ADG	Description	Total	Group 1	Group 2
	continued			
17	Chronic Specialty: UnStable-ENT	0.0%	0.0%	0.1%
18	Chronic Specialty: UnStable-Eye	1.6%	0.8%	5.2%
19	No Longer in Use	0.0%	0.0%	0.0%
20	Dermatologic	4.5%	4.4%	5.0%
21	Injuries/Adverse Effects: Minor	10.8%	10.2%	13.7%
22	Injuries/Adverse Effects: Major	9.3%	8.1%	14.3%
23	Psychosocial: Time Limited, Minor	3.5%	3.0%	5.5%
24	Psychosoc:Recur or Persist: Stable	9.8%	7.4%	20.3%
25	Psychosoc:Recur or Persist: UnStable	5.8%	2.5%	20.1%
26	Signs/Symptoms: Minor	16.9%	15.3%	24.4%
27	Signs/Symptoms: Uncertain	17.5%	14.1%	32.3%
28	Signs/Symptoms: Major	14.8%	11.6%	28.9%
29	Discretionary	5.8%	4.8%	10.4%
30	See and Reassure	1.8%	1.3%	3.8%
31	Prevention/Administrative	43.5%	46.7%	29.5%
32	Malignancy	1.0%	0.3%	4.0%
33	Pregnancy	2.2%	2.6%	0.3%
34	Dental	1.4%	1.4%	1.7%

Table 1 illustrates how ADGs, the building blocks of the ACG system, can quickly demonstrate differences in types of morbidity categories across sub-groupings within your organization. In this example, the case-mix profile of Group 2 tends to be more complex than that of Group 1, with the prevalence of the chronic medical and psychosocial ADGs being especially high.

An advantage of ADGs is they can quickly identify clinically meaningful morbidity trends that may be obscured at the disease-specific or relative morbidity index levels.

As discussed in the "Overview of New Features in ACG Release 6.0" section of this document, the ACG software will automatically assign a six-level (Low to High) simplified morbidity category we term *RUBs* (for Resource Utilization Bands). The six RUBs are formed by combining the ACG mutually exclusive cells that measure overall morbidity burden.

Utilizing the Release 6.0 RUB categories, Table 2 demonstrates how a simple RUB-based analysis highlights differences in the distribution of morbidity of the Group 1 and Group 2 exemplary sub-populations. Confirming the impression drawn from Table 1, the Group 2 population clusters in the bands associated with higher overall morbidity burdens.

Table 2: Percentage Distribution of Two Sub-groups, by RUB Categories

RUB Category	*Total*	**Group 1**	**Group 2**
1 - Non-users	25.8	35.6	22.5
2 - Healthy Users	13.9	17.5	11.1
3 - Low Morbidity	28.3	30.1	25.0
4 - Moderate	27.6	13.8	33.5
5 - High	3.7	2.5	7.4
6 - Very High	.7	.5	1.5

As discussed earlier in this document, through use of disease-specific EDCs a "standardized morbidity ratio" report is now included as part of the standard ACG print file.[3] Based on the "Major" subheadings of Expanded Diagnosis Clusters, this report presents MEDC level disease prevalence of a sub-population of interest after taking into account the age and gender mix of the group relative to the underlying population. Thus, this report will assist users in isolating statistically significant (demographically adjusted) disease category differences within a sub-population of interest.

The diagnostic/morbidity distribution reports outlined here should be useful for many clinically oriented applications within your organization. These could include population clinical needs assessments and targeting where disease management or outreach programs might be developed.

[3] See "Refinements to the Johns Hopkins Expanded Diagnosis Clusters (EDCs)" in the *Version 6.0 Release Notes* and Chapter 13, "Dino-Clusters: The Johns Hopkins Expanded Diagnosis Clusters (EDCs)" in the *Version 5.0 Documentation and Application Manual* for additional details on interpreting this table and how to generate this table on other desired population groups.

Table 3. Observed to Expected Standardized Morbidity Ratio (SMR) by MEDC

Population: AW Number of persons=167109

Major EDC	Observed Prevalence per 1000 Population	Age/Gender Expected per 1000 Population	SMR	Approximate 95 percent confidence interval (low)	(high)
Administrative	269.87	280.93	0.961	0.952	0.969
Allergy	75.56	63.50	1.190	1.169	1.211
Cardiovascular	86.29	79.18	1.090	1.072	1.108
Dental	6.65	7.60	0.876	0.824	0.927
Ears, Nose, Throat	172.29	211.01	0.817	0.807	0.826
Endocrine	40.65	31.44	1.293	1.262	1.324
Eye	54.53	121.67	0.448	0.439	0.457
Female Reproductive	88.28	81.09	1.089	1.071	1.106
Gastrointestinal/Hepatic	67.47	57.13	1.181	1.159	1.203
General Signs and Symptoms	80.15	70.37	1.139	1.120	1.158
General Surgery	108.65	100.40	1.082	1.066	1.098
Genetic	0.25	0.24	1.045	0.729	1.360
Genito-urinary	50.53	48.01	1.053	1.030	1.075
Hematologic	11.49	10.53	1.091	1.042	1.139
Infections	*28.20*	*36.80*	*0.766*	*0.744*	*0.788*
Malignancies	14.01	11.10	1.263	1.212	1.314
Musculoskeletal	164.24	184.12	0.892	0.881	0.903
Neurologic	66.96	58.69	1.141	1.120	1.162
Nutrition	10.04	10.86	0.924	0.880	0.969
Psychosocial	51.25	40.68	1.260	1.233	1.287
Reconstructive	24.36	27.22	0.895	0.867	0.922
Renal	8.87	5.27	1.684	1.598	1.770
Respiratory	126.73	140.04	0.905	0.893	0.917
Rheumatologic	14.72	12.44	1.183	1.136	1.230
Skin	144.07	149.81	0.962	0.950	0.974
Toxic Effects	4.49	5.51	0.815	0.756	0.873
Unassigned	128.37	99.34	1.292	1.275	1.309

c. Profiling Resource Use

One of the most popular uses of the ACG software is to set risk-adjusted resource consumption norms for sub-groups of patients/members within an organization. These norms are compared to actual resource use in order to "profile" provider efficiency and to help suggest where over-use and under-use may be a problem.

Profiling applications are very amenable to simple actuarial cell strategies for risk adjustment. Most ACG users apply the ACG mutually exclusive cells for this purpose while others have chosen to combine ACGs and use RUBs for these applications. The simpler RUB method is sometimes selected when the population's numbers are small or when the need to communicate the inner-workings of the methods to a wide audience of providers is critical.

If a user has historical claims data (or other similar data sources), it is generally preferable to calculate "expected" resource use values for each ACG (or RUB) for each resource measure of interest (e.g., total cost, hospital use, specialist referrals, pharmacy) based on actual

patterns of practice within your organization. If such data are unavailable or inadequate, then the relative weights supplied as part of this release can be used as a proxy.[4]

Table 4 presents a summary of the most common profiling statistics:[5] 1) the actual to group average resource use (unadjusted efficiency ratio); 2) the expected to plan average (the case-mix index or morbidity factor); and 3) the actual to expected average resource use (efficiency ratios). The first is a measure of how the profiling group compares to the "average" population. The second, the morbidity factor provides an indication of how "sick" the profiling population is compared to the "average" population. The last statistic, the observed to expected ratio ("O/E Ratio") provides an indication of how many health care resources were consumed by this group compared to how many resources they would have consumed had they utilized the "average" resource use of the population based on their case-mix characteristics. All three of these statistics are expressed as relative values with the "average" or normative value centered at 1.0. Scores greater than 1.0 indicate higher than average whereas those less than 1.0 indicate lower than average. Tests of statistical significance can be developed to assess outlier status. Clearly the use of risk adjustment provides a dramatically different basis for assessing the performance of the three profiled sites.

Profiles such as those summarized above are a useful tool for evaluating performance and allocating resources within a wide range of ACG users. The most common applications include:

- financial exchange between MCO and providers,
- assessing provider efficiency,
- resource planning,
- evaluating access to care, and
- detecting fraud, waste, and abuse.

In the absence of local resource data that can be used to determine local weights, the concurrent weights available within the ACG System can be used to develop summary measures of case-mix for comparisons between groups.

---

[4] See Chapter 8, "Calibrating ACGs for Intended Use: ACG Weights, Resource Utilization Bands, and Carve-outs," of the *Version 5.0 Documentation and Application Manual* and the "Using the Available Relative Value Weights" section of this document for a detailed discussion of relevant methodologic issues related to weight calculation.

[5] For further details see Chapter 12, "ACG Risk Adjustment and Provider Profiling," in the *Version 5.0 Documentation and Application Manual.*

Table 4. Comparison of Observed to Expected Visits and Calculation of Three Profiling Ratios

		SITE A	SITE B	SITE C
1)	Actual Visits per Person **(Observed)**	5.35	6.10	6.90
2)	Plan Average	5.50	5.50	5.50
3)	Actual to Group Average* **(Unadjusted Efficiency Ratio)**	0.97	1.11	1.26
4)	Number of Expected Visits**	4.30	6.25	5.54
5)	Expected to Plan Average*** **(Morbidity Factor)**	.78	1.14	1.01
6)	Observed to Expected Ratio**** **(Adjusted Efficiency Ratio)**	1.24	0.98	1.25

* Row 1 divided by Row 2
** Expected based on ACG characteristics at each site
*** Row 4 divided by Row 2
**** Row 1 divided by Row 4

d. Disease Management and Case Management Applications

As discussed previously, concurrent ACG / RUB morbidity information can be combined with EDCs to control for morbidity differences across a given disease-specific group of interest (e.g., diabetics enrolled in a disease management program). EDCs will be useful in portraying the disease characteristics of a population of interest. Within disease management programs, if significant differences in expected resource consumption exist across the morbidity sub-classes, this analytic approach should be quite useful in better targeting interventions towards sub-groups at higher risk.

Along these lines, two new tables (see Tables 5 and 6 for examples) are now part of the print file that the ACG software produces. Each row of these tables represents persons falling into EDC (or MEDC) disease-specific categories; the columns array these individuals into RUB co-morbidity categories according to their ACG assignment. Table 5 presents the percentage distribution for a series of selected EDCs across the five RUB categories. Table 6 presents the expected relative resource use within each RUB. This table illustrates co-morbidity's profound influence on resource use within individual disease groups. The ACG-based RUBs do a very good job of explaining variations in resource use within specific diseases. The ACG software

automatically generates these reports based on nationally representative weights, but such tables are likely to become even more useful when calibrated to local cost and practice patterns.[6]

Table 5: Percentage Distribution of Each Co-morbidity Level within an EDC (Samples)

EDC	RUB-1 Very Low	RUB-2 Low	RUB-3 Average	RUB-4 High	RUB-5 Very High
ADM01:General medical exam	19.8	32.9	39.9	6.2	1.3
ADM02:Surgical aftercare	4.7	19.3	46.6	18.9	10.4
ADM03:Transplant status	3.8	7.7	32.9	26.6	29.1
ADM04:Complications of mechanical	0.0	10.3	32.4	25.5	31.8
ALL01:Allergic reactions	0.0	36.2	53.6	8.5	1.6
ALL03:Allergic rhinitis	0.0	34.5	56.0	8.2	1.3
ALL04:Asthma, w/o status asthmati	0.0	23.6	63.2	10.7	2.5
ALL05:Asthma, with status asthmat	0.0	20.9	58.0	15.6	5.4
ALL06:Disorders of the immune sys	0.0	6.5	47.6	25.5	20.4
CAR01:Cardiovascular signs and sy	0.0	14.5	64.2	15.2	6.1
CAR03:Ischemic heart disease (exc	0.0	0.5	55.7	27.3	16.6
CAR04:Congenital heart disease	0.0	17.9	45.9	23.9	12.4
CAR05:Congestive heart failure	0.0	0.4	36.6	31.1	31.9
CAR06:Cardiac valve disorders	0.0	7.6	59.1	22.2	11.1
CAR07:Cardiomyopathy	0.0	2.2	43.8	30.1	23.9
CAR08:Heart murmur	12.3	25.8	44.5	11.9	5.4
CAR09:Cardiac arrhythmia	0.0	3.7	58.4	24.5	13.3
CAR10:Generalized atherosclerosis	0.0	7.0	43.7	25.4	23.9
CAR11:Disorders of lipoid metabol	0.0	17.3	68.0	10.4	4.2
CAR12:Acute myocardial infarction	0.0	0.2	21.3	39.3	39.2
CAR13:Cardiac arrest, shock	0.0	5.4	19.2	31.2	44.2
CAR14:Hypertension, w/o major com	0.0	20.6	64.7	10.2	4.5
CAR15:Hypertension, with major co	0.0	4.1	55.4	24.1	16.3

As discussed elsewhere, EDCs are very useful for many purposes. If users so choose, they can develop their own reports, and the EDCs that define the rows in Tables 5 and 6 could be replaced by episodes of illness categories that an organization may obtain from other sources. ACG-based RUBs are equally effective in explaining variations in resource use within episodes of care.

[6] See Chapter 13, "Dino-Clusters: The Johns Hopkins Expanded Diagnosis Clusters (EDCs)," in the *Version 5.0 Documentation and Application Guide* for detailed instructions on how to create these tables calibrated to your own data.

Table 6: Estimated Concurrent Resource Use by RUB by MEDC (Samples)

EDC	RUB-1 Very Low	RUB-2 Low	RUB-3 Average	RUB-4 High	RUB-5 Very High
ADM01:General medical exam	0.19	0.54	1.97	7.12	24.76
ADM02:Surgical aftercare	0.20	0.63	2.31	7.94	27.30
ADM03:Transplant status	0.20	0.65	2.39	8.23	29.89
ADM04:Complications of mechanical	0.00	0.69	2.35	7.97	29.84
ALL01:Allergic reactions	0.00	0.54	2.07	7.49	25.41
ALL03:Allergic rhinitis	0.00	0.54	2.13	7.43	25.40
ALL04:Asthma, w/o status asthmati	0.00	0.62	2.03	7.43	26.10
ALL05:Asthma, with status asthmat	0.00	0.62	2.13	7.50	28.23
ALL06:Disorders of the immune sys	0.00	0.74	2.39	7.71	29.63
CAR01:Cardiovascular signs and sy	0.00	0.60	2.43	7.96	26.56
CAR03:Ischemic heart disease (exc	0.00	0.68	2.25	8.12	25.35
CAR04:Congenital heart disease	0.00	0.73	2.20	7.11	25.56
CAR05:Congestive heart failure	0.00	0.81	2.62	8.30	28.83
CAR06:Cardiac valve disorders	0.00	0.56	2.42	7.86	27.10
CAR07:Cardiomyopathy	0.00	0.73	2.37	8.23	28.69
CAR08:Heart murmur	0.21	0.64	2.22	7.20	23.05
CAR09:Cardiac arrhythmia	0.17	0.61	2.37	8.07	25.82
CAR10:Generalized atherosclerosis	0.00	0.46	2.47	8.23	27.06
CAR11:Disorders of lipoid metabol	0.00	0.49	2.29	8.17	25.14
CAR12:Acute myocardial infarction	0.00	0.82	1.85	7.87	26.28
CAR13:Cardiac arrest, shock	0.00	0.62	2.12	7.74	27.84
CAR14:Hypertension, w/o major com	0.00	0.48	2.28	8.16	25.75
CAR15:Hypertension, with major co	0.00	0.62	2.35	8.31	27.40

e. High-risk Case Identification for Case Management

The various components of the "acgPM"—the new ACG predictive modeling module—represent a real advance for users wishing to establish or augment care management programs within their organization. Furthermore, existing ACG measures have many applications in this domain as well.

There are many ways to adapt the ACG suite of tools in the pursuit of improved patient care. This sub-section provides a summary and overview of some of the recommended approaches that an organization may wish to consider in the care-management and quality improvement (QI) domains.

As discussed in some detail earlier in this document, the new acgPM risk measurement tools provide information at the individual patient level to help identify persons who potentially would be well served by special attention from the organization's care management infrastructure. This "high-risk case identification" process could be used to target a person for interventions such as a referral to a case-manager, special communication with the patient's physician, structured disease management programs, or educational outreach.

As part of the new acgPM module the Release 6.0 software now includes a report that provides a disease-specific (based on selected individual and aggregated EDCs) distribution of

risk probability scores and average expected resource use for different risk cohorts. This latter report, shown here as Table 7, will be potentially be useful in helping to frame a strategy for targeting various risk cohorts within disease management programs.

Table 7. Number of Cases and acgPM Predicted Relative Resource Use, by Risk Probability Thresholds for Selected Chronic Conditions

	Number of Cases				Predicted Relative Resource Use			
		Probability Score Category			Probability Score Category			
Disease Category (EDC)	Total	≥0.4	≥0.6	≥0.8	<0.4	≥0.4	≥0.6	≥0.8
Arthritis	17,679	940	463	172	2.18	6.82	9.31	15.71
Asthma	27,863	764	386	136	1.43	6.75	9.29	14.85
Diabetes	16,991	1,307	716	345	2.67	7.59	10.62	17.36
Hypertension	50,122	2,064	1,011	457	2.06	7.25	10.27	17.57
Ischemic Heart Disease	9,330	971	514	242	3.27	7.40	10.35	17.33
Congestive Heart Failure	1,634	460	292	184	5.17	8.81	12.26	19.61
Hyperlipidemia	31,240	1,170	529	186	1.97	7.13	9.49	15.46
Low Back Pain	61,980	1,493	723	279	1.76	6.53	8.77	14.27
Depression	10,190	599	298	113	2.09	6.63	9.03	14.30
Chronic Renal Failure	742	308	253	183	13.11	16.48	19.40	25.21
COPD	6,204	545	301	147	2.58	7.71	10.24	16.68

The acgPM's probability score was fine-tuned to identify persons who will likely be the ones in your organization who would most benefit from special attention. To capitalize on this new method, an organization will want to develop periodic reports of members with high acgPM scores who also meet other organizational criteria, such as

- enrollment with certain providers,
- falling into certain eligibility categories,
- residing in certain geographic areas, or
- meeting previous patterns of utilization.

After these other stratifiers are taken into consideration as appropriate, a "case finding" report should list all in-scope individuals arrayed from highest to lowest based on the overall acgPM high-risk probability score within your organization.

In addition to running the report automatically generated by the software, users are encouraged to develop their own "individual risk summary" reports on each potential case over a certain threshold (say the top 1% of individuals). This target group can be winnowed further by

case managers on the basis of various sources of information available from the ACG software and elsewhere. These additional data might include primary care provider information, service history, history of prior inclusion in care management programs, and results from any ongoing surveys (such as health-risk appraisals). Please see the "The ACG Predictive Model: Helping to Manage Care for Persons at Risk for High Future Cost" section of this document for a comprehensive discussion of the new acgPM module and its applications.

f. Capitation, Actuarial Underwriting, and Rate Setting

ACGs have been successful for so long because they do a good job at capturing the complex interplay of co-morbidities that explain the impact of case-mix on resource use. This factor clearly distinguishes ACGs from other risk-adjustment strategies that treat diseases individually, as if they each have a completely independent effect on health. The validity of the ACG view of illness has been borne out through the successful application of this risk-adjustment strategy across a range of applications for over a decade.

i. Using ACGs as "Actuarial Cells"

The ACG System has made it possible to accomplish risk adjustment with fairly simple and straightforward analytic strategies. For example, ACGs can readily be used as actuarial cells, which have long been the primary actuarial method for both capitation rate setting and underwriting. Actuarial cells represent a fixed number of discrete categories into which individuals are placed based on their expected use of resources. ACGs are very well suited for assigning individuals into these types of actuarial cells.

There are a number of advantages associated with using an actuarial cell-based approach to risk adjustment for capitation and underwriting:

- **Simplicity**. Once the population has been classified into around 100 ACG "cells," it is possible to risk-adjust the population by using a spreadsheet. Some users have chosen to simplify this approach even further by collapsing the ACGs into smaller homogeneous groupings, resource utilization bands (RUBs). Even when grouped into RUBs, studies indicate that ACGs retain much of their explanatory power.

- **Less prone to gaming or manipulation.** Particularly in applications involving rate setting, there could be incentives to "game" risk-adjustment strategies to increase payment. Unlike some other disease-specific risk adjusters, aggressive efforts to capture additional diagnostic codes on the part of providers will have a more limited impact on ACG assignments. Where "code creep" associated with general increases in completeness and accuracy of coding exists, the simplicity of the ACG system makes it very easy to identify this trend and to implement appropriate action, such as re-calibration of weights.

- **Stability.** The conceptual elegance and underlying simplicity of ACGs have made the system very stable over long periods. The underlying clinical "truth" captured by ACGs does not change dramatically with each new data set and each new application.

- **Ease of making local calibrations.** It is very easy to recalibrate ACG-based actuarial cells to reflect local differences in patterns of practice, benefit structure, and provider fees. Especially for capitation and rate-setting tasks, we encourage users to calibrate the ACG output to reflect the unique nature of the local cost structure. The same simplicity that makes it possible to risk-adjust using a spreadsheet makes it equally possible to accomplish recalibration using the same types of simple tools.

The ultimate testimony to the value of ACGs used as the basis of actuarial cells is the fact that for almost a decade they have been used to facilitate the exchange of many billions of dollars within numerous private and public health plans in both the United States and Canada.

ii. ACGs in Multivariate Models

Multivariate regression for risk adjustment has been used for many years by some of the more sophisticated users of ACGs. If additional risk descriptors are available beyond diagnosis, age, and sex, this approach has the potential for improved predictive models.

The strength of regression-based strategies is the ease with which additional risk factor information can be incorporated and thereby introduce better control for the effects of case-mix. This ease is also a potential drawback since regression may introduce some assumptions and statistical pitfalls that can be troublesome without seasoned analytical support. Their inherent

complexity makes them difficult to calibrate to local cost patterns, and regression models are also potentially easier to game because more factors can be manipulated. Finally, while it is possible to introduce a wide range of variables that improve the model's explanatory power, this explanatory power is often confined to the data set and time period on which the model is based. The model's results may end up differing significantly from year to year depending on the inter-relations of the myriad risk factors that have been included, a phenomenon referred to as "over-fitting."

To address some of these analytic challenges, the acgPM provided as part of this release represents a regression-based strategy that can be applied for prospective financial applications. As discussed, one acgPM output, the probability score, has been specifically tailored for case and disease-management applications. The other acgPM output, the predicted resource index (PRI), assigns a relative value that can be readily converted to dollars. This PRI output is most relevant for financial risk-adjustment applications and can be considered a substitute for ACG cells for prospective rate setting or payment.

One important caveat is worth noting here. Prior pharmacy cost has been made an optional "risk factor" variable in the new acgPM. Although it is useful for calculating the most accurate predictions for future costs, we do NOT recommend that models using the optional pharmacy cost predictor be applied to capitation rate setting. Instead, we suggest that the acgPM model relying only on ICD input variables be used for such a purpose. We take this position for the same reason we believe that episode groupers that rely on procedure codes (such as CPT) and "Rx-groupers" based on use of specific medications (as defined by NDC codes) should not be used for rate-setting purposes or efficiency profiles. Risk factor variables of this type, which are directly defined by the providers' clinical practices, are potentially intertwined with patterns of "over use" or "under use." Risk-adjusted rates based on these factors may, in a circular manner (termed endogoneity by the economists), lead to setting rates that are inappropriate—either too high or too low. Moreover, when risk factors are determined by such drug use (or procedural) delivery patterns, providers who practice "efficiently" could potentially be penalized for their efficiency. This circularity issue is not a major concern when only diagnostic information (not linked to specific types or settings of service) is used as the main source of information on risk factors.

iii. To Regress or Not to Regress: That Is the Question

One of the key decision points in using risk adjustment for financial applications is whether to use a simple actuarial cell approach or a more complex multivariate model. If you have been applying ACG-based actuarial cells successfully for some time, there may be little incentive to change since ACGs alone remain a highly effective case-mix adjustment tool. "If it ain't broke, don't fix it." If you are just starting out in your selection of methods, you will need to balance the stability and ease of use of ACG-based actuarial cells against the potentially enhanced ability to explain variations in resource use by applying regression modeling strategies. If you have access to additional well-validated risk factor data and if you have previous experience using regression models within your organization, then you should consider using regression. In regression strategies, ACGs, ADGs, and EDCs remain valuable as distinct risk factors to be supplemented by additional data. Although EDCs are useful for identifying individuals with specific high impact diseases, it is important to note that they do not account for burden of co-morbidity as do ACGs. Therefore, we do not generally recommend that EDCs be used as the only means of controlling for case-mix in regression analysis.

The ultimate choice of risk-adjustment approach depends on the specific application, and it is prudent to compare both actuarial cell and regression approaches over a span of several years before making a final decision. For multivariate models, the R-squared statistic is often used as an indicator of performance. In fact, extreme caution is recommended when evaluating models based on R-squared values. The R-squared statistic is very sensitive to outlier individuals. Aside from considering measures of model fit, such as the R-squared value, you should consider whether the results are reasonably stable over time. It is also advisable to simulate the degree to which each approach results in over- or under-payment to key segments of your population.

iv. Concurrent versus Prospective Applications

The time frame used for most rate setting and other financial analyses is a "prospective" or predictive one. That is, this year's diagnostic information is used to determine risk factors and expected resource consumption in some future period. Thus the weights associated with each risk factor are calibrated to that future period. But this is not the only temporal approach that organizations can use for rate setting. Some ACG users have implemented concurrent rating processes for financial exchanges. In such cases, this year's expected resource use among the

benchmark population is attached to each ACG cell as a relative value rather than next year's resource use. While we do encourage experienced actuaries and financial analysts to learn more about the advantages and challenges of these innovative concurrent approaches, we do not recommend that organizations apply concurrent approaches to payment without first simulating the impact that these methods might have on the rate-setting process.

A real-world example of a concurrent approach to rate setting is one being implemented in Minnesota Medicaid where plan-level payments are based on concurrent ACG-adjusted profiles of the plan. Under this scenario, payment to a health plan is the same for each individual enrollee within a particular plan; however, the amount paid is case-mix adjusted by the plan's *overall* morbidity burden (relative to an average, across the population, of 1.0). This approach assumes that the morbidity burden of large groups (i.e., any individual health plan) is fairly stable and that the group's overall morbidity does not change much by the addition/exit of any one individual.

For additional discussion on this and related issues related to risk adjustment as applied to financial exchanges, we encourage readers to review our chapter incorporated into Charles Wrightson's recently published book *Financial Strategy for Managed Care Organizations: Rate Setting, Risk Adjustment, and Competitive Advantage* (see http://www.ache.org/pubs/wrightson.cfm for ordering details). Our chapter is available online at http://www.acg.jhsph.edu. Readers are also encouraged to review the ACG bibliography at that site for a variety of articles illustrating ACGs used for capitation.

g. In Closing

As part of our ongoing commitment to furthering the international state-of-the-art of risk-adjustment methodology and supporting ACG users worldwide, we will continue to perform evaluation, research, and development. We will look forward to sharing the results of this work with our user-base via white papers, our web site, peer-reviewed articles, and in-person presentations. After you have carefully reviewed the documentation supplied with this software release, we would welcome your inquiries on any topic of relevance to your use of ACGs within your organization. (Contact us at askacg@jhsph.edu.) We thank you for using ACGs and for

helping us to work toward meeting the Johns Hopkins University's ultimate goal of improving the quality, efficiency, and equity of health care across the United States and around the globe.

Section 6
Installation and Usage

Intended for old and new users alike, this chapter is written for the programmer/analyst who will be using the software. This chapter begins with an overview of the technical enhancements new to Release 6.0 and is required reading for all users of the software. The remainder of the chapter is divided into the following main sections:

- installing the software,
- using the software,
- required components of the input files and how to pass data to the software, and
- output files.

Readers are referred to the relevant sections of the *ACG Version 6.0 Release Notes* and the *Version 5.0 Documentation and Application Manual* for additional details on calculating or using ACG weights, discussions of the built-in reporting features and how to customize these reports, and other relevant criteria for implementing ACG technologies within your organization.

a. For Users Already Familiar with the ACG Software

Although the functionality of Release 6.0 has been greatly enhanced, these improvements only slightly affect the installation and use of the software. This section highlights the few important changes so that current ACG System users will be able to get up and running quickly. A full set of current installation and usage instructions follows so that new (and current) users will have all the information that they need readily at hand.

What is needed to take advantage of ACG System enhancements involves two new "control card keywords" and four "names" that have been added to the OUTREC control card. The control card keywords allow for additional (and optional) data to be passed to the grouper from the input file; the four names accommodate additional (and optional) output fields produced by the software.

i. New Control Card Keywords

Two new record layout control card keywords permit the input of two new data fields:

1. **POP** is used for a group membership identifier such as underwriting group or PCP assignment; and
2. **PCOST** is used for pharmacy cost data.

The group identifier, called POP, serves as a stratifier for producing age/sex-adjusted prevalence rates and standard morbidity ratio reports using the Johns Hopkins Expanded Diagnosis Clusters (EDCs) typology.[7]

The optional pharmacy cost information, called PCOST, is a useful adjunct that improves performance of the acgPM (see description of the HRCI card below). If pharmacy cost information is available, we recommend its inclusion.

ii. New OUTREC Features

Four new names have been added to the control card that controls the output file. **OUTREC** additions are as follows:

1. **CWT**, to output a set of fixed concurrent ACG weights based on our nationally representative database (written to the output file as ##.###).
2. **RUB**, to output six resource use levels (Resource Utilization Bands, or RUBs for short) expressed as an ordinal number with values between zero and five as follows:

   0 = nonusers
   1 = healthy users
   2 = low morbidity
   3 = moderate morbidity
   4 = high morbidity
   5 = very high morbidity

3. **HOS**, to output a Boolean indicator (e.g., a value of zero or one) for the presence of conditions likely to lead to a hospitalization.
4. **HRCI**, to output the four scores for predictive modeling (written to the output file as ###.###) in the following order:

---

[7] Note: Please refer to Section 2 of this Addenda *Additions and Refinements to the Expanded Diagnosis Cluster (EDC) Methodology (including new reports) and ICD-9 Coding Updates* and Chapter 13, " 'Dino-Clusters': The Johns Hopkins Expanded Diagnosis Clusters (EDCs)," of the *Version 5.0 Documentation and Application Manual* for technical specifications necessary to customize these tables to your application.

a. *Total cost predicted resource index* – an estimate for Year 2 total expenditures (including pharmacy charges) expressed as a relative weight;
b. *Pharmacy cost predicted resource index* – an estimate for Year 2 predicted pharmacy expenditures also expressed as a relative weight;
c. *Probability of being in the high total cost cohort* – a probability score with values between zero and one, indicating the likelihood that a person will have high cost in the subsequent time period; and
d. *Probability of being in the high pharmacy cost cohort* – a probability score indicating the likelihood that a person will have high pharmacy cost in the subsequent time period.

Also note that users may include the POP and PCOST control cards in the OUTREC statement if they are interested in having this information appear in the output (as well as the input) file produced by the software.

iii. Processing Speed and Space Requirements

Unlike prior versions of the software that allocated memory dynamically for each individual, the expanded reporting features and modeling components of Release 6.0 require more processing time. Processing time depends on a variety of factors including but not limited to CPU speed, disk read-write speed, available memory and disk space, as well as the size of the input files. As a general guideline the software can be expected to take two to three times longer than previous versions. For groups that are fewer than 100,000 the increase in processing speed is negligible, but for very large populations the software may take substantially longer than prior versions to process similarly sized input files. Note that processing time may be reduced if duplicate diagnoses cards are removed from the input data stream.

The expanded reports and model-building component of Release 6.0 now also necessitate the use of temporary files. While the ultimate output file(s) produced by the software will be similar in size to those produced by prior versions of the software (with the addition of space needed to accommodate new OUTREC fields), sufficient disk space must be available for the writing of temporary files. At a minimum the software will need 1 MB of available disk space for general overhead. On top of this, an estimated additional 600 bytes per person for member identifier, age, and sex fields plus an additional 51 bytes per each unique ICD per person are

required. To simplify things somewhat, as a general guideline Release 6.0 requires temporary disk space approximately five to six times the size of the input file.

Because of the modeling component of the software, there is now for the first time a limit to the maximum number of people that can be processed in one run. We estimate the upper limit to be approximately 3.5 million individuals. The limit is imposed not on the maximum number of individuals, but rather on the maximum allowable temporary file size, which has been set at 2 GB in most operating systems. If you have membership in excess of 3.5 million, please contact your software distributor for further guidance on how best to divide your input file into smaller population subgroupings.

iv. The Print File

The number of fixed reports has increased in this release, and so the size of the software's print file can be expected to be larger. The size of the print file will be at least partially determined by the number of levels provided in the population stratifier control card (e.g., POP discussed in the preceding paragraphs) because these reports are generated for each individual stratifier. If hundreds of stratifiers are used, the software will generate hundreds of pages of printed reports. Even if hundreds of stratifiers are included, the print file can be easily managed by loading it into text processing software and globally changing the print font to Courier 8. With this simple adjustment, reports should be legible and page breaks should appear in logical places. Extracts of this file can easily be extracted by using the cut and paste feature to separate files as desired.

v. Other

The input file or files no longer need to be sorted. Consequently, the NOSORT card is not needed (and is no longer provided). The age reference date must be provided on the DOB card and must be of the form CCYYMMDD. Previously this information could be left blank, and the system date would be used, or it could be provided as YYMMDD. The format of the EDC output file has changed, so current users must change their programs to read data from these files. Although the EDCs are still written from columns one through five, the unique member ID does not begin until column 12 so as to allow for future expanded EDC categories.

b. Installing the Software

The ACG grouper software is supplied on a diskette or CD. Installation involves copying all files from this disk to your hard disk; other installation steps depend on the platform. If you are *upgrading from an earlier release* and want to maintain the prior release for comparison purposes, then *rename the old executable file before installing the new version.* This can be done by using the DOS **rename** command or the UNIX **mv** command. Consult your system documentation for more information on renaming files. Alternatively, you could create separate directories for each version of the software, although you should be sure to use the **appropriate version**. The version of the software used at any one time is listed in the heading of the print file output.

i. UNIX Platforms

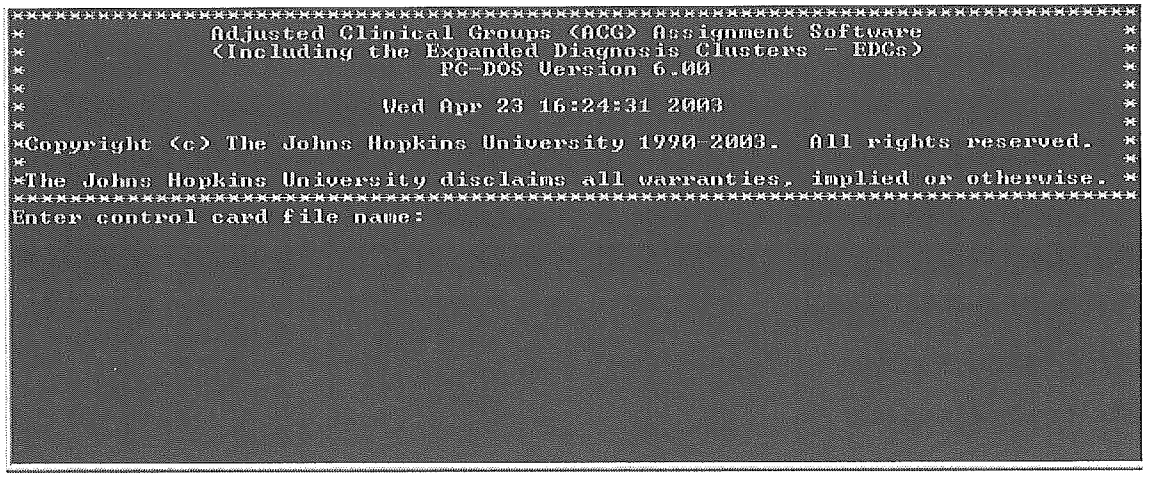
The file supplied on the disk is named ACGGROUP. Copy this file to your UNIX partition. You may then need to change the file mode to allow "execute permission." This is usually done with the chmod command, e.g., **chmod +x acggroup**. The software can then be invoked by merely entering **acggroup controlcardfilename**. By pressing the break key combination for your system (e.g., Ctrl-Break or Ctrl-C), the system can be halted. That completes the UNIX installation.

ii. PC Platforms

The file supplied on the disk is named ACGGROUP.EXE. Copy this file to your hard disk for the PC-DOS installation.

c. Using the Software

Figure 1. Opening Screen for ACG Version 6.0.



After copying the file to the appropriate subdirectory, to access the software (acggroup.exe) simply double click[8] on the file listing or icon using Windows Explorer or File Manager. Figure 1 is a screen capture of what should appear. The software prompts for the controlcardfilename. The controlcardfilename specifies the location of a file containing a series of ACG control cards that communicate to the software a) the location of the input and output files and what fields to read from the input file, b) which ACG branching options are to be used, and c) what ACG-based risk assessment variables are to be written to the output file(s). (See Table 1. More detail to follow.) After you type in the full filename and hit return, the software executes and a series of progress bars will appear at the bottom of the screen indicating the percentage of data processed. When 100% of the data is processed, the window automatically closes, and files created by the software will reside in the appropriate directories (as indicated by the control card). If a problem is found with any of the control cards, then an error message is written to the screen and the user must press *enter* to halt execution of the program (at which time viewing the print file may help users to better ascertain where the problem resides and/or at what point the software stopped executing).

[8] Alternatively, type the following from the command line: ACGGROUP CONTROLCARDFILENAME (where CONTROLCARDFILENAME is the filename that contains all of the control cards)