ambulatory Care Group (ACG) case-mix system was developed by a team of researchers at the Johns Hopkins University to provide a conceptually simple, statistically valid, and clinically relevant measure useful in predicting the need for health care services. ACGs are based on the premise that a measure of a population's "illness burden" can help explain variation in health care resource consumption. ACGs represent a method for categorizing persons based on their age, gender, and diagnoses assigned by their providers during their contacts with the delivery system.

The original development and validation of the ACGs relied on the analysis of computerized encounter/claims data obtained from five large "enrolled" population groups (total sample size 150,000). Four of the populations came from HMOs; the fifth from the Medicaid program of the State of Maryland.

ACGs represent a methodology for clustering ICD-9-CM diagnostic codes. The overall theoretical goal of ACG assignment is to cluster similar conditions based on their expected impact on health care services resource consumption. The first step requires the assignment of ICD-9 codes to one of 34 clusters known as Ambulatory Diagnostic Groups or ADG's. The assignment of an ICD-9 code to an ADG is based on statistical analysis and the following clinical criteria:

- Expected persistence/recurrence of a condition over time;
- The likelihood the patient would make a return visit to continue treatment;
- The likelihood a specialty referral would be required;
- The expected need and cost of diagnostic and therapeutic procedures associated with a condition;
- The likelihood the patient would require hospitalization for a condition during the near term; and
- Likelihood it would result in either short or long term disability.

During a single year, a patients' diagnoses may place them into anywhere from 1 to 34 distinct ADGs. That is, the ADGs are not mutually exclusive and there are potentially as many different unique ADG combinations as there are combinations of the 34 categories (i.e., over a million). In order to develop a system for practical use, it was necessary to collapse this huge number of possibilities into a manageable number.

The DRG development process served as the general model for the reduction of the data into ACGs. Using actual claims data to estimate the resource consumption in each ADG, statistical variance reduction techniques, and the incorporation of age and sex, resulted in the assignment of every ADG combination to a single ACG.

Ultimately, each person was categorized into 1 of 51 mutually exclusive ACGs. All of this is performed automatically by a software program that operates in the IBM-compatible PC, IBM-compatible mainframe or UNIX environments.

The ACG system was intended as a methodology that could be used to measure the health status or "disease burden" of insured populations. It was designed: 1) as a mechanism for normalizing utilization management profiles, 2) to help adjust capitation rates across providers or sites; 3) to assist in the adjustment of risk to compensate for potential biased-selection across insured populations; and 4) as a general adjustment tool for a broad range of claims-oriented quality improvement, management and evaluation research activities.

OVERVIEW OF INDIRECT ADJUSTMENT

There are numerous mathematical and statistical adjustment methods that can be employed to apply ACGs to provider profiling. Although multivariate regression models are the technique most often employed in research settings, such statistical rigor is often not appropriate or necessary in the operational environment of an insurer or managed care organization. The provider profiling example described in this paper employs a nonstatistical mathematical technique known as indirect adjustment to adjust for the patient's health status. Although not as elegant as the use of multivariate statistics, indirect adjustment will yield results that are similar to those obtained from multivariate regression models.

The SAS programs presented in this paper implement the epidemiologic technique known as indirect adjustment to control for health status (ACG) when comparing utilization rates across gatekeepers or delivery sites. A SAS program and several SAS macros will produce resource consumption reports and graphs comparing utilization with and without factoring in the health status of patients. For a detailed description of the mathematics involved in performing the ACG adjustment, please see appendix 2.

USING SAS TO PERFORM ADJUSTMENT

The ACG Adjustment System is a group of SAS macros that allow users who are not SAS experts to produce utilization profiles that are adjusted for health status. All programs were developed using SAS version 6.04 under PC-DOS. The programs are also "backward" compatible with SAS version 5.18 for IBM-compatible mainframe computers running MVS and CMS, as well as versions 6.06/6.07/6.08 on IBM mainframes and IBM-compatible PCs under Windows 3.1. A complete listing of the ACG Adjustment program is provided in Appendix 1.

Using the ACG Adjustment Program Module requires access only to the BASE/SAS module of SAS. The SAS/STAT (statistics) module is not required. Also not required is the SAS/GRAPH module, but its use will result in the production of presentation quality charts.

PREPARING THE INPUT DATA

The input to the program is a SAS dataset containing one record for each patient. The records can be created by joining the resource consumption measures to be analyzed to the output records from the Johns Hopkins ACG Grouper software. At a minimum, the dataset must contain the following data elements:

- Patient identifier
- The ACG assigned to the patient
- A variable containing the delivery site or gatekeeper identifier (or other "stratifier")
- Resource consumption variable(s)

Figure 1 displays an example SAS Proc Contents output for the input dataset.

The ACG Adjustment program will generate ACG-adjusted expected resource consumption reports for up to

four variables. Resource consumption measures can include total charges, ambulatory charges, laboratory or radiology charges, or ambulatory visits. In figure 1, we see that two resource consumption measures are included in the file, TOTCHRG (total charges) and AMBCHRG (ambulatory charges). The user may include from one to four resource consumption measures for analysis by the program. The input file must have at least one stratifying variable, such as PCP identifier, delivery site, etc.

Figure 1
Contents Procedure for Input SAS Dataset

CONTENTS PROCEDURE						
Data Set Name:	E.ANALYSIS			Type:		
Observations:	29248			Record Len:	34	
Variables:	7					
Label:						
-----Alphabetic List of Variables and Attributes-----						
Variable	Type	Len	Pos	Format	Label	
ACG	Num	3	4		AMBULATORY CARE GROUP	
AGE	Num	3	115		PATIENT'S AGE	
AMBCHRG	Num	8	133	7.2	TOTAL AMBULATORY CHARGES	
TOTCHRG	Num	8	141		TOTAL CHARGES	
RECIPID	Char	11	119		PATIENT IDENTIFIER	
SEX	Char	1	118		GENDER	
SITE	Char	1	149		GATEKEEPER OR SITE IDENTIFIER	

RESULTS

Using a sample of claims provided by John Alden Life Insurance Company, the utilization profile of 13,000 enrollees was compared across geographic location. Table 1 displays the results of profiling five geographic regions. Column four of this table shows that region 01 appears to be consuming 21 percent more resources than the company-wide average. Looking at column six, we should note that even after controlling for the health status of the patients in region 01, consumption is still 19 percent higher than the average. Since the ACG adjustment controls for the patient's health status, we should conclude from this finding that some other factor(s), perhaps provider behavior, are responsible for this overutilization of health care resources.

In contrast to the findings for region 01, the ACG-adjusted utilization profile for region 05 reduces the 13 percent overutilization level to the company-wide average (actually one percent below average). This suggests that the patient's in region 05 are sicker than the average patient, and the observed overutilization is justified by actual health care needs.

Table 1

COMPARISON OF ACTUAL VERSUS EXPECTED UTILIZATION: TOTAL CHARGES					
Region	Average Actual Utilization	Company-wide Average	Ratio of Actual to Overall Average	ACG-Adjusted Expected Utilization	Ratio of Actual to Expected Utilization
01	\$1594.37	\$1313.51	1.21	\$1344.82	1.19
02	\$1104.02	\$1313.51	0.84	\$1244.89	0.89
03	\$1274.25	\$1313.51	0.97	\$1399.89	0.91
04	\$1134.90	\$1313.51	0.86	\$1185.05	0.96
05	\$1482.21	\$1313.51	1.13	\$1490.27	0.99

CONCLUSION

The Johns Hopkins ACG case-mix system represents a new population-oriented method to adjust for case-mix differences. Unlike inpatient case mix tools which focus on a single encounter, such as DRGs, ACGs focus on the health status of a patient over time as the means to gauge the need for health care resources.

Insurance companies and managed care organizations will have to develop straightforward means of producing utilization profiles from claims data. This paper demonstrated that BASE/SAS and commonly available claims data can be used to produce case-mix adjusted utilization profiles without building multivariate regression models or undertaking costly abstraction of medical records.

REFERENCES

Weiner, J., Starfield, B., and Richard Lieberman (1992). "Johns Hopkins Ambulatory Care Groups (ACGs): A Case-Mix System for UR, QA and Capitation Adjustment." *HMO Practice*. 6(1):13-19.

ACKNOWLEDGEMENTS

I am grateful to Debbie Hoppe and Ned Crocker of the Group Actuarial Department of the John Alden Life Insurance Company, who were gracious enough to provide the data that was used for the analysis presented in this paper. Additional thanks to Jonathan Weiner, Dr. P.H., who provided The Johns Hopkins ACG Case-Mix System software and documentation and provides continuing support for the development of ACG implementation tools for the insurance and managed care industry.

CONTACT INFORMATION

Richard N. Lieberman
Health Data Systems, Inc.
5840 Pimlico Road
Baltimore, MD 21209
(410) 367-1500
Internet: 0004279015@mcimail.com

The Johns Hopkins ACG Case-Mix System is © 1990-1992 by Johns Hopkins University, Baltimore, MD, USA. SAS® is a registered trademark of SAS Institute Inc., Cary, NC, USA.

APPENDIX 1 SAS PROGRAM CODE TO PERFORM INDIRECT ADJUSTMENT FOR HEALTH STATUS

```

/*****
This program implements the epidemiologic technique
known as "indirect" adjustment to control for
health status (ACG) when comparing utilization
across gatekeepers or delivery sites. The user is
required to insert that names of the resource
consumption variables that are to be analyzed and
the name of the stratifying variable(s), such as
physician identifier, location, etc. Profiles for
a maximum of four resource consumption measures
can be produced at one time.
*****/

/***** SAS Version 6 ONLY *****/
libname in 'e:\acg';
/*****

/***** SAS/GRAPH Use ONLY *****/
goptions device=HPLJS2; /* HP Laserjet II Printer */
*options device=HPLJL3;
/*****

/***** SAS System Options *****/
options symbolgen mprint nodate nonumber;
/*****

/***** Assign Reports to External Files *****/
** filename report1 'd:\acg\report1.out';
/*****

/***** USER DEFINED DATA ELEMENTS *****/
/* &analysis contains the SAS names for the
resource consumption measures */
%LET analysis = ambchrg totchrg;
/* &infile is the SAS dataset containing
the data to be profiled */
%LET infile = in.nesug3;
/* &report1-4 are the labels for the
resource consumption variables
defined by &analysis */
%LET report1 = Total Ambulatory Charges;
%LET report2 = Total Charges;
%LET report3 = ;
%LET report4 = ;
/* &strata defines the name of the
stratifying variable(s) */
%LET strata = region;

```

```

/* &sdesc is the label for the stratifying
variable(s) */
%LET sdesc = 'Region';
/*****

/***** COMPILE SAS MACROS *****/
%macro words (string,root=w);
%local word;
%global count;
%let count=1;
%let word=%nrquote(%scan(&string,&count,%str(
)));
%do %while(&word^=);
%global &root&count;
%let &root&count = &word;
%let count=%eval(&count+1);
%let word=%nrquote(%scan(&string,&count,%str(
)));
%end;
%let count=%eval(&count-1);
%mend words;

%macro makevar;
%global expect weight plan site adjust charge adjmean
e_ratio a_ratio
acgsite;
%let i=%eval(&count);
%if &i=4 %then %do;
%let acgsite = acgsite1 acgsite2 acgsite3 acgsite4;
%let site = site1 site2 site3 site4;
%let expect = expect1 expect2 expect3 expect4;
%let weight = weight1 weight2 weight3 weight4;
%let plan = plan1 plan2 plan3 plan4;
%let adjust = adjust1 adjust2 adjust3 adjust4;
%let charge = charge1 charge2 charge3 charge4;
%let adjmean = adjmean1 adjmean2 adjmean3
adjmean4;
%let e_ratio = e_ratio1 e_ratio2 e_ratio3 e_ratio4;
%let a_ratio = a_ratio1 a_ratio2 a_ratio3 a_ratio4;
%end;
%else %if &i=3 %then %do;
%let acgsite = acgsite1 acgsite2 acgsite3;
%let site = site1 site2 site3;
%let expect = expect1 expect2 expect3;
%let weight = weight1 weight2 weight3;
%let plan = plan1 plan2 plan3;
%let adjust = adjust1 adjust2 adjust3;
%let charge = charge1 charge2 charge3;
%let adjmean = adjmean1 adjmean2 adjmean3;
%let e_ratio = e_ratio1 e_ratio2 e_ratio3;
%let a_ratio = a_ratio1 a_ratio2 a_ratio3;

```

```
%end;
%else %if &i=2 %then %do;
  %let acgsite = acgsite1 acgsite2;
  %let site = site1 site2;
  %let expect = expect1 expect2;
  %let weight = weight1 weight2;
  %let plan = plan1 plan2;
  %let adjust = adjust1 adjust2;
  %let charge = charge1 charge2;
  %let adjmean = adjmean1 adjmean2;
  %let e_ratio = e_ratio1 e_ratio2;
  %let a_ratio = a_ratio1 a_ratio2;
%end;
%else %if &i=1 %then %do;
  %let acgsite = acgsite1;
  %let site = site1;
  %let expect = expect1;
  %let weight = weight1;
  %let plan = plan1;
  %let adjust = adjust1;
  %let charge = charge1;
  %let adjmean = adjmean1;
  %let e_ratio = e_ratio1;
  %let a_ratio = a_ratio1;
%end;
```

```
%mend makevar;
```

```
%macro initvar(left,right);
```

```
  %let i = &count;
  %do %until(&i=0);
    &left&i = &right&i;
    %let i = %eval(&i-1);
```

```
  %end;
```

```
%mend initvar;
```

```
%macro comprat(left,top,bottom);
```

```
  %let i = &count;
  %do %until(&i=0);
    &left&i = &top&i / &bottom&i;
    %let i = %eval(&i-1);
```

```
  %end;
```

```
%mend comprat;
```

```
  %macro report(i);
```

```
  %local iterat;
```

```
  %let iterat=1;
```

```
  %let i = %eval(&count-1);
```

```
  %let width = %length(&&report&i);
```

```
  %let col = %eval((140-&width)/2);
```

```
  %do %until(&i=0);
```

```
  /***** Doesn't work under version 5.18 *****/
```

```
  %if %EVAL(&SYSVER > 5) %then %do;
    footnote "Report Generated on: &SYSDATE";
  %end;
```

```
  *****/
```

```
  data _null_;
```

```
  set expect2 end=eof;
```

```
  %if &iterat=1 %then %do;
```

```
    file report1 lrecl=140 header=newpage;
```

```
  %end;
```

```
  %else %do;
```

```
    file report1 mod lrecl=140 header=newpage;
```

```
  %end;
```

```
  put @3 &strata @35 site&i @55 plan&i @74
  a_ratio&i
```

```
    @91 adjmean&i @112 e_ratio&i;
```

```
  return;
```

```
  newpage;
```

```
  put _page_;
```

```
  put @47 'COMPARISON OF EXPECTED vs.
  ACTUAL UTILIZATION';
```

```
  put @&col "&&report&i";
```

```
  PUT //;
```

```
  put @35 'Average' @89 'ACG-Adjusted';
```

```
  put @35 'Actual' @54 'Overall' @69 'Actual to
  Overall'
```

```
    @89 'Expected' @106 'Actual to Expected';
```

```
  put @35 'Utilization' @54 'Average' @69 '
  Average'
```

```
    @89 'Utilization' @109 'Utilization' /;
```

```
  return;
```

```
proc sort data=acgsite;
```

```
  by acg;
```

```
run;
```

```
proc sort data=acgmeans;
```

```
  by acg;
```

```
run;
```

```
data acgrpt;
```

```
  merge acgsite(in=in1) acgmeans(in=in2);
```

```
  by acg;
```

```
  label weight&i = 'Overall/Average'
```

```
    acgsite&i = 'Site/Average'
```

```
    &strata = "&sdesc"
```

```
    acg = 'ACG/Category'
```

```
    numpat = "Provider's/ Patients"
```

```
    spopnum = 'Total/Patients'
```

```
  ;;;;;;
```

```
run;
```

```

proc sort data=acgrpt;
  by &strata acg;
run;

proc print data=acgrpt noobs split='';
  by &strata;
  pageby &strata;
  var acg acgsite&i numpat weight&i spopnum;
  title1 "&&report&i";
  title2 "Mean Values and Patient Counts by ACG";
run;

data graph1;
  set expect2;
  response = a_ratio&i;
  respcode = 1;
  output;
  response = e_ratio&i;
  respcode = 2;
  output;
run;

proc format;
  value respfmt 1='Actual' 2='Expected';
run;

proc gchart data=graph1;
  pattern1 v=m3 c=red;
  pattern2 v=s c=blue;
  axis1 label=none value=none;
  axis2 label=('Ratio');
  format respcode respfmt.;
  title1 'Comparison of Unadjusted and Adjusted Ratios';
  title2 "&&report&i";
  vbar respcode/midpoints= 1 2 type=sum
  group=&strata sumvar=response
  patternid=midpoint maxis=axis1 raxis=axis2;
run;

%let i = %eval(&i-1);
%let iterat = %eval(&iterat+1);
%end;

%mend report;

/***** PROGRAM BEGINS HERE *****/
*****
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
*****/

```

```

/*****
** This macro identifies the names
** of the resource consumption variables
*****/
%words(&analysis,root=acgsite);

/*****
** This macro creates the variables
** used internally by this program
*****/
%makevar;

/*****
** Tabulate means for resource consumption
** variables, stratified by delivery site
** and ACG
*****/
Proc Summary Data=&infile missing;
  class &strata acg;
  var &analysis;
  output out = adjust mean = &acgsite;
run;

/*****
** Extracts plan-wide means, ACG-wide
** means, site-wide means, and site/acg
** wide means.
*****/
data acgmeans(keep=&weight spopnum acg)
  acgsite(keep=&strata acg &acgsite numpat &plan)
  sitemean(keep=&strata &site);
  format &weight &acgsite &plan 9.2;
  retain &plan &site;
  set adjust end=eof;
  if _type_ = 0 then do;
    %initvar(plan,acgsite);
  end;
  if _type_ = 1 then do;
    %initvar(weight,acgsite);
    spopnum = _freq_;
    output acgmeans;
  end;
  if _type_ = 2 then do;
    %initvar(site,acgsite);
    output sitemean;
  end;
  if _type_ = 3 then do;
    numpat = _freq_;
    output acgsite;
  end;
  drop _freq_ _type_ ;

```

```

run;

proc sort data = sitemean;
  by &strata;
run;

proc sort data = acgsite;
  by &strata;
run;

/*****
** Joins site-wide means with acg/site-
** wide means.
**
*****/
data acgsite;
  merge acgsite(in=in2) sitemean(in=in1);
  by &strata;
  if in2;
run;

proc sort data = acgmeans;
  by acg;
run;

proc sort data = acgsite;
  by acg;
run;

/*****
  Tabulates ACG expected TOTAL
  utilization, assuming plan-wide means
  apply to acg/site-wide patient counts
*****/
data match;
  merge acgsite acgmeans end=eof;
  by acg;
  %initvar(expect,numpat*weight);
  format &expect 9.2;
run;

proc summary data=match;
  class &strata;
  id &acgsite &plan &site;
  var &expect numpat;
  output out=expected sum=&charge patients;
run;

data expect2;
  set expected;
  if _type_ = 1;
  %initvar(adjmean,(1/patients)*charge);

```

```

%comprat(a_ratio,site,plan);
%comprat(e_ratio,site,adjmean);
format &a_ratio &adjmean &e_ratio 9.2 &charge 9.;
drop &acgsite _type_ _freq_;
run;

/***** Call to Report Generator Macro
*****/
%report;

```

APPENDIX 2 INDIRECT ADJUSTMENT TECHNICAL DESCRIPTION

Associated with each ACG is a value that represents its resource consumption in comparison to other ACGs. This value, known as the ACG relative weight, is calculated as the mean resource use associated with all cases in the ACG divided by the mean resource use for all cases. It is the ACG relative weights that form the building blocks for calculation of the case-mix index.

Calculating the relative weight for a given ACG is a straightforward mathematical process. The average resource consumption for all patients in the stratum is divided by the average resource consumption for all patients in all strata. Mathematically:

$$W_i = \frac{\frac{1}{n} * \sum_{j=1}^n M_{ij}}{\frac{1}{N} * \sum_{i=1}^{51} \sum_{j=1}^n M_{ij}}$$

where:

- W_i = Relative Weight (Mean resource consumption for the i th ACG)
- M_{ij} = Resource consumption measure for Patient j in ACG i
- n = Number of patients in ACG i
- N = Total number of patients in all ACGs

The case-mix index (CMI) represents the gatekeeper's mix of patient types, where patient type is classified by ACG. Case-mix index values are distributed around 1.0, which represents the ACG with the "average" mix of patient type in the population. Therefore, the case-mix index allows for the direct comparison of gatekeepers or delivery sites on the basis of relative mix of patients. For example, a gatekeeper with a case-mix index of .95 would be characterized as treating a less resource-intensive mix of patients than a gatekeeper with a case-mix index of 1.05.

As with relative weights, the computation of CMI is a straightforward process. The number of cases in each ACG is computed for each gatekeeper. This value,

when summed across all ACGs, becomes the denominator in the CMI formula. When multiplied by the corresponding mean resource consumption measure and summed across all ACGs, it becomes the numerator.

Therefore, the formula for CMI can be displayed as follows:

$$CMI_j = \frac{\sum_{i=1}^{51} W_i * D_{ij}}{\sum_{i=1}^{51} D_{ij}}$$

where:

- CMI_j = Case-Mix index for gatekeeper j
- W_i = ACG Relative Weight for ACG i
- D_{ij} = Number of ACG i Cases for Gatekeeper j