The Prevalence and Cost of Select Chronic Diseases

Prepared for:
The Pharmaceutical Research and Manufacturers of America (PhRMA)

Prepared by:
The Lewin Group
Brian Simonson, M.S.
Soumita Lahiri, M.S.

March 2007

Table of Contents

Introduction

Chronic disease conditions place an enormous strain on the lives of people. The prevalence rate of many conditions has increased at an alarming rate in recent years. These diseases not only affect the quality of life for individuals, but also place a large economic burden on the community. Consequently, there exists great interest in ascertaining the prevalence and economic burden of chronic diseases within various communities. This report discusses the prevalence of diabetes, cardiovascular disease, depression, asthma, and cancer. In addition, estimates of the increased costs associated with these diseases are discussed.

The following represents standardized reports for South Carolina, Iowa, and New Hampshire. Each report addresses the specific demographics, estimated prevalence, and cost statistics for the state. Estimates were created using a model The Lewin Group developed. The methodology for the model is presented following the state reports.

Reports

South Carolina

Chronic diseases place major strains on the community in terms healthcare costs and loss of productivity. Measuring the financial impact of major chronic diseases on individual communities can help quantify this strain on the community. The purpose of this memo is to provide the summary results for a study that analyzed the prevalence of cancer, asthma, diabetes, cardiovascular disease, and depression for various geographic regions. In addition, this memo also identifies the increased costs associated with these diseases. In particular, this memo focuses on results for South Carolina.

Statistical models were formulated using several well-known data sources[1]. Results specific to South Carolina were achieved by utilizing the demographics specifically for the district.

As of July 2006, it was estimated that South Carolina had a population of 1,225,304 African Americans, 125,238 Hispanics, and 2,740,604 Whites. The estimated demographics can be seen below in Table1.

[1] Analysis was performed by utilizing the National Health Interview Survey (NHIS), the Medical Expenditure Panel Survey (MEPS), the Current Population Survey (CPS), and Census 2000 Data.

Table 1. Demographics:

	African American	Hispanic	White	All Three Groups	National (In Millions)
Total	1,224,120	125,302	2,740,445	4,089,867	293.8
Age					
Less than 20 years old	397,298 (32.4%)	44,042 (35.2%)	646,679 (23.6%)	1,088,020 (26.6%)	81.6 (27.76
20 to 29 years old	183,058 (14.9%)	29,320 (23.4%)	362,854 (13.2%)	575,233 (14.1%)	40.5 (13.79
30 to 39 years old	155,330 (12.7%)	27,176 (21.7%)	361,139 (13.2%)	543,645 (13.3%)	40.1 (13.65
40 to 49 years old	192,074 (15.7%)	13,232 (10.6%)	428,730 (15.6%)	634,036 (15.5%)	44.9 (15.27
50 to 59 years old	138,938 (11.3%)	8,033 (6.4%)	392,352 (14.3%)	539,323 (13.2%)	38.1 (12.97
Greater than 60 years old	158,605 (12.9%)	3,435 (2.7%)	548,850 (20.%)	710,890 (17.4%)	48.7 (16.56
Sex					
Male	568,274 (46.4%)	63,194 (50.4%)	1,327,893 (48.5%)	1,959,361 (47.9%)	144.2 (49.07
Female	655,846 (53.6%)	62,109 (49.6%)	1,412,551 (51.5%)	2,130,506 (52.1%)	149.6 (50.93
Household Income					
Less than $20,000	389,216 (31.8%)	37,140 (29.6%)	494,751 (18.1%)	921,106 (22.5%)	57.0 (19.4
$20,000 to $34,999	297,590 (24.3%)	28,837 (23.%)	428,368 (15.6%)	754,795 (18.5%)	48.3 (16.44
$35,000 - $54,999	226,463 (18.5%)	32,022 (25.6%)	573,400 (20.9%)	831,885 (20.3%)	53.9 (18.34
$55,000 - $74,999	154,553 (12.6%)	16,645 (13.3%)	417,299 (15.2%)	588,497 (14.4%)	41.2 (14.03
$75,000 and over	156,299 (12.8%)	10,658 (8.5%)	826,628 (30.2%)	993,585 (24.3%)	93.4 (31.78

Table 2 below shows age adjusted diagnosed prevalence rates for the United States. Hispanics are 61% more likely to have diabetes compared to Whites, adjusting for age. Additionally, African Americans are 70% more likely to have diagnosed diabetes than Whites, adjusting for age[2]. It should be noted that these are only diagnosed case, and it has been documented that African Americans and Hispanics have a larger proportion of undiagnosed diabetes cases. Thus, the true difference in diabetes prevalence is truly greater.

Table 2. National Age-Adjusted Prevalence Rates

	Black/African American	Hispanics	Whites*
Heart Disease	26.3%	17.0%	22.3%
Asthma	10.9%	6.9%	9.5%
Cancer	2.0%	1.5%	4.8%
Diabetes	8.2%	7.7%	4.8%
Depression	4.9%	5.2%	8.7%
Any of the five	38.3%	26.8%	36.6%

*All groups excluding African American or Hispanic. Primarily White.

[2] Increased likelihoods calculated using ratio of hazards, or a ratio of prevalence rates.

Table 3 below shows the prevalence for cancer, asthma, diabetes, cardiovascular disease, and depression in the state[3]. For African Americans in South Carolina, Heart Disease had the highest estimated prevalence (302,574 people). Heart Disease had the highest estimated prevalence total for Hispanics (12,517 people). Heart Disease had the highest estimated prevalence total for Whites (678,249 people)[4].

Table 3. Analytical Results: Prevalence Totals by Disease

	Black/African American	Hispanic	White	Overall	Prevalence or Uninsured Rate	National (In Millions)
Disease						
Heart Disease	302,574	12,517	678,249	993,340	24.3%	64.5
Cancer	21,548	1,202	139,352	162,102	4.0%	12.2
Asthma	137,832	8,872	264,049	410,753	10.0%	2.7
Diabetes	90,695	5,615	150,381	246,691	6.0%	23.6
Depression	58,633	5,842	247,991	312,466	7.6%	22.7
Any of the five	466,051	25,615	1,097,431	1,589,098	38.9%	104.0
Uninsured	236,640	50,636	396,849	684,125	16.7%	49.2

In addition, Table 3 shows the projected uninsured totals for the state. These estimates are point in time estimates, meaning the presented statistics are for at any given point in time. In the state, African Americans have an estimated 236,640 uninsured. Hispanics have an estimated 50,636 uninsured. Finally, Whites have a projected 396,849 uninsured.

Finally, Table 4 below displays the estimated incurred costs due to the five chromic diseases. These averages are national averages, taken from recent cost of illness studies, and are not specific to the state. The estimates include both direct and indirect costs, meaning they include both direct healthcare cost s as well as indirect costs such as lost earnings. Additionally, these costs have, to the extent possible, factored out costs due to other related illnesses. In other words, the cost estimates attempt to estimates the costs due to that specific disease, excluding the effects of other co-morbidities.

Table 4. Cost of illness

[3] Prevalence refers to diagnosed prevalence. For diabetes there are studies that suggest approximately one third of all diabetes cases are undiagnosed, but information on all the chronic diseases discussed in this paper do not have analogous information available. For consistency, only diagnosed prevalence is reported.

[4] Note that these estimates were formulated from survey data, which tends to underreport prevalence due to the fact respondents were unaware they had a particular disease. Through literature review, and adjustment factor was applied to diabetes results to correct for the underreporting issue. The remaining diseases remain unchanged and should be view as reported prevalence, which may serve as a lower bound estimate for true prevalence.

	Per Capita Cost of Illness (National Average)	Increased Costs for District (In Millions)			
	Difference	African American	Hispanic	White	Overall
Disease					
Heart Disease	$5,438	$1,645.5	$68.1	$3,688.5	$5,402.1
Cancer	$19,281	$415.5	$23.2	$2,686.8	$3,125.5
Asthma	$2,584	$356.2	$22.9	$682.3	$1,061.4
Diabetes	$10,909	$659.6	$40.8	$1,093.7	$1,794.1
Depression	$3,971	$232.8	$23.2	$984.8	$1,240.9
Any of the Five	$3,917	$3,007.7	$165.2	$7,824.8	$10,997.7

Study Methodology

We utilized Census data to attain demographic information by region. We used the Medical Expenditure Panel Survey (MEPS) to collect information concerning disease prevalence and insurance coverage. To model prevalence and insured status, a predictive model was created that adjusted for various demographic factors, including, age, sex, and income. The characteristics of each region of interest were placed into the model to made region specific projections.

Iowa

Chronic diseases place major strains on the community in terms healthcare costs and loss of productivity. Measuring the financial impact of major chronic diseases on individual communities can help quantify this strain on the community. The purpose of this memo is to provide the summary results for a study that analyzed the prevalence of cancer, asthma, diabetes, cardiovascular disease, and depression for various geographic regions. In addition, this memo also identifies the increased costs associated with these diseases. In particular, this memo focuses on results for Iowa.

Statistical models were formulated using several well-known data sources[5]. Results specific to Iowa were achieved by utilizing the demographics specifically for the district.

As of July 2006, it was estimated that Iowa had a population of 61,913 African Americans, 141,251 Hispanics, and 2,604,416 Whites. The estimated demographics can be seen below in Table1.

Table 1. Demographics:

	African American	Hispanic	White	All Three Groups	National (In Millions)
Total	61,887	141,199	2,604,384	2,807,471	293.8
Age					
Less than 20 years old	25,873 (41.8%)	63,046 (44.6%)	630,734 (24.2%)	719,654 (25.6%)	81.6 (27.76
20 to 29 years old	10,711 (17.3%)	33,186 (23.5%)	354,490 (13.6%)	398,386 (14.2%)	40.5 (13.79
30 to 39 years old	9,178 (14.8%)	19,012 (13.5%)	307,721 (11.8%)	335,911 (12.%)	40.1 (13.65
40 to 49 years old	8,483 (13.7%)	15,049 (10.7%)	406,909 (15.6%)	430,441 (15.3%)	44.9 (15.27
50 to 59 years old	3,755 (6.1%)	8,468 (6.%)	396,120 (15.2%)	408,343 (14.5%)	38.1 (12.97
Greater than 60 years old	3,911 (6.3%)	2,489 (1.8%)	508,443 (19.5%)	514,843 (18.3%)	48.7 (16.56
Sex					
Male	34,542 (55.8%)	73,123 (51.8%)	1,273,048 (48.9%)	1,380,714 (49.2%)	144.2 (49.07
Female	27,346 (44.2%)	68,076 (48.2%)	1,331,336 (51.1%)	1,426,758 (50.8%)	149.6 (50.93
Household Income					
Less than $20,000	38,052 (61.5%)	44,897 (31.8%)	408,925 (15.7%)	491,874 (17.5%)	57.0 (19.4
$20,000 to $34,999	8,252 (13.3%)	36,822 (26.1%)	378,926 (14.5%)	424,000 (15.1%)	48.3 (16.44
$35,000 - $54,999	6,288 (10.2%)	26,796 (19.%)	538,104 (20.7%)	571,188 (20.3%)	53.9 (18.34
$55,000 - $74,999	4,221 (6.8%)	16,969 (12.%)	460,873 (17.7%)	482,063 (17.2%)	41.2 (14.03
$75,000 and over	5,075 (8.2%)	15,716 (11.1%)	817,557 (31.4%)	838,348 (29.9%)	93.4 (31.78

[5] Analysis was performed by utilizing the National Health Interview Survey (NHIS), the Medical Expenditure Panel Survey (MEPS), the Current Population Survey (CPS), and Census 2000 Data.

Table 2 below shows age adjusted diagnosed prevalence rates for the United States. Hispanics are 61% more likely to have diabetes compared to Whites, adjusting for age. Additionally, African Americans are 70% more likely to have diagnosed diabetes than Whites, adjusting for age[6]. It should be noted that these are only diagnosed case, and it has been documented that African Americans and Hispanics have a larger proportion of undiagnosed diabetes cases. Thus, the true difference in diabetes prevalence is truly greater.

Table 2. National Age-Adjusted Prevalence Rates

	Black/African American	Hispanics	Whites*
Heart Disease	26.3%	17.0%	22.3%
Asthma	10.9%	6.9%	9.5%
Cancer	2.0%	1.5%	4.8%
Diabetes	8.2%	7.7%	4.8%
Depression	4.9%	5.2%	8.7%
Any of the five	38.3%	26.8%	36.6%

*All groups excluding African American or Hispanic. Primarily White.

Table 3 below shows the prevalence for cancer, asthma, diabetes, cardiovascular disease, and depression in the state[7]. For African Americans in Iowa, Heart Disease had the highest estimated prevalence (8,733 people). Asthma had the highest estimated prevalence total for Hispanics (9,983 people). Heart Disease had the highest estimated prevalence total for Whites (624,552 people)[8].

Table 3. Analytical Results: Prevalence Totals by Disease

	Black/African American	Hispanic	White	Overall	Prevalence or Uninsured Rate	National (In Millions)
Disease						
Heart Disease	8,733	9,472	624,552	642,758	22.9%	64.5
Cancer	761	1,221	132,848	134,829	4.8%	12.2
Asthma	7,202	9,983	250,399	267,583	9.5%	2.7
Diabetes	3,209	4,764	136,955	144,928	5.2%	23.6

[6] Increased likelihoods calculated using ratio of hazards, or a ratio of prevalence rates.

[7] Prevalence refers to diagnosed prevalence. For diabetes there are studies that suggest approximately one third of all diabetes cases are undiagnosed, but information on all the chronic diseases discussed in this paper do not have analogous information available. For consistency, only diagnosed prevalence is reported.

[8] Note that these estimates were formulated from survey data, which tends to underreport prevalence due to the fact respondents were unaware they had a particular disease. Through literature review, and adjustment factor was applied to diabetes results to correct for the underreporting issue. The remaining diseases remain unchanged and should be view as reported prevalence, which may serve as a lower bound estimate for true prevalence.

Depression	2,603	5,618	229,233	237,454	8.5%	22.7
Any of the five	19,485	25,906	1,031,932	1,077,323	38.4%	104.0
Uninsured	18,044	53,881	357,567	429,493	15.3%	49.2

In addition, Table 3 shows the projected uninsured totals for the state. These estimates are point in time estimates, meaning the presented statistics are for at any given point in time. In the state, African Americans have an estimated 18,044 uninsured. Hispanics have an estimated 53,881 uninsured. Finally, Whites have a projected 357,567 uninsured.

Finally, Table 4 below displays the estimated incurred costs due to the five chromic diseases. These averages are national averages, taken from recent cost of illness studies, and are not specific to the state. The estimates include both direct and indirect costs, meaning they include both direct healthcare cost s as well as indirect costs such as lost earnings. Additionally, these costs have, to the extent possible, factored out costs due to other related illnesses. In other words, the cost estimates attempt to estimates the costs due to that specific disease, excluding the effects of other co-morbidities.

Table 4. Cost of illness

	Per Capita Cost of Illness (National Average)	**Increased Costs for District (In Millions)**			
	Difference	**African American**	**Hispanic**	**White**	**Overall**
Disease					
Heart Disease	$5,438	$47.5	$51.5	$3,396.5	$3,495.5
Cancer	$19,281	$14.7	$23.5	$2,561.4	$2,599.6
Asthma	$2,584	$18.6	$25.8	$647.0	$691.4
Diabetes	$10,909	$23.3	$34.6	$996.0	$1,054.0
Depression	$3,971	$10.3	$22.3	$910.4	$943.0
Any of the Five	$3,917	$109.3	$150.7	$7,342.2	$7,602.3

Study Methodology

We utilized Census data to attain demographic information by region. We used the Medical Expenditure Panel Survey (MEPS) to collect information concerning disease prevalence and insurance coverage. To model prevalence and insured status, a predictive model was created that adjusted for various demographic factors, including, age, sex, and income. The characteristics of each region of interest were placed into the model to made region specific projections.

New Hampshire

Chronic diseases place major strains on the community in terms healthcare costs and loss of productivity. Measuring the financial impact of major chronic diseases on individual communities can help quantify this strain on the community. The purpose of this memo is to provide the summary results for a study that analyzed the prevalence of cancer, asthma, diabetes, cardiovascular disease, and depression for various geographic regions. In addition, this memo also identifies the increased costs associated with these diseases. In particular, this memo focuses on results for New Hampshire.

Statistical models were formulated using several well-known data sources[9]. Results specific to New Hampshire were achieved by utilizing the demographics specifically for the district.

As of July 2006, it was estimated that New Hampshire had a population of 11,241 African Americans, 18,652 Hispanics, and 1,228,606 Whites. The estimated demographics can be seen below in Table1.

Table 1. Demographics:

	African American	Hispanic	White	All Three Groups	National (In Millions)
Total	11,237	18,683	1,228,436	1,258,355	293.8
Age					
Less than 20 years old	5,084 (45.2%)	6,929 (37.1%)	307,617 (25.%)	319,630 (25.4%)	81.6 (27.76
20 to 29 years old	1,010 (9.%)	3,014 (16.2%)	145,168 (11.8%)	149,192 (11.9%)	40.5 (13.79
30 to 39 years old	1,895 (16.9%)	3,295 (17.7%)	169,963 (13.8%)	175,153 (13.9%)	40.1 (13.65
40 to 49 years old	1,574 (14.%)	2,978 (16.%)	205,220 (16.7%)	209,772 (16.7%)	44.9 (15.27
50 to 59 years old	1,379 (12.3%)	1,377 (7.4%)	187,920 (15.3%)	190,676 (15.2%)	38.1 (12.97
Greater than 60 years old	299 (2.7%)	1,060 (5.7%)	212,718 (17.3%)	214,077 (17.%)	48.7 (16.56
Sex					
Male	5,554 (49.4%)	8,843 (47.3%)	608,158 (49.5%)	622,555 (49.5%)	144.2 (49.07
Female	5,682 (50.6%)	9,840 (52.7%)	620,278 (50.5%)	635,800 (50.5%)	149.6 (50.93
Household Income					
Less than $20,000	2,373 (21.1%)	2,854 (15.3%)	140,161 (11.4%)	145,389 (11.6%)	57.0 (19.4
$20,000 to $34,999	2,044 (18.2%)	1,174 (6.3%)	180,635 (14.7%)	183,854 (14.6%)	48.3 (16.44
$35,000 - $54,999	3,069 (27.3%)	4,368 (23.4%)	192,546 (15.7%)	199,983 (15.9%)	53.9 (18.34
$55,000 - $74,999	2,345 (20.9%)	3,601 (19.3%)	185,857 (15.1%)	191,803 (15.2%)	41.2 (14.03
$75,000 and over	1,406 (12.5%)	6,685 (35.8%)	529,236 (43.1%)	537,326 (42.7%)	93.4 (31.78

[9] Analysis was performed by utilizing the National Health Interview Survey (NHIS), the Medical Expenditure Panel Survey (MEPS), the Current Population Survey (CPS), and Census 2000 Data.

Table 2 below shows age adjusted diagnosed prevalence rates for the United States. Hispanics are 61% more likely to have diabetes compared to Whites, adjusting for age. Additionally, African Americans are 70% more likely to have diagnosed diabetes than Whites, adjusting for age[10]. It should be noted that these are only diagnosed case, and it has been documented that African Americans and Hispanics have a larger proportion of undiagnosed diabetes cases. Thus, the true difference in diabetes prevalence is truly greater.

Table 2. National Age-Adjusted Prevalence Rates

	Black/African American	Hispanics	Whites*
Heart Disease	26.3%	17.0%	22.3%
Asthma	10.9%	6.9%	9.5%
Cancer	2.0%	1.5%	4.8%
Diabetes	8.2%	7.7%	4.8%
Depression	4.9%	5.2%	8.7%
Any of the five	38.3%	26.8%	36.6%

*All groups excluding African American or Hispanic. Primarily White.

Table 3 below shows the prevalence for cancer, asthma, diabetes, cardiovascular disease, and depression in the state[11]. For African Americans in New Hampshire, Heart Disease had the highest estimated prevalence (1,361 people). Heart Disease had the highest estimated prevalence total for Hispanics(1,923 people). Heart Disease had the highest estimated prevalence total for Whites (271,825 people)[12].

Table 3. Analytical Results: Prevalence Totals by Disease

	Black/African American	Hispanic	White	Overall	Prevalence or Uninsured Rate	National (In Millions)
Disease						
Heart Disease	1,361	1,923	271,825	275,109	21.9%	64.5
Cancer	156	223	61,679	62,058	4.9%	12.2
Asthma	1,245	1,246	115,163	117,654	9.3%	2.7
Diabetes	473	771	57,831	59,075	4.7%	23.6

[10] Increased likelihoods calculated using ratio of hazards, or a ratio of prevalence rates.

[11] Prevalence refers to diagnosed prevalence. For diabetes there are studies that suggest approximately one third of all diabetes cases are undiagnosed, but information on all the chronic diseases discussed in this paper do not have analogous information available. For consistency, only diagnosed prevalence is reported.

[12] Note that these estimates were formulated from survey data, which tends to underreport prevalence due to the fact respondents were unaware they had a particular disease. Through literature review, and adjustment factor was applied to diabetes results to correct for the underreporting issue. The remaining diseases remain unchanged and should be view as reported prevalence, which may serve as a lower bound estimate for true prevalence.

Depression	404	774	102,876	104,054	8.3%	22.7
Any of the five	3,282	3,945	466,430	473,656	37.6%	104.0
Uninsured	1,819	4,426	153,093	159,337	12.7%	49.2

In addition, Table 3 shows the projected uninsured totals for the state. These estimates are point in time estimates, meaning the presented statistics are for at any given point in time. In the state, African Americans have an estimated 1,819 uninsured. Hispanics have an estimated 4,426 uninsured. Finally, Whites have a projected 153,093 uninsured.

Finally, Table 4 below displays the estimated incurred costs due to the five chromic diseases. These averages are national averages, taken from recent cost of illness studies, and are not specific to the state. The estimates include both direct and indirect costs, meaning they include both direct healthcare cost s as well as indirect costs such as lost earnings. Additionally, these costs have, to the extent possible, factored out costs due to other related illnesses. In other words, the cost estimates attempt to estimates the costs due to that specific disease, excluding the effects of other co-morbidities.

Table 4. Cost of illness

	Per Capita Cost of Illness (National Average)	**Increased Costs for District (In Millions)**			
	Difference	**African American**	**Hispanic**	**White**	**Overall**
Disease					
Heart Disease	$5,438	$7.4	$10.5	$1,478.3	$1,496.1
Cancer	$19,281	$3.0	$4.3	$1,189.2	$1,196.5
Asthma	$2,584	$3.2	$3.2	$297.6	$304.0
Diabetes	$10,909	$3.4	$5.6	$420.6	$429.6
Depression	$3,971	$1.6	$3.1	$408.6	$413.2
Any of the Five	$3,917	$18.2	$25.2	$3,306.4	$3,349.8

Study Methodology

We utilized Census data to attain demographic information by region. We used the Medical Expenditure Panel Survey (MEPS) to collect information concerning disease prevalence and insurance coverage. To model prevalence and insured status, a predictive model was created that adjusted for various demographic factors, including, age, sex, and income. The characteristics of each region of interest were placed into the model to made region specific projections.

Analytical Methodology

The prevalence rate model for diseases by region was developed in two stages. In the first stage, we developed a model for chance of occurrence of a disease based on 2004 (Medical Expenditure Panel Survey (MEPS) data. The second stage involves a "case-mix" adjustment to project the prevalence rate of each disease by race/ethnicity and region (Congressional district, State, MSA). Case-mix adjustment is by which projections for specific area are made by specifying the region specific demographic statistics in the national model.

A. Prevalence Rate Modeling

Chronic diseases place major strains on the community in terms healthcare costs and loss of productivity. Measuring the financial impact of major chronic diseases on individual communities can help quantify this strain on the community. The five major chronic diseases studied in this report are – asthma, cancer, depression, diabetes and cardiovascular disease.

Let 'p_i' denote the chance of occurrence of the "i^{th}" disease. A purpose of the study was to get an estimate of the prevalence rate of each disease by race/ethnicity. A logistic model was fit for the prevalence rate –

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 income + \beta_2 age + \beta_3 race + \beta_4 gender$$

with the demographic variables income, age, race and gender as explanatory variables for prevalence of a disease.

MEPS data from 2005 was used to fit the models for the 5 diseases. Other than the 5 models for the 5 diseases, we also modeled the chance of occurrence of either one of these diseases (an overall model).

B. "Case-mix" Adjustment Prediction

"Case-mix" Adjustment for the Congressional Districts (CD) Prediction –

Having modeled the chance of occurrence of a disease, the next goal was to project the prevalence rate of the disease by race/ethnicity and region. To obtain the projections, it is important to have the totals comparable to that reported in the census. Adjustments were made based on census numbers and CPS data to compute the prevalence rate of a disease by race/ethnicity and region (let us denote it by pvr_{eri}). CPS data was used to get the total population for MSA

regions and by race/ethnicity (let us denote that by n_{er}). The total projected prevalence number for a disease by region and race/ethnicity type can be calculated as $pvr_{eri} \times n_{er}$.

Note that CPS data from 2005 and 2006 were pooled to together to make MSA level estimates on the demographics. In order for the projections to make sense, the weights were rescaled so that the total population, as well as the population totals by age category, sex, and income summed to totals reported for the 2000 CPS March Supplement. This ensures that the national level demographics remain current. Next, 2006 MSA level projections were used to again rescale projections so that pooled MSA level projections added up to the known 2006 projections. This two stage rescaling should ensure a sample with weights that reflect the demographics of the area as well as accurate total population projections.

For congressional district, data from the 110th Congressional District 2000 Census Sample File was used for the demographic information need for the projections. Although this file was mapped to the current congressional district, the Census left the total population estimates at 200 levels. For this reason, the population estimates were adjusted to reflect the estimated population growth in the state that the congressional district resides in.

Calibrating Case Mix Adjustment Projections

A pitfall of regression based case mix adjustment is that projections are made "on the average". Essentially, a projected prevalence rate can be interpreted statistically the predicted probability of being covered for the average person in the region. The average person in the county is defined by the demographics of the region, thus this mythical person is 60% white, 30% African American, 51% male, 49% female, etc. The problem with this is that the average person does not typically have these diseases. More often, those with these chronic diseases fall into numerous high risk categories, as opposed a certain percentage of them falling into these categories.

In order to circumvent this pitfall, the prevalence projections are rescaled by a factor that reflects this systematic under-prediction by the logistic regression based case mix adjustment method. In addition, the logistic regression is dominated by the white population, even with the adjustment for ethnicity in the model. Hence, an adjustment factor for each ethnicity is created. The adjustment factor is determined by the ratio of the estimated prevalence for the United States over the national projected prevalence rate using the logistic model. For the national projected prevalence from the logistic regression model, the national demographics for the control categories were utilized.

C. Projection of Healthcare cost

MEPS data was used to calculate the average healthcare cost due to each disease by race/ethnicity. It is to be noted that healthcare cost for each disease type is not readily available and it is difficult to parse out the expense by each disease type for an individual. Hence the calculation for healthcare cost for a disease type is not straight forward. The method we used to predict the cost for each disease is as follows –

a) For each race/ethnicity calculate the average healthcare cost for having a particular disease (let us denote this by C^{D}_{ei}) and not having the disease (C^{ND}_{ei}). The difference between these two costs (C^{D}_{ei} - C^{ND}_{ei}) gives an estimate of the average cost per person due to the i^{th} disease.
b) The total estimated cost for the disease by race/ethnicity is then –

> Total prevalence number × estimate of average cost per person due to the disease
>
> $= (pvr_{ei} \times n_e) \times (C^{D}_{ei} - C^{ND}_{ei})$

An alternate strategy would be to utilize published studies on the cost of these various diseases. These studies typically measure both direct and indirect costs of disease. The direct costs for these diseases typically attempt to measure the healthcare costs incurred due to an illness, excluding the effect of other co-morbidities. Second, the indirect costs tend to measure loss of income and productivity due to the disease.

So which method is preferred? That is debatable, but much of that decision lies in the trust and belief in the previous studies. Much work goes into determining the causality of diseases. Many might argue that by simply "adjusting" for co-morbidities, one would greatly understate the effect of the chronic disease being modeled. Very complex modeling techniques could be employed to attempt to differentiate this causality issue. However, very rarely does the literature explain in depth the exact methods employed in the modeling procedure. Therefore, if one trusts the source of the literature, then using this meta-analysis approach seems sound. If however, one is not certain of the reliability of these estimates, then the first method outlined in this section is perfectly sound. This first method makes no presumption of causality, and simply states the statistics on healthcare costs derived from the MEPS data.

For this project, we utilized the existing literature to make inference on the cost of the chronic diseases. Sources for these costs are listed below:

1. Heart Disease and Stroke Statistics — 2007 Update, American Heart Association
2. Estimating the Cost of Cancer: Results on the Basis of Claims Data Analyses for Cancer Patients Diagnosed With Seven Types of Cancer During 1999 to 2000. Stella Chang, Stacey R. Long, Lucie Kutikova,

Lee Bowman, Denise Finley, William H. Crown, and Charles L. Bennett. Journal of Clinical Oncology. September 2004.
3. A National Estimate of the Economic Costs of Asthma. David H. Smith, Daniel C. Malone., Kenneth A Lawson, Lynn J. Okamoto, Carmelina A. Batista, and William B. Saunders. AM J RESPIR CRIT CARE MED 1997;156:787–793.
4. Economic Costs of Diabetes in the U.S. in 2002. Paul Hogan, Tim Dall, Plamen Nikolov. The Lewin Group. DIABETES CARE, VOLUME 26, NUMBER 3, MARCH 2003
5. Lost Productive Time and Cost Due to Common Pain Conditions in the US Workforce**.** Walter F. Stewart; Judith A. Ricci, Elsbeth Chee, David Morganstein, Richard Lipton. *JAMA*. 2003, 290:2443-2454.

D. National Statistics in Reports

In this section, we address a noted discrepancy in the tables produced for the individual reports. Tables 2 and 3 reference the prevalence and cost estimates for congressional districts. Additionally, national figures are also included. The National prevalence rates are projections from the 2004 Medical Expenditure Survey. The national cost figures shown are from historical reports on each disease. Thus, the two national figures are not tied into one another. The reason for this is that we wanted to give the most recent prevalence projections available form the singular MEPS source. The report could have applied an average per person cost estimate to the prevalence rates according to MEPS, but this would result in slightly different total cost estimates compared to historical report findings. As opposed to making adjustments to these well known studies, this total cost column is a simple reference to these historical reports.