STATISTICAL NEWS

PA Department of Health ◆ Bureau of Health Statistics and Research ◆ Vol. 31 No. 1 ◆ Jan/Feb 2008

Starting Prenatal Care in the 1st Trimester Declines

Young Mothers (<20) Had Lowest Percentages for 1st Trimester Care

Among Pennsylvania resident live births, the percent of mothers who received prenatal care in the first trimester of pregnancy decreased in 2005. This marked the second consecutive percent decrease since the 2003 revision to the certificate of live birth. Young mothers (under 20) tended to have the lowest percentages of first trimester prenatal care, compared to mothers aged 20 and older. It is beneficial that mothers begin prenatal care in the first trimester of pregnancy. For Pennsylvania resident live births in 2005 to mothers who obtained prenatal care in the first trimester, only 7.2 percent resulted in low birth weight babies. For births to mothers who did not receive prenatal care in the first trimester, a significantly higher percentage (9.1) resulted in low birth weight babies.

In 2005, 81.1 percent (97,194) of live births were to mothers who received prenatal care in the first trimester of pregnancy among Pennsylvania residents – just slightly lower than the 81.3 percent (97,316) reported in 2004.

Pennsylvania mothers who did not receive prenatal care until the second trimester accounted for 14.4 percent of resident live births in 2005. Mothers who waited until the third trimester of pregnancy to receive prenatal care accounted for 3.2 percent of resident live births and 1.3 percent were to mothers who did not receive any prenatal care in 2005. The following reviews statistics of resident live births to mothers who obtained prenatal care in the first trimester by various demographic features.

This (2005) marked the second consecutive percent decrease since the 2003 revision to the certificate of live birth.

Race/Ethnicity:

The percent of mothers who received prenatal care in the first trimester was lower among Black and Hispanic residents than Whites and Asians. During 2005, the percent of first trimester prenatal care was only 66.4 for Blacks and 66.7 for those of Hispanic origin (note that persons of Hispanic origin can be of

Go to Page 4 or click here...

The Use of Prescription Antibiotics in Pennsylvania

Four 2006 BRFSS Questions on Prescription Antibiotics Reviewed

Bacteriological resistance to antibiotics poses a potential public health threat. Two major factors in promoting bacteriological resistance to antibiotics are physicians over-prescribing them and patients inappropriately using the antibiotics. To obtain some basic information on the medical utilization of antibiotics by Pennsylvanians, four questions regarding prescription antibiotics were added to the 2006 Pennsylvania Behavioral Risk Factor Surveillance System (BRFSS) survey. The following is a review of the answers from those four questions.

When Pennsylvania adults were asked if they had been prescribed antibiotics for an illness in the past 12 months, 38.1 percent (CI: 36.4 - 39.8%) of the residents responded *'YES'*. Groups with a statistically significantly lower likelihood of using antibiotics in the past year are males, African Americans, adults without health care coverage, non-smokers and those reporting good to excellent heath (Chart 2 on page 6).

Of those who had been prescribed antibiotics in the past year, 43.9 percent (CI: 41.1

Go to Page 5 or click here...

Two major factors in promoting bacteriological resistance to antibiotics are physicians over-prescribing them and patients inappropriately using the antibiotics.

INSIDE THIS ISSUE

DEPARTMENT OF HEALTH

Edward G. Rendell, Governor

Data on Gestational Diabetes Reviewed

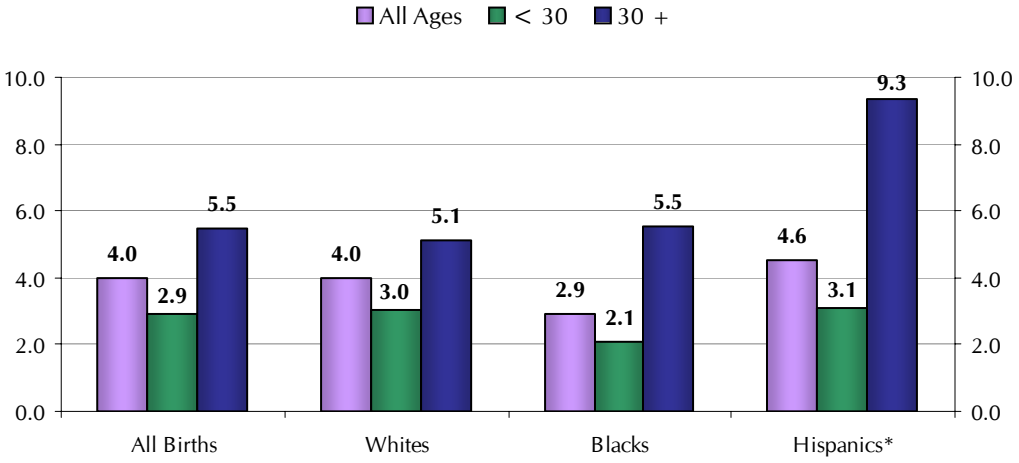
Gestational Diabetes Increases Risk for Type 2 Diabetes After Childbirth

In 2003, the Certificate of Live Birth was revised to include more detailed questions to capture more analytical information. One of the questions modified on the new certificate was the question concerning diagnosis of diabetes in the mother. The birth certificate prior to 2003 simply asked if diabetes was present in the mother during pregnancy. The certificate now asks if diabetes in the mother was present prior to pregnancy (pre-pregnancy diabetes) or as a result of the pregnancy (gestational diabetes). As a result of collecting more specific information, all Pennsylvania data concerning diabetes in the mother collected via the Certificate of Live Birth prior to 2003 is not comparable to data collected during or after 2003.

During the period of 2003-2005, there were a total of 20,143 live births to Pennsylvania mothers with pre-pregnancy or gestational diabetes out of 434,712 resident births – a percentage of 4.6. Of the 20,143 resident mothers with some form of diabetes, 17,285 (or 85.8 percent) were mothers who had gestational diabetes. The remainder of this article will focus on gestational diabetes.

In Pennsylvania, 4.0 percent of mothers who had a live birth experienced gestational diabetes during the years 2003, 2004, and 2005.

Chart 1
Percent of Live Births with Gestional Diabetes Listed as a Risk Factor By Race/Ethnicity and Age, Pennsylvania Residents, 2005



* Hispanics can be of any race

According to a report released by the Centers for Disease Control and Prevention (CDC) in 2003, "estimates of the overall prevalence of gestational diabetes in the United States range from at least 2.5% to 4% of pregnancies that result in singleton live births, with higher percentages among some ethnic groups and older women. Most gestational diabetes occurs in women with risk factors for type 2 diabetes; they are unable to secrete sufficient insulin to overcome the increased insulin resistance that normally results as pregnancy proceeds. Gestational diabetes usually ends after the baby is born, but women with gestational diabetes have a 20% to 50% chance of developing type 2 diabetes in the 5-10 years after childbirth."

In Pennsylvania, 4.0 percent of mothers who had a live birth experienced gestational diabetes during the years 2003, 2004, and 2005. Among births to older mothers in Pennsylvania (aged 30 and older), the percentage with gestational diabetes was 5.5 in 2005, compared to 2.9 for all resident mothers under the age of 30 – almost twice as high. It is important to note, however, that diabetes is a condition somewhat more common among older persons.

Table 1 (page 3) displays a definitive divide between mothers under 30 and those mothers 30 and older. Between 2003 and 2005, there was a slight increase in the percentage of older mothers who experienced gestational diabetes compared to their younger counterparts (from 3.0 in 2003 to 2.9 in 2005 for younger mothers compared to 5.3 in 2003 to 5.5 in 2005 for older mothers).

As shown in Table 1, the divide between younger and older mothers also holds true for Whites and was even more prominent among Black and Hispanic mothers for the period of 2003 through 2005. During 2005, White and Black mothers age 30 and older with gestational diabetes (5.1 percent and 5.5 percent, respectively) had percentages 1.5 to 2.5 times greater than for mothers under 30 years of age (3.0 percent for Whites and 2.1 percent for Blacks). Hispanics had an even higher difference between younger and older mothers with gestational diabetes in 2005 (3.1 percent of mothers under 30 compared to 9.3 percent of mothers 30 and over – a difference three times higher for older than for younger Hispanic mothers). This is illustrated in Chart 1.

Another interesting point is the fact that both Whites and Hispanics experienced increases in the percentage of older mothers with gestational diabetes between 2004 and 2005. They also both experienced decreases in the percentage of younger mothers during the same time frame. Black mothers, on the other hand, saw a decrease in the percentage of older mothers between 2004 and

Continued on Page 3...

Tools of the Trade:

Strategy for Determining Sample Size

One of the most frequently asked questions in survey statistics is, "What size sample do I need?". A sample that is too large may waste time, money, and resources. A sample that is too small may lead to inaccurate and unreliable results. Most people expect a simple answer to this question but in reality the methods used to determine the sample size are influenced by a number of factors which include the precision, confidence level, expected eligibility rate, expected completion rate, expected variability of the results, survey design and the way the data will be analyzed. In the following paragraphs, a method to determine sample size will be presented using the Synar survey as an example. This method can be used for a number of different designs with minor modifications.

Before we begin, some background information about the Synar survey must be established. The survey uses a stratified and clustered design with the sampling units defined as retail outlets that sell cigarettes over-the-counter. The objective of the survey is to produce a statewide estimate for the rate that retail outlets sell cigarettes to minors. The sampling design is two stages with the first stage involving the selection of clusters within stratum using probability proportionate to size of the cluster. The second stage involves selecting 17 outlets from the sampled clusters. The survey protocol requires youth to enter the outlet and attempt to purchase cigarettes. The Synar survey is federally regulated and must adhere to federal requirements that aren't normally present in other surveys. Lastly, the survey does not analyze sub-populations. If sub-populations are the goal, more sample would have to be added.

Determining Sample Size

Step 1: Set precision and confidence levels.
Step 2: Calculate the effective sample size.
Step 3: Determine design effect and adjust sample size.
Step 4: Estimate eligibility & completion rates; adjust sample size.
Step 5: Adjust sample for variability and equal sample allocation.

Precision and Confidence Levels

The first step is to decide the level of precision and confidence to use for the survey. The confidence level refers to the probability value associated with a confidence interval. A confidence level is necessary when the results will be presented using confidence intervals. Common confidence levels are 90 or 95 percent for either a two-sided or one-sided interval. The confidence level is directly proportional to sample size, i.e. more sample is required for higher levels of confidence. The Synar survey is required by a federal regulation to use a one-sided 95% confidence interval.

Precision refers to the survey error. The amount of error that you are willing to allow must be set prior to the start of the survey. The

Continued on Page 8...

Continued from Page 2...

Data on Gestational Diabetes...

2005 (6.3 and 5.5, respectively). The percentage of younger Black mothers with gestational diabetes remained at 2.1 percent for both years.

Between 2003 & 2005, there was a slight increase in the percentage of older mothers who experienced gestational diabetes compared to their younger counterparts...

If you have any questions regarding national diabetes data, please visit the Centers for Disease Control and Prevention website at ***www.cdc.gov/diabetes***. If you have questions concerning the data presented in this article on the incidence of diabetes as a risk factor for Pennsylvania women giving birth, please contact the Bureau at 717-783-2548 or send an email from the Health Statistics web pages at ***www.health.state.pa.us/stats***. Also, additional Pennsylvania birth statistics can be accessed using the above link for the Health Statistics web pages.

Table 1
Percent of Live Births with Gestational Diabetes Listed as a Risk Factor, By Race/Ethnicity, Age, and Year Pennsylvania Residents, 2003-2005

	2003	2004	2005
All Births	4.0	4.0	4.0
Births to Mothers <30	3.0	3.0	2.9
Births to Mothers 30+	5.3	5.4	5.5
Whites	4.0	4.0	4.0
Births to Mothers <30	3.1	3.2	3.0
Births to Mothers 30+	5.0	5.0	5.1
Blacks	3.1	3.2	2.9
Births to Mothers <30	2.4	2.1	2.1
Births to Mothers 30+	5.3	6.3	5.5
Hispanics*	4.5	4.6	4.6
Births to Mothers <30	3.3	3.5	3.1
Births to Mothers 30+	8.5	7.8	9.3

*Hispanics can be of any race.

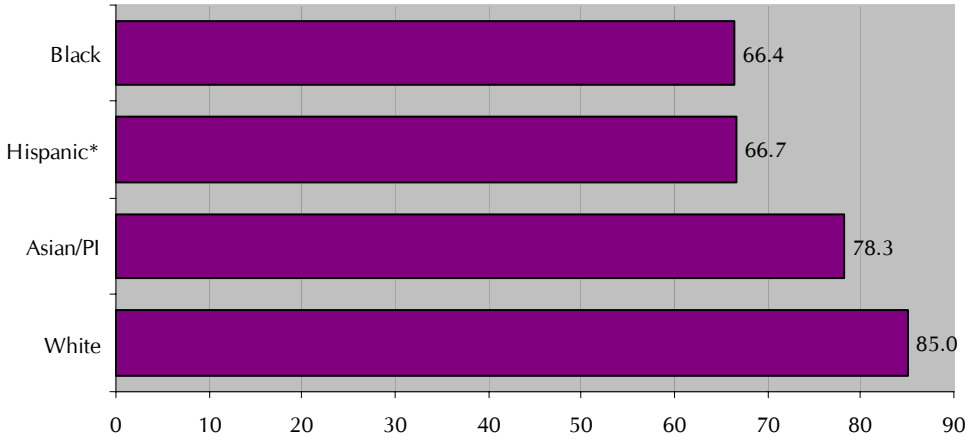
Continued from Page 1...

Starting Prenatal Care in the 1st Trimester Declines

any race). These percentages are much lower than the percentage for Whites (85.0). The percentage for Asian/Pacific Islanders (78.3) was lower than the White percentage and the overall state percentage (81.1), but not nearly as low as the percentages for Blacks and those of Hispanic origin (see Chart 1). The percentage of Black mothers who did not receive any prenatal care in 2005 was 4.1, while the overall percentage for no prenatal care for the state was 1.3.

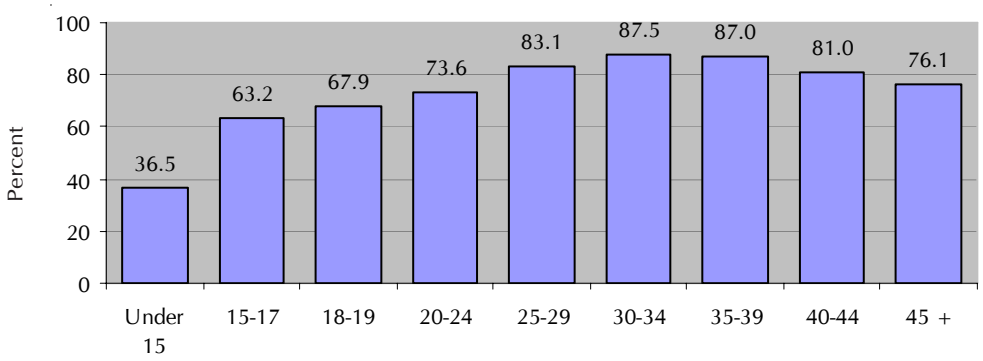
The percentage of Black mothers who did not receive any prenatal care in 2005 was 4.1, while the overall percentage for no prenatal care for the state was 1.3.

Chart 1
Percent of Mothers Obtaining First Trimester Prenatal Care by Race/Ethnicity
Pennsylvania Resident Live Births, 2005



* Hispanics can be of any race

Chart 2
Percent of Mothers Obtaining First Trimester Prenatal Care by Age Group
Pennsylvania Resident Live Births, 2005



Age:
In 2005, resident live births to mothers in their 30s had the highest percentages of obtaining prenatal care in the first trimester compared to other age groups. The age group 30-34 had the highest percentage (87.5), followed closely by those ages 35-39 (87.0). Mothers less than 20 years old had the lowest percentages of first trimester prenatal care, with the lowest percentage coming from mothers under 15 years old (36.5). Chart 2 shows the percentages were very low for young mothers but then increased steadily with age until age 35, when they slowly declined.

Marital Status/Medicaid:
Mothers who were married as well as mothers who were not on Medicaid were more likely to obtain prenatal care in the first trimester of pregnancy than those who were not married and those who were on Medicaid. For Pennsylvania resident live births in 2005, 86.8 percent of married mothers obtained prenatal care in the first trimester, while only 70.6 percent of unmarried mothers obtained first trimester prenatal care. Almost 85 percent of mothers who were not on Medicaid obtained first trimester prenatal care. The same percentage for mothers who were on Medicaid was only 69.3 percent.

County:
More counties in the western part of the state had higher percentages of births to mothers who obtained first trimester prenatal care in 2005. The county with the highest percentage in 2005 was Allegheny County (91.2 percent). The second highest percentage was a tie between residents of Armstrong and Elk counties (89.4 percent). Rounding out the top five was Westmoreland County followed by Clearfield County. The percentages for these counties were all significantly higher than the state percentage.

Continued on Page 5...

Continued from Page 1...

The Use of Prescription Antibiotics in Pennsylvania

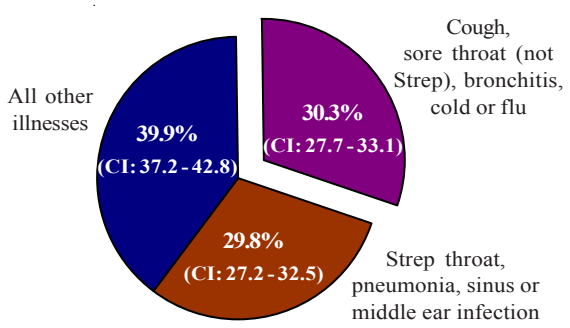
46.7%) had antibiotics prescribed on more than one occasion. Of these, females were significantly more likely than males to have antibiotics prescribed more than once.

The adults who had been prescribed antibiotics in the past 12 months were asked: *"What illness did you have the last time you took antibiotics?"* After removing the approximately three percent who responded *'don't know/not sure'*, about three out of ten Pennsylvania adults responded that they took antibiotics for a cough, sore throat, bronchitis, cold or the flu. These illnesses are most often caused by viruses, for which antibiotics are ineffective. For the purposes of our examination we considered these uses to be inappropriate for treatment with antibiotics while assuming all other uses of antibiotics to be appropriate.

Males were significantly more likely than females to re-

Continued on Page 6...

Chart 1
Illness for Which Antibiotics were Prescribed
Pennsylvania BRFSS, 2006

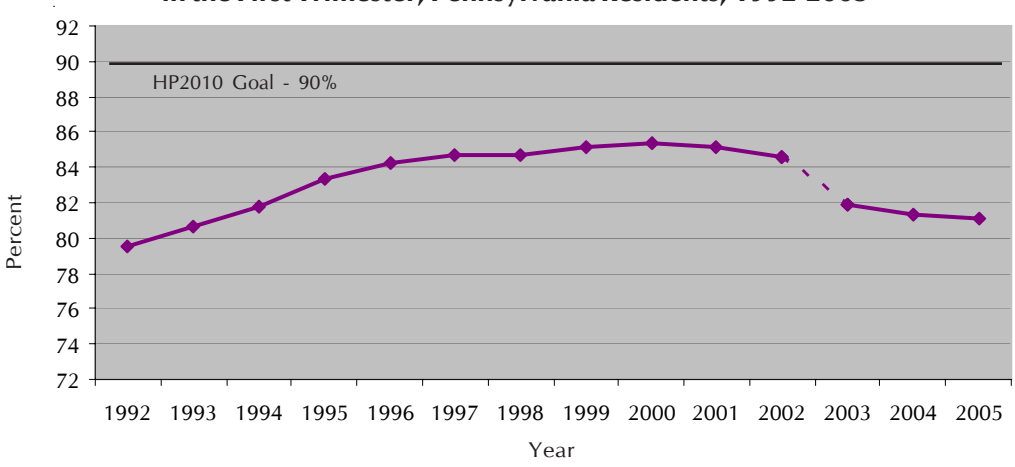


Continued from Page 4...

Starting Prenatal Care in the 1st Trimester...

Philadelphia and counties in the central portion of the state had the lowest percentages of births to mothers who obtained prenatal care in the first trimester. Philadelphia County had the lowest percentage (66.6) in 2005. The next lowest percentage was for residents of Mifflin County (71.1), followed by Juniata (72.0), Pike (74.1), and Franklin (74.9) counties. These five counties all had significantly lower percentages compared to the state.

Chart 3
Percent of Live Births to Mothers Who Obtained Prenatal Care in the First Trimester, Pennsylvania Residents, 1992-2005



Trends:

Please note that in years prior to 2003 the mother or prenatal care provider reported the month of pregnancy in which the mother began prenatal care. Starting in 2003, this item was replaced by the exact dates of first and last prenatal visit. Therefore the month prenatal care began is now calculated from the last normal menses date and the date of first prenatal care visit. This change will cause some inconsistency with the trend data displayed in Chart 3.

The percent of first trimester prenatal care among live births to Pennsylvania resident mothers decreased in 2005 for the fifth consecutive year (or second consecutive year since the 2003 revision to the certificate of live birth). The percent has dropped from 85.4 in 2000 to 81.1 in 2005. However, the percent has increased since 1992 when the percent was 79.5. The percent of prenatal care in the first trimester increased or held steady each year from 1992 to 2000, but starting in 2001 it has declined each year. The Healthy People 2010 goal for births to mothers beginning prenatal care in the first trimester is 90 percent. With the percent decreasing in recent years, Pennsylvania is headed in the wrong direction to meet this important goal.

For questions about this article, please contact the Bureau at 717-783-2548. Additional birth statistics can be accessed on the Health Statistics web pages at ***www.health.state.pa.us/stats***.

Continued from Page 5...

The Use of Prescription Antibiotics...

ceive antibiotics for an inappropriate illness. However, males were significantly less likely than females to be prescribed antibiotics and, of those who were prescribed antibiotics in the past year, males were significantly less likely to receive antibiotics more than once.[1]

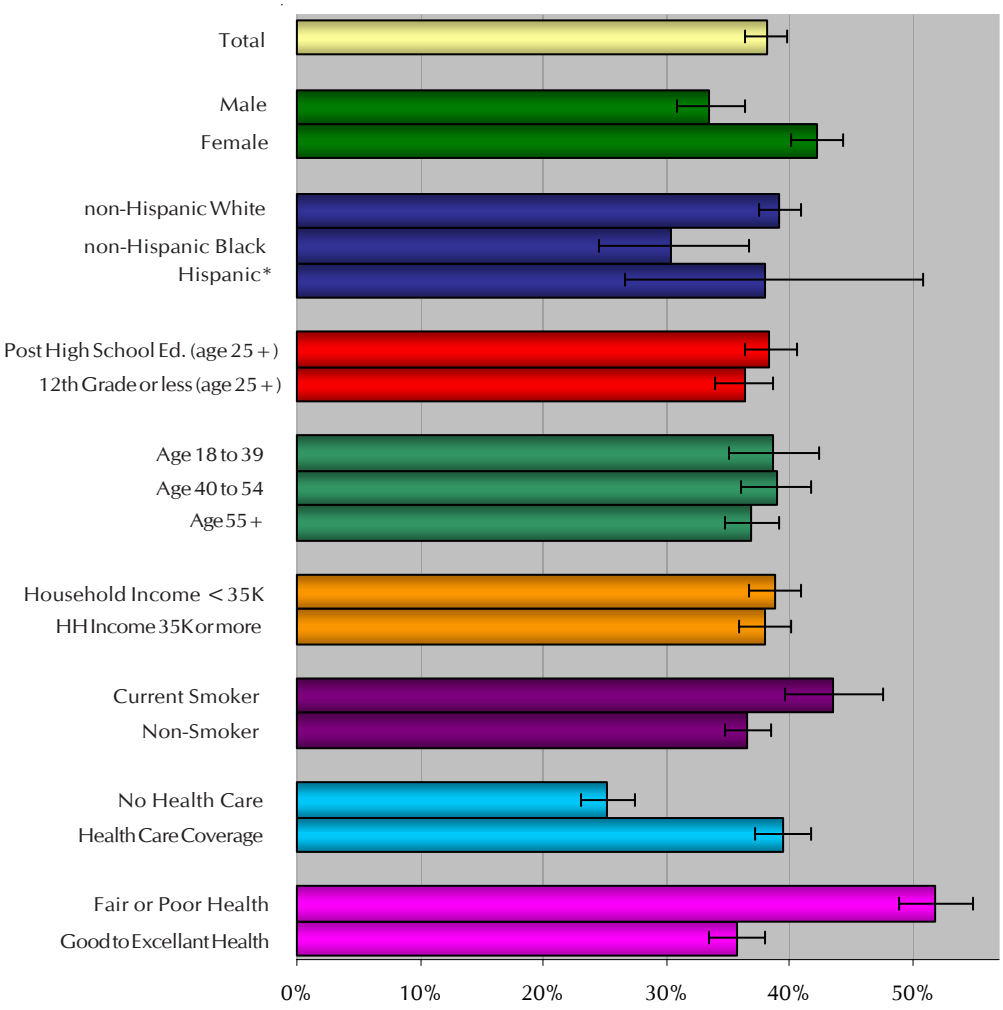
In this age of increasing strains of multiple antibiotic resistant bacteria, it is disturbing to find that a quarter to one third of antibiotics may be inappropriately prescribed. Further aggravating the problem of multiple resistant bacteria is that 11.4 percent (CI: 9.6 - 13.6%) of persons prescribed antibiotics did not complete the course of antibiotics prescribed.

Please contact the Bureau at 717-783-2548 if you have any questions about this article. You can access additional BRFSS data by using our interactive web tool called EpiQMS at ***www.health.state.pa.us/stats***.

[1] Statistical significance was tested using SUDAAN to account for the sample error in a manner analogous to a Chi-Square test of association. Unless otherwise noted the probability of no association is rejected at the P<.01 level.

Chart 2

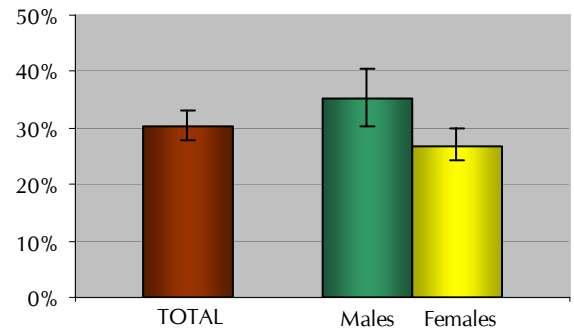
Percent Prevalence of PA Adults Prescribed Antibiotics within Past 12 Months
Pennsylvania BRFSS 2006



* Hispanics can be of any race

Chart 3

% of Adults Prescribed Antibiotics Inappropiately*
For Most Recent Time in the Past 12 Months
Pennsylvania BRFSS, 2006



* prescibed for: cough, sore throat (not Strep), bronchitis, or cold/flu

The **Behavioral Risk Factor Surveillance System (BRFSS)** is a public health surveillance system that is conducted in Pennsylvania and in all other states with support from the Centers for Disease Control and Prevention (CDC). Its purpose is to collect sample data on risk behaviors linked to chronic disease, injury, and infectious diseases as well as preventive health practices supportive of community health. The BRFSS survey consists of telephone interviews using randomly generated telephone numbers to determine the households contacted. The survey contains a core set of questions provided by CDC to gather comprehensive, standard information nationwide and questions added by the Pennsylvania Department of Health to obtain information needed by the Department that is unavailable elsewhere.

Update: Healthy People 2010 Objectives

Focus Area 19: Nutrition and Overweight

19-01 - Increase % healthy weight adults (ages 20+).....HP2010 Target: 60%

19-02 - Reduce % obese adults (ages 20+).....HP2010 Target: 15%

Healthy Weight Adults:
The percent of healthy weight adults (ages 20+) has not changed much between 2002 and 2006 according to the Pennsylvania Behavioral Risk Factor Surveillance System (BRFSS) annual sample survey of adults in the state. Among all adults, the percentages ranged from 36 in 2005 to 38 in 2003 and 2002.

The percentages of healthy weight adults were much lower for non-Hispanic Blacks (31 in 2006) and males (30), compared to non-Hispanic Whites (38) and, especially, females (45). Figures for Hispanics had large fluctuations due to small sample sizes – the highest (37) in 2002 and lowest (27) in 2006 and 2003.

Percentages for non-Hispanic Blacks appear to be on the increase, but it is unlikely that they or any of the other groups will achieve the HP2010 goal.

Obese Adults:
In 2006, 24 percent of the Pennsylvania adult population was considered obese, down from 26 percent in 2005. The percentages for non-Hispanic Black adults were consistently higher than for non-Hispanic Whites, males, and females – 33 percent in 2006, compared to 24 percent for non-Hispanic Whites, 24 for females, and 25 for males. However, the percentages for non-Hispanic Blacks has declined in recent years. No obvious trends were observed among the other sex/race groups. The percentages for Hispanic residents are difficult to interpret due to the small sample sizes for this ethnic group.

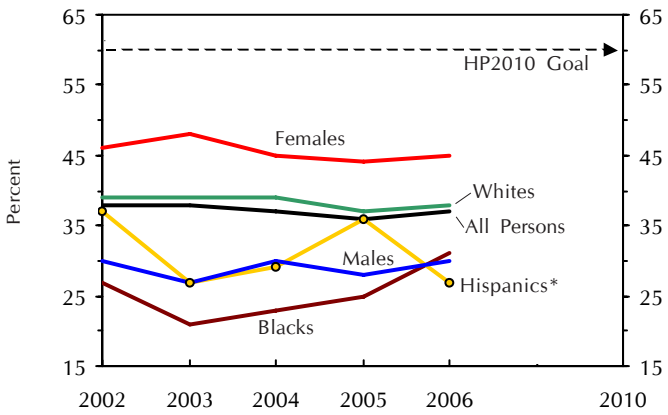
The Healthy People 2010 goal is set for 15 percent. In Pennsylvania, it is unlikely for any of the groups to achieve this goal over the next four years.

Percent Healthy Weight and Obese Adults Ages 20+ By Race/Ethnicity and Sex, Pennsylvania 2002-2006

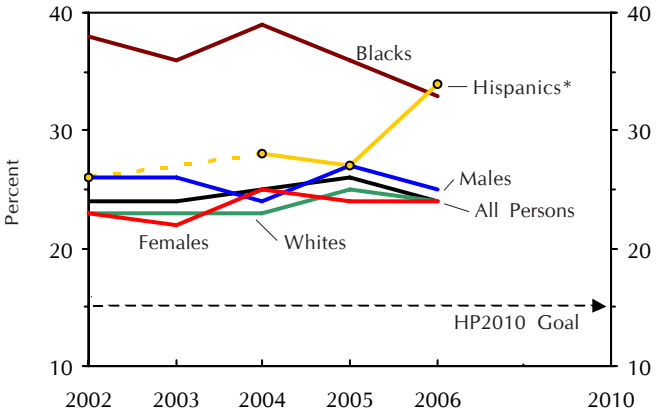
Healthy Weight	2002	2003	2004	2005	2006
All Adults 20+	38±1	38±2	37±2	36±1	37±2
Non-Hispanic Whites	39±1	39±2	39±2	37±1	38±2
Non-Hispanic Blacks	27±5	21±7	23±6	25±5	31±8
Hispanics	37±7	27±12	29±10	36±10	27±10
Males	30±2	27±3	30±2	28±2	30±3
Females	46±2	48±3	45±2	44±2	45±2
Obese					
All Adults 20+	24±1	24±2	25±1	26±1	24±1
Non-Hispanic Whites	23±1	23±2	23±1	25±1	24±2
Non-Hispanic Blacks	38±5	36±8	39±6	36±5	33±7
Hispanics	26±8	DSU	28±11	27±8	34±11
Males	26±2	26±3	24±2	27±2	25±2
Females	23±1	22±2	25±2	24±1	24±2

NOTES: ± denotes 95% confidence interval. Hispanics can be of any race.
DSU = data statistically unreliable. Percents are age-adjusted to 2000 std population.

Percent Healthy Weight Adults Ages 20+ By Race Ethnicity and Sex, Pennsylvania 2002-2006



Percent Obese Adults Ages 20+ By Race/Ethnicity and Sex, Pennsylvania 2002-2006



* Hispanics can be of any race

HP2010 State and County Data on the Web

To access the Department of Health's web page of Healthy People 2010 statistics for the state and counties, go to *www.health.state.pa.us/stats*. The latest available statistics as well as trend data are shown. You can view data for the state, all counties, a specific demographic element (age, sex, race, etc.) or just for a specific county. Complete data sets for the state and counties can be downloaded. There is also a link to the national HP2010 web site.

Continued from Page 3...

Strategy for Determining Sample Size

precision level is inversely proportional to sample size. The lower the precision level is set, the more sample will be needed to reach that goal. Obviously survey error is undesirable and every survey should strive for as little error as possible when conducting the survey. While determining the sample size, the level should be set as low as possible without increasing the sample size to more than your resources will allow. A federal regulation requires the Synar error to be less than or equal to three percent.

Effective sample size

Calculate the effective sample size using the precision and confidence levels established in Step 1. The effective sample size is the minimum sample size needed to meet the precision and confidence requirements under a simple random design. Ignoring the finite population correction factor, the effective sample size equation is derived from the precision equation "$w = z(s.e.)$" and the standard error equation "$s.e. = \frac{\sqrt{p(1-p)}}{\sqrt{n_e}}$". The effective sample size equation is:

$$n_e = \left(\frac{z}{w}\right)^2 p(1-p) \qquad \textbf{(E1)}$$

Where z is the critical value of the standard normal distribution for a one-sided 95% confidence interval, w is the precision and p is the expected rate for the variable being measured.

For the Synar survey, z is 1.645 according to a critical value table and w is .03 since the precision is set at three percent. The expected sale rate must be estimated and should be set at a level that will produce the highest number of sample that can be afforded. It is better to estimate too high than to estimate too low. As displayed in the equation, the highest amount of sample occurs when the sale rate is 50 percent. The Synar survey isn't accepted if the statewide estimate is greater than 23 percent (20%+3% error=23%). Therefore, 23 percent was used for the value of p since it is the closest we can get to 50 percent. The effective sample size for the 2007 Synar survey is determined by substituting these numbers into **(E1)** and solving for n_e:

$$n_e = \left(\frac{z}{w}\right)^2 p(1-p) = \left(\frac{1.645}{.03}\right)^2 \times .23(1-.23) = 532.5 \approx 533$$

Design Effect

Since most surveys are not truly random, the sample size must be adjusted to account for deviations from a simple random design. These deviations are known as the design effect (DEFF). The higher the design effect, the larger the sample size needs to be. The DEFF is calculated by dividing the variance of the survey with the complex design by the variance of a simple random sample of the same size. Since the DEFF can not be calculated before the survey is conducted, the DEFF has to be estimated by gathering information from surveys with similar designs and calculating their DEFFs. With this information and the knowledge of your own survey's circumstances, make an informed estimate for the DEFF. The DEFF equation is:

$$n_t = \text{Deff}_h \times n_e \qquad \textbf{(E2)}$$

Where n_e is the effective sample size and $Deff_h$ is the highest DEFF of the last three surveys. Solving **(E2)** with $Deff_h$=1.29, the increase sample size for design effect is:

$$n_t = \text{Deff}_h \times n_e = 1.29 \times 533 = 687.57 \approx 688$$

Eligibility and Completion Rates

The sample size is adjusted to account for ineligible sample and non-completions. The eligibility and completion rates have to be estimated. The eligibility rate is the percent of the sample that is eligible and the completion rate is the percent of eligible sample completed. The best estimates come from previous versions of the same survey or other surveys with similar designs. If this information is available it should be used. If past surveys aren't available, then make your best guess. It is better to underestimate these rates than to overestimate them. If you assume your completion/eligibility rates will be high, you will need less sample but if they actually turn out to be low, you won't meet your precision goal because the sample will be too small.

For the Synar survey, the rates of the past three surveys were reviewed and 68 percent was used for the eligibility rate and 80 percent for the completion rate. The equation used to increase the sample to account for ineligible sample and non-completions is:

$$n_{ec} = \frac{n_t}{r_l r_c} \qquad \textbf{(E3)}$$

Where r_l= lowest eligibility rate of historical Synar surveys of a similar design (2004-2006) and r_c = 80 percent, the lowest completion rate allowed by the Federal government. Solving **(E3)** the sample size is increased to:

$$n_{ec} = \frac{n_t}{r_l r_c} = \frac{688}{(.68)(.80)} = 1264.7 \approx 1265$$

Original Sample Size

Lastly, the sample size needs to be adjusted for variability and equal sample allocation. This is the last step and the result is called the original sample size. Variability refers to the variability between clusters. One method to test variability is to simulate survey results repeatedly, using a computer programming language such as SAS (Statistical Analysis System), and study the results. SAS programs were used to test the variability for the Synar survey. If SAS isn't available, manually simulate the results a few times and examine them. It is recommended to add to the sample if you can afford it.

Continued on Page 9...

Continued from Page 8...

Strategy for Determining Sample Size

If the survey is clustered, it will have to be adjusted to fit the design so it can be equally distributed among the clusters. Consider a clustered design with 17 samples per cluster. The total sample size cannot be 1,265 because 1,265 can not be divided by 17 evenly. Therefore, the total sample has to be increased so each cluster has 17 samples.

After testing, it was determined that more sample was needed for the Synar Survey. The amount of sample was determined by viewing the program results and spreading it out among the clusters and random areas of the design. n_A =287 sampling units were added. The original sample size is determined using the following equation:

$$n_o = n_{ec} + n_A \qquad \textbf{(E4)}$$

By solving this equation, it was determined that the 2007 Synar survey needed the following sample size:

$$n_o = n_{ec} + n_A = 1265 + 287 = 1552$$

Final Equation

Although the above equations can be combined into one equation, each of the variables will still have to be determined individually. Combining the equations gives:

$$n_o = \left(\frac{Deff_h}{r_l r_c}\right) n_e + n_A \qquad \textbf{(E5)}$$

By solving the equation, one obtains a number that differs slightly due to rounding error.

$$n_o = \left(\frac{Deff_h}{r_l r_c}\right) n_e + n_A = \left(\frac{1.29}{(.68)(.80)}\right) 533 + 287 \approx 1551$$

For questions about this article, please contact the Bureau of Health Statistics and Research at 717-783-2548. To access additional "Tools of the Trade" reports, go to the Health Statistics web pages at ***www.health.state.pa.us/stats*** and select "Technical Assistance".

Statistical News *is published bimonthly by the Bureau of Health Statistics and Research, Pennsylvania Department of Health, 555 Walnut St., 6th Floor, Harrisburg, PA, 17101. Please write, telephone (717-783-2548) or FAX (717-772-3258) us if you have any questions regarding the contents of this newsletter. Visit the Health Statistics section of the Department's web site at* ***www.health.state.pa.us/stats*** *to access additional health statistics and reports.*

The Department of Health is an equal opportunity provider of grants, contracts, services and employment.

Click on the EpiQMS logo above to access our interactive health statistics web site.

H106.830P