Relationship between Insured Rate and Proportion of Individuals with a Personal Doctor

Andrew Deighan

April 17, 2017

Abstract

In this report we use a log-linear Poisson regression model to investigate the relationship between the rate of individuals having a personal doctor and the insured rate in states within the United States. In the past several years the insured rate in the United States has increased and this has given rise to many important questions. The topic of interest in this report is whether primary care access increases with insured rates and how so. Our analysis shows that with an increase in a state's insured rate by 1 percentage point, we expect an 1.012 (95% CI: 1.009, 1.015) increase in the mean rate of individuals with a personal doctor. This analysis serves as evidence to support the notion that increases in insured rates are accompanied by increases in primary care access.

<

>

Introduction

We are interested in how the rate of individuals with insurance is related to improved primary care access. Improving access to primary healthcare has been shown time and time again to improve health outcomes and lower overall healthcare system costs. From the county level to the national level, greater accessibility of primary care is strongly associated with increased general health along with decreased mortality rates, hospital admissions, overall healthcare system costs, and socioeconomic disparities [1]. Countries with governmental policies that are supportive of primary care have better primary care systems and subsequently better health outcomes at lower costs. One of the policies most consistently linked with better primary care is universal financial coverage of healthcare costs, provided by or regulated by the goverment [1]. Most industrialized countries provide universal or near-universal access to primary care. Notably, the United States is not included in this group. The United States does have some public programs (Medicaid and Medicare) that help pay healthcare costs for the elderly and poor, but for most Americans private health insurance is needed to cover healthcare costs.

In March 2010 the Patient Protection and Affordable Care Act (ACA) was signed into law. One of the primary objectives of this legislation was to decrease the uninsured rate in the United States. Although some reforms began in 2010 directly after the signing of the bill, most major reforms did not take effect until 2014 [2]. Beginning in 2014 the government funded Medicaid program was expanded in most states, new private insurance marketplaces selling government subsidized insurance plans were opened, and new consumer protections were placed on private insurance companies. During the first open-enrollment period of the ACA (Oct 1st 2013 through April 15th 2014) an estimated 20 million Americans gained insurance under the ACA. In roughly the same period of time the uninsured rate dropped from around 18% to about 13.4% [2]. Since then the uninsured rate has continued to steadily drop, down to an estimated 9.1% in 2015 [3].

With this recent boost in the proportion of Americans with health insurance come important questions about the overall effects of more people having health insurance. The primary question for us is whether or not an increse in the insured rate is correlated with an increase in primary care access, and, if so, to what degree. In an attempt to address this question we will investigate the relationship between the number of individuals with a personal doctor and the number of individuals with insurance at the state level. Thus we are using the number of individuals with a personal doctor as a proxy for primary care access. Specifically we are use Poisson regression to regress the estimated count of individuals with a personal doctor onto the insured rate. We are standardizing the estimated count to an estimate of the count *per 100 individuals* which in essence is the same as the percentage of individuals with a personal doctor. If there is indeed a positive correlation between the rate (percentage) of individuals with a personal doctor and the insured rate, then the recent reduction in the uninsured rate could herald a future overall improvement of health outcomes and lowere healthcare system costs.

In the Methods and Results section which follows we first discuss the datasets used, the cleaning of the data, the weighting of the data, and the standardizing of counts. We then briefly investigate the change in the insured rate from 2011 through 2015 to ensure that it matches with the previously described figues. After verifying the trend of increasing in insured rates we then move onto a discussion of the fitting and diagnostics processes used in building and validating our models. Finally, in the last section of the Methods and Results we interpret our models. In the Discussion section of the report we summarize our findings and explain their meanings and importance.

Methods and Results

Data Sets

The data used in this study were directly taken from the Behavioral Risk Factor Surveillance System (BRFSS) annual surveys from 2011 through 2015. The BRFSS survey is a joint effort between individual state and territory health departments and the CDC [4]. Every year, each state does a land-line and cell-phone survey of selected telephone numbers of residents in their state and asks a set of questions. The core questions asked by each state are the same. The data used in this study came from responses to core questions of the BRFSS Survey that all states were asked. Any refusals to respond or responses of "Unsure/Don't Know" for the response variable or the explanatory variable of interest were ignored. Thus, only full responses were included. However, the data sets are large so the fraction of incomplete responses was negligible.

The BRFSS data sets from 2011 through 2015 have data for each state as well as the District of Columbia, Guam, and Puerto Rico. Since we are working at the state level this gives us a total of 53 observations for each year. Within the data sets there were two questions that were of interest to us:

1. Insurance Status: "Do you have any kind of health care coverage, including health insurance, prepaid plans such as HMOs, or government plans such as Medicare, or Indian Health Service?""
   - 1: Yes
   - 2: No
   - 7: Don't know/Not Sure
   - 9: Refused
2. Personal Doctor: "Do you have one person you think of as your personal doctor or health care provider? (If"No" ask "Is there more than one or is there no person who you think of as your personal doctor or health care provider?".)"
   - 1: Yes, only one
   - 2: More than one
   - 3: No
   - 7: Don't know/Not Sure
   - 9: Refused

For the personal doctor question we counted both responses as "Yes, only one" and "More than one" as a count of an individual with a personal doctor.

Cleaning Data

Firstly we need to clean our data sets by removing all missing values. We will check how many observations we end up eliminating to be sure that this removal will not impact our analysis significantly. Table 1 reports the number of data points before and after cleaning for each year as well as the percentage removed. Since only a small fraction (at most 0.79%) were removed for any given year we accept that the effect of these removals will be negligible. We will also split each data set in half so we will have one half for fitting our model and one half for testing.

Table 1: **Table 1:** Number of missing values removed from data sets.

Year	Total Before Cleaning	Total After Cleaning	Number Removed	Percentage Removed
2011	506467	503325	3142	0.6204
2012	475687	472780	2907	0.6111
2013	491773	488240	3533	0.7184
2014	464664	461057	3607	0.7763

Year	Total Before Cleaning	Total After Cleaning	Number Removed	Percentage Removed
2015	441456	437959	3497	0.7922

Calculating Weighted and Standardized Counts

We are interested in how the percentage of individuals with a personal doctor in a state is related to the percentage of people with insurance. Viewing percentage of people with a personal doctor as a rate (a per capita rate) we can use Poisson regression to regress this rate on the pecentage of individuals with insurance. We will use each individual state/territory included in the BRFSS as an observation. In order to do this we need to get the counts of individuals with a personal doctor and the counts of individuals with insurance for each state/territory. When doing this though we need to make use of the recomended weights in the BRFSS.

For each observation (state/territory) we will have three values:

- Weighted counts of respondents with a personal doctor in that state
- Weighted counts of respondents with insurance in that state
- Total number of respondents from that state

For instance, the observation for the state of Maine for 2011 would look as follows

---

Table 2: 2011 observation for state of Maine

	Personal Doctor Count	Insurance Count	Total Respondents
ME	5800	5741	6578

We will standardize the values to counts per 100 respondents so the Maine 2011 observation will become:

Table 3: Standardized 2011 observation for state of Maine

	Personal Doctor Count	Insurance Count	Total Respondents
ME	88.17	87.28	6578

Since we are including all 50 states, the Distric of Colombia, Puerto Rico, and Guam we have 53 observations for each year.

Figure 1 below plots the standardized personal doctor counts against the standardized insurance counts for each state. The figure gives plots for each year, all years combined, the five year averages for each state, and the difference between 2015 and 2011. From all these plots it is clear that there is a positive correlation between the standardized counts of individuals with a personal doctor and the the standardized insured rate.

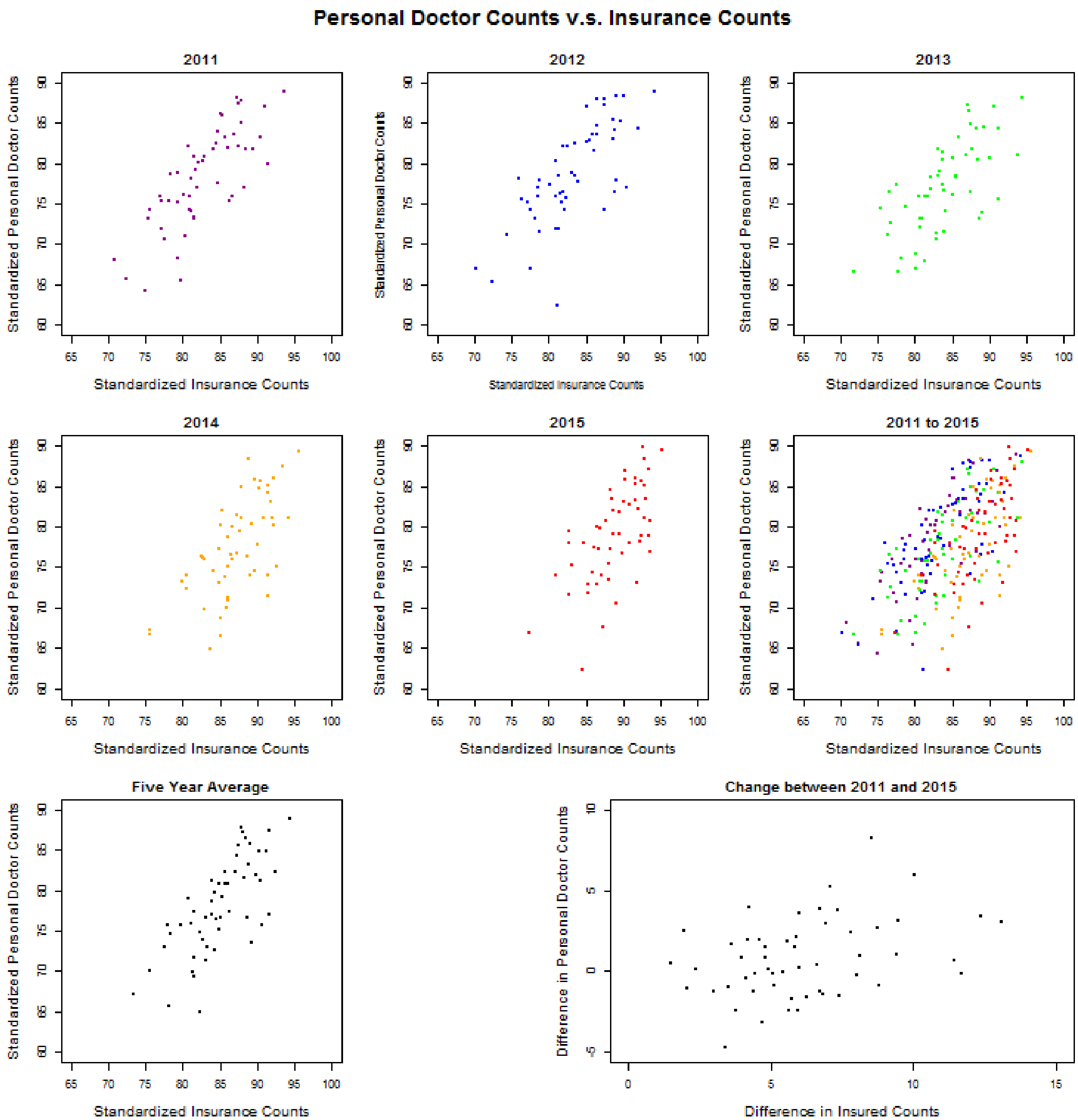

Figure 1: **Figure 1:** The plot above shows the standardized personal doctor counts plotted against the standardized insurance counts. The points in purple correspond to data from 2011, blue from 2012, green from 2013, orange from 2014, and red from 2015. The first five plots show each of these years individually while the sixth plot shows all the years combined. From the 6th plot it appears that there is a general shift upwards to the left which would correspond with increaseing insured rates and increasing counts of individuals with a personal doctor. The 7th plot shows the 2011 to 2015 average for all states. The 8th plot shows the difference in personal doctor counts and insurance counts between 2015 and 2011 for each state.

Yearly Insured Rates

Tables 2, 3 and 4 report the estimated yearly insured rates of adults for the nation as a whole, states that adopted medicaid expansion on January 1st 2014, and those states that had not adopted the expansion as of the end of 2015. These insured rates were estimated from the 2011 to 2015 BRFSS which is restricted to adults. This restriction may account for the estimated insured rates reported here being slightly less than those reported by the US Census Bureau. The percentage of insured adults does indeed rise from 2011 to 2015, even in states that did not adopt the Medicaid expansion. Moreover, there appears to be a particularly sharp increase from 2013 to 2015, which corresponds temporally with the initiation of most of the ACA reforms. However, states that did not adopt the Medicaid expansion had lower insured rates than those that did. Figures 2 and 3 show this graphically.

Table 4: **Table 2:** Estimated yearly adult insured rates (in percentages) for the nation as a whole. The table gives 95% confidence intervals for the percentage of insured adults as estimated from the 2011 to 2015 BRFSS surveys.

	Lower Bound	Estimate	Upper Bound	Standard Error
2011	83.03	83.13	83.24	0.0555
2012	83.66	83.77	83.88	0.0572
2013	83.86	83.97	84.08	0.0552
2014	87.21	87.31	87.41	0.0524
2015	88.87	88.97	89.07	0.0498

Table 5: **Table 3:** Estimated yearly adult insured rates (in percentages) for states that adopted the Medicaid expansion on January 1st 2014. The table gives 95% confidence intervals for the percentage of insured adults in states that adopted the Medicaid expansion, as estimated from the 2011 to 2015 BRFSS surveys.

	Lower Bound	Estimate	Upper Bound	Standard Error
2001	85.35	85.51	85.66	0.0782
2012	85.41	85.56	85.72	0.0805
2013	86.03	86.19	86.34	0.0788
2014	89.54	89.69	89.84	0.0757
2015	91.13	91.26	91.4	0.0676

Table 6: **Table 4:** Estimated yearly adult insured rates (in percentages) for states that had not adopted the Medicaid expansion as of the end of 2015. The table gives 95% confidence intervals for the percentage of insured adults in states that did not adopt the Medicaid expansion, as estimated from the 2011 to 2015 BRFSS surveys.

	Lower Bound	Estimate	Upper Bound	Standard Error
2011	80.16	80.33	80.51	0.0905
2012	80.91	81.1	81.3	0.0995
2013	81.07	81.25	81.42	0.0906
2014	84.57	84.74	84.91	0.0875

	Lower Bound	Estimate	Upper Bound	Standard Error
2015	86.05	86.21	86.38	0.0848

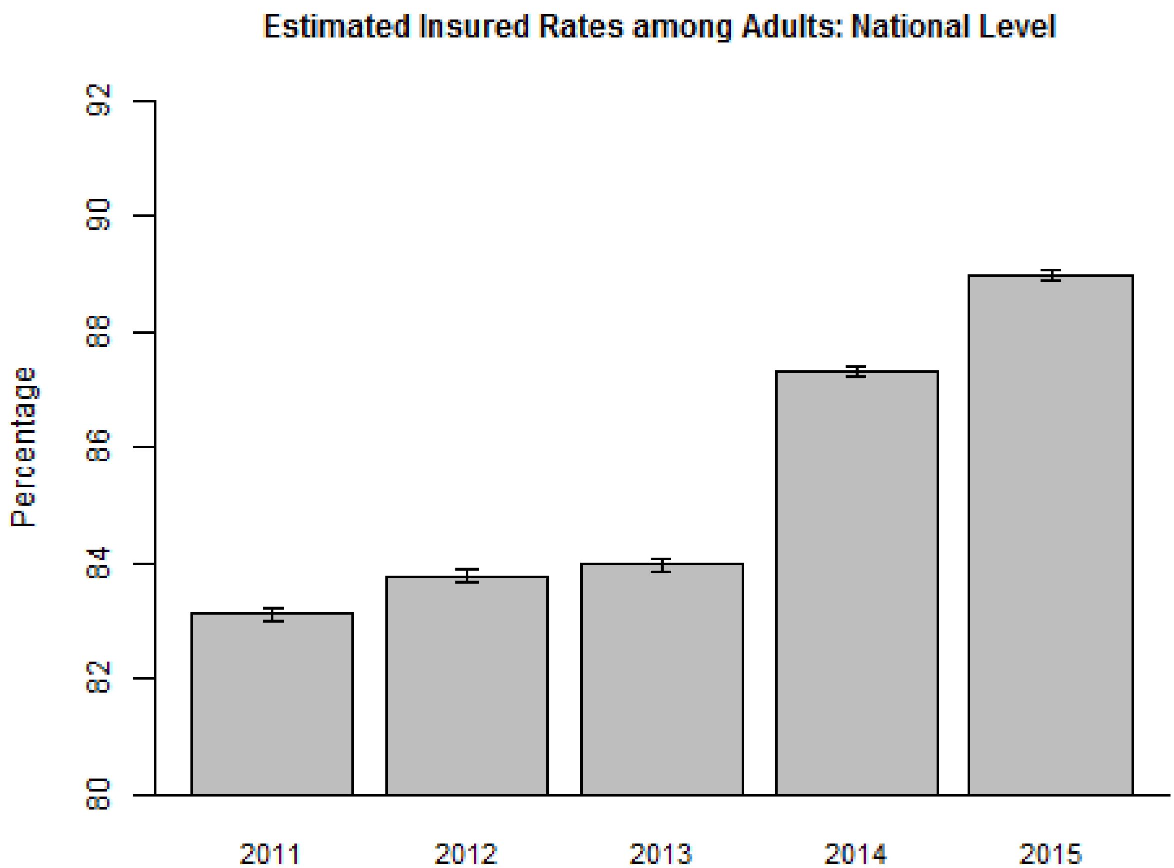

Figure 2: **Figure 2:** The plot above shows the estimated percentage of insured adults from 2011 to 2105. The line segments at the top of each bar represent the estimated 95% confidence intervals for the percentage of insured adults. Note the jump in percentage from 2013 to 2014.

Model Fitting and Diagnostics

To fit our Poisson models, we will use the built in R function for fitting generlized linear models ("glm()"). We will fit one model for each year and then one model for all years combined (five year averages). The model for the five year averages will still be fit using only 53 observations because each state still only counts for 1 observation. We do this because we cannot count different years for the same state as different observations because we cannot reasonably assume that these observations would be independent. Figure 4 shows the fitted log-linear models over scatterplots of the data.

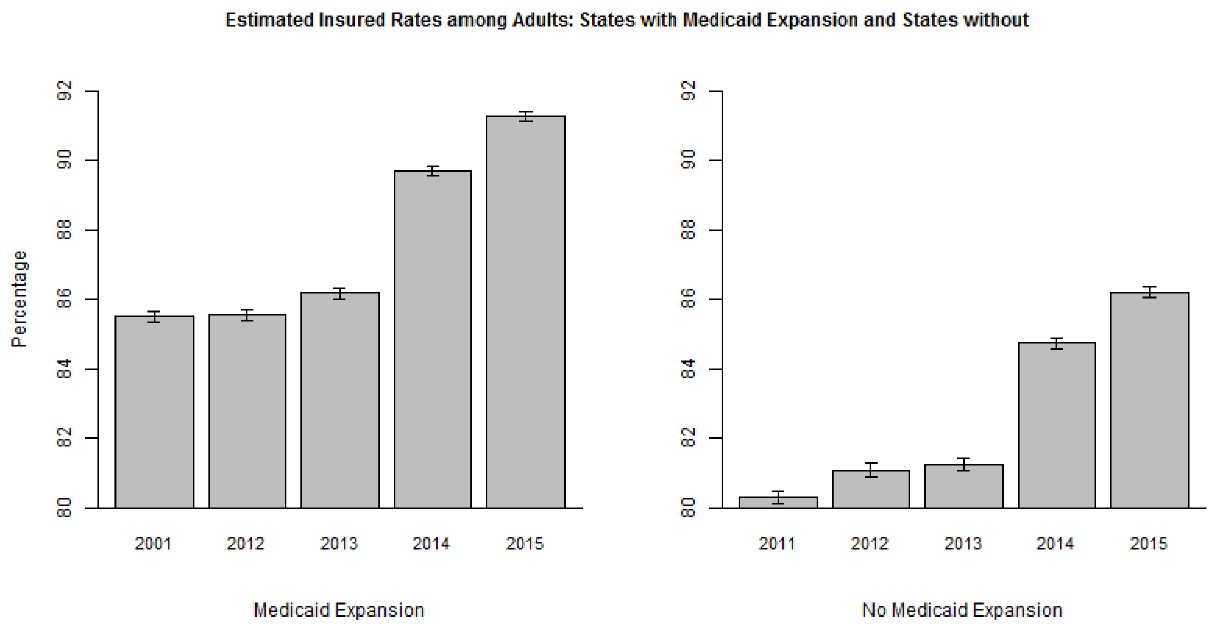

Figure 3: **Figure 3:** The plots above show the estimated percentage of adults with insurance in states that adopted the Medicaid expansion at the beginning of 2014 and those that had not adopted the expansion as of the end of 2015. The line segments at the top of each bar indicate the 95% confidence for the percentage of insured adults for that year. It is clear that in both states with and without the Medicaid expansion, insured rates have increased from 2011 to 2015 and there was a sharp boost in insured rates from 2013 to 2014 that corresponds with the initiation of most of the ACA reforms. Also, it is clear that while insured rates increaed for both Medicaid expansion states and non-Medicaid expansion states, overall the percentage of insured adults in states without the expansion is much lower..

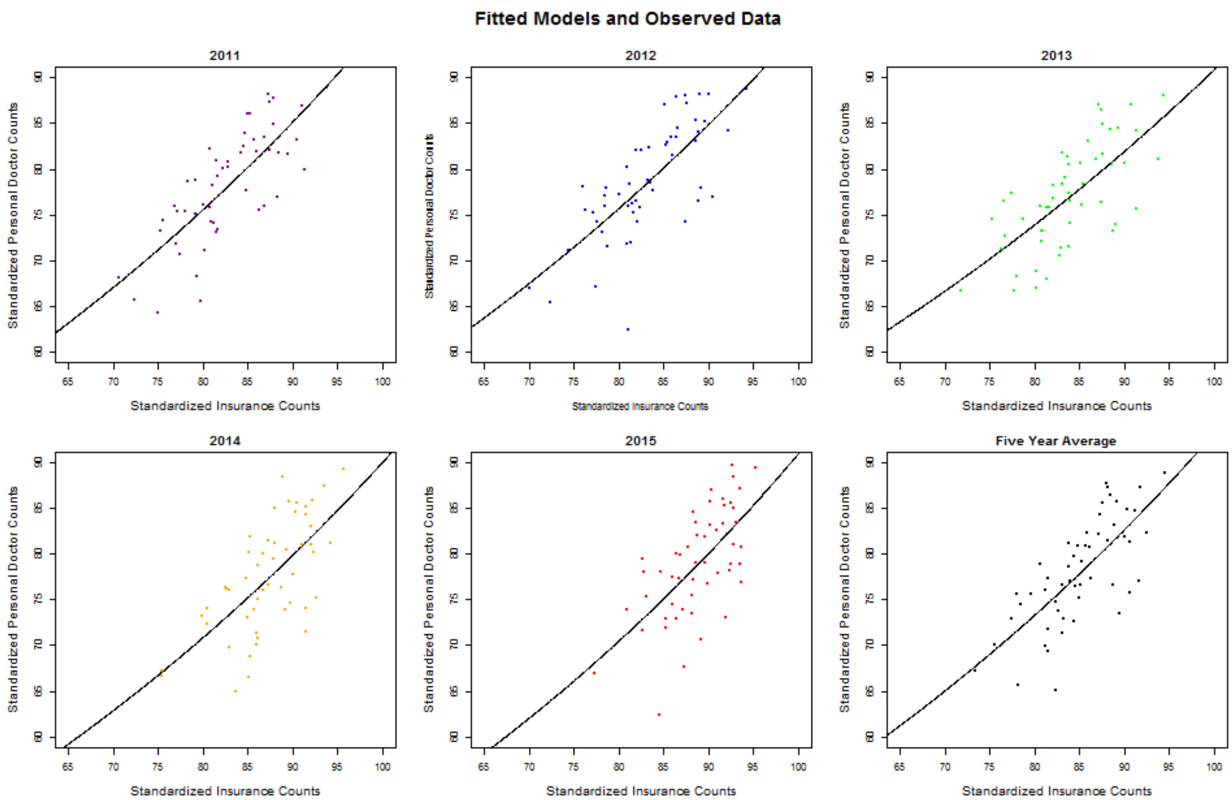

Figure 4: **Figure 4:** Log-linear models of standardized personal doctor counts regressed onto standardized insured rates.

After fitting our models we used the dispersion test ("dispersiontest()") from the Applied Econometrics with R ("AER") package to test if our data was over- or under-dispersed for a Poisson distribution. We found that our data for each model was significantly underdispersed. To correct for this, we refit the model using the "quasipoisson" family option which uses a Poisson distribution with an estimated dispersion parameter. This does not effect the estimates of the regression parameters, but it does effect the estimated standard errors of said parameters. In our case, since our data were under-dispersed, our estimated standard errors decreased. Table 5 below reports the estimated model parameters, their standard errors, and their p-values.

Table 7: **Table 5:** Estimated model parameters, standard errors, and p-values for each fitted model.

		Estimate	Standard Error	P-value
2011	Intercept	3.369	0.1146	1.194e-33
	Slope	0.01196	0.001378	1.295e-11
2012	Intercept	3.41	0.1215	1.133e-32
	Slope	0.01146	0.001454	2.211e-10
2013	Intercept	3.476	0.1274	4.278e-32
	Slope	0.01033	0.001512	9.928e-09
2014	Intercept	3.303	0.1655	9.508e-26
	Slope	0.01198	0.001891	6.127e-08
2015	Intercept	3.234	0.1896	1.013e-22
	Slope	0.01277	0.002128	2.033e-07
Five Year Average	Intercept	3.335	0.1346	4.295e-30
	Slope	0.01201	0.001576	5.774e-10

Residual Analysis

Figure 5 below shows the residuals plotted agains the fitted values for each model. There does not appear to be any pattern to the residual scattering for any of the models.

Global Fit Tests

For global fit tests we are testing the null hypothesis that the model fits the data against the alternative hypothesis that the model does not fit the data (i.e. there is significant lack of fit). We performed two global fit tests for each model, the Pearson test and the likelihood ratio test. All six of our models passed both tests, indicating that there is no significant lack-of-fit.

Pearson Goodness-of-Fit Test

The test-statistic for the Pearson test is:

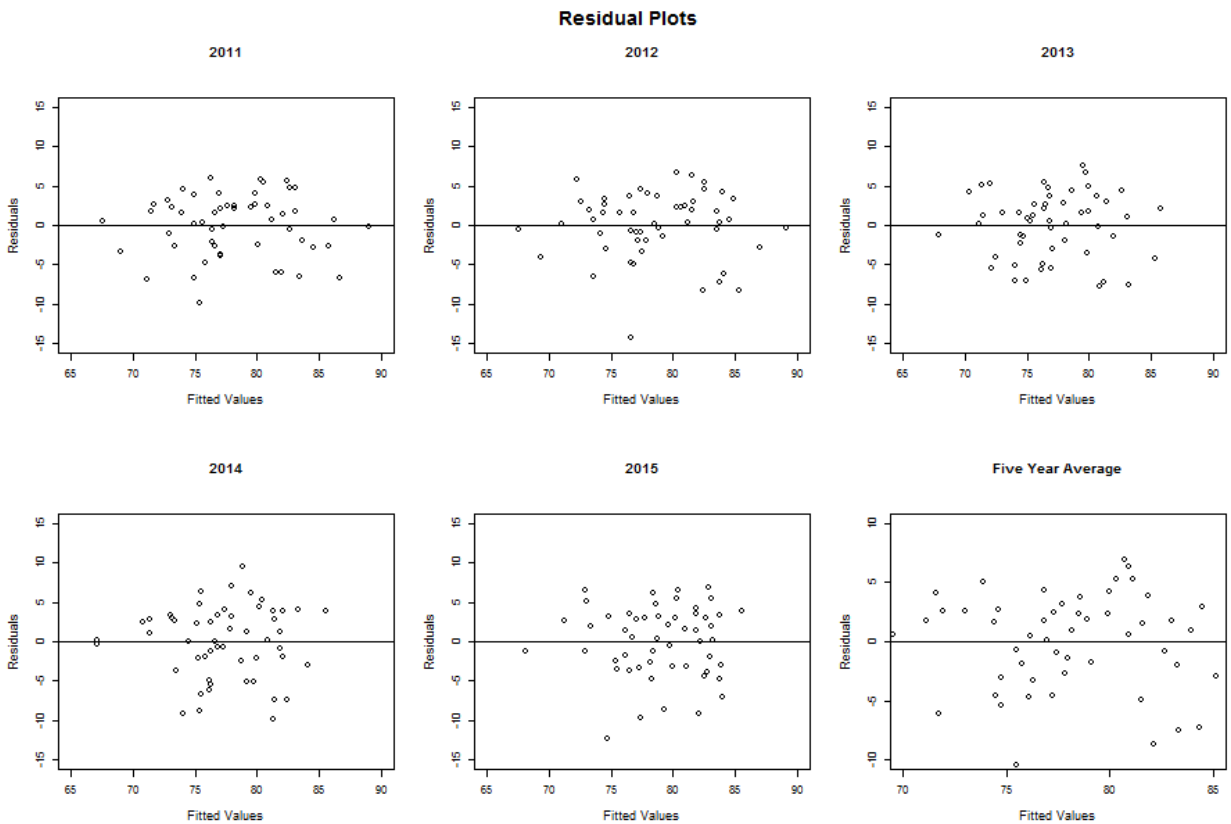

Figure 5: **Figure 5:** The plots above show the residuals plotted against the fitted values for each model.

$$X^2 = \sum_{i=1}^{n} \frac{(y_i - \widehat{\mu}_i)^2}{\widehat{\mu}_i} \sim \chi^2_{n-p} n = \text{number of observations} p = \text{number of model parameters}$$

For all of our model, the number of observations in the data set is 53 and the number of parameters is 2 (slope and intercept coefficients) so the degrees of freedom are 51. We used the following R function to calculate the test-statistic and corresponding p-value for each distribution.

```r
#Function takes fitted values and observed values as arguments

Pearson.Global.Fit <- function(fitval, obsval, df = 51){

  TS <- sum( (obsval - fitval)^2 / fitval )
  p <- pchisq(TS, df, lower.tail = FALSE)

  out <- data.frame(c(TS, p))
  out <- t(out)
  colnames(out) <- c("Test-Statistic", "P-value")

  out
}
```

Table 6 below reports the test-statistics and corresponding p-values for each model. The p-values are all approximately 1 indicating that we should accept the null hypothesis that the models fit the data.

Table 8: **Table 6:** Pearson goodness-of-fit test-statistics and p-values for each model

	Test-Statistic	P-value
2011	9.911	1
2012	11.55	1
2013	11.06	1
2014	13.61	1
2015	13.37	1
Five Year Average	10.58	1

Likelihood Ratio Goodness-of-Fit Test

The test-statistic for the likelihood ratio test is:

$$G^2 = 2 \sum_{i=1}^{n} y_i \ln \frac{y_i}{\widehat{\mu}_i} \sim \chi^2_{n-p} n = \text{number of observations} p = \text{number of model parameters}$$

Again all our models were fitted using data sets of 53 observations and all of our models have 2 parameters so our degrees of freedom are 51. We used the following R function to calculate the likelihood ratio test-statistic and corresponding p-value for each model.

```r
#Function takes fitted values and observed values as arguments

LR.Global.Fit <- function(fitval, obsval, df = 51){

  TS <- abs( 2*sum( (obsval) * log( (obsval)/(fitval) ) ) )
  p <- pchisq(TS, df, lower.tail = FALSE)

  out <- data.frame(c(TS, p))
  out <- t(out)
  colnames(out) <- c("Test-Statistic", "P-value")

  out
}
```

Table 7 reports the calculated likelihood ratio test-statistics and corresponding p-values for each model. The p-values of virtually 1 indicate that, as with the Pearson test, the likelihood ratio test found no evidence of lack of fit for any of our models. Thus, we accept the null hypothesis that our models fit the data.

Table 9: **Table 7:** Likelihood ratio goodness-of-fit test-statistics and p-values for each model

	Test-Statistic	P-value
2011	10.01	1
2012	11.78	1
2013	11.12	1
2014	13.73	1
2015	13.56	1
Five Year Average	10.7	1

Comparative Fit

From looking at the residual plots and the results of the global fit tests we have no reason to think that our models poorly fit the data. Now we want to see if they give us better fit than just the null model (a horizontal line at the global mean of the data set). To do this we will test the difference between deviance of the null model and the fitted model. The null hypothesis of this test is that the fitted model gives no better fit than the null model, the alternative is that it does give significantly better fit. The test statistic is the difference between the deviance of the null model and the deviance of the fitted model. The smaller the difference the less likely it is that the fitted model gives better fit than the null model. The test statistic has a chi-squared distribution with degrees of freedom equal to the difference between the number of parameters in the null model and the fitted model. So for us, the degrees of freedom will equal 1.

$$TS = \text{Deviance}_{null} - \text{Deviance}_{fitted} \sim \chi^2_1$$

We can use the "summary()" function to find the deviance of each of our fitted models and their corresponding null models. Table 8 below reports the test-statistics and p-values for each of our models. For each model the p-value was less than .01. This indicates that even at a 99% percent significance level we should reject the null hypothesis that the fitted model gives no better fit than the null model and conclude that the fitted model does indeed provide better fit.

Table 10: **Table 8:** Comparative fit test-statistics and p-values for each model

	Test-Statistic	P-value
2011	14.66	0.0001289
2012	14.14	0.0001694
2013	10.14	0.001453
2014	10.81	0.001009
2015	9.551	0.001999
Five Year Average	12.1	0.0005035

Diagnostics Using Split Data Sets

Earlier on we mentioned that we split the BRFSS data sets for each year in half so we could use half of a data set to fit a model and the other half to test the model. Now we will use those data sets to test the models. We will begin by performing the last three tests we just performed.

- Pearson global fit test
- Likelihood ratio global fit test
- Deviance comparative fit test

We will also test the predictive power of our models using the concordance statistic gamma.

Global and Comparative Fit Tests

We will use the R function "predict()" to generate fitted values for observations in the diagnostic data sets and then use these fitted values to perform the above tests in the same manner as before. Tables 9 and 10 report the global fit statistics calculated using the test data sets. The p-values are still all approximately 1 indicating that there is still no evidence that our models fit the data poorly.

Table 11: **Table 9:** Pearson goodness-of-fit test-statistics and p-values for each model calculated using the testing half of the data sets.

	Test-Statistic	P-value
2011	10.43	1
2012	12.48	1
2013	12.23	1
2014	13.63	1
2015	10.97	1
Five Year Average	10.92	1

Table 12: **Table 10:** Likelihood ratio goodness-of-fit test-statistics and p-values for each model calculated using the testing half of the data sets.

	Test-Statistic	P-value
2011	9.429	1
2012	9.807	1
2013	23.36	0.9997
2014	7.081	1
2015	8.258	1
Five Year Average	6.908	1

Predictive Power

Next we will measure the predictive power of our models using the concordance statistic Goodman and Kruskal's gamma. To do this we first predict the number of personal doctor counts for each observation in each testing data set using the corresponding model. Then we use the predicted number of personal doctor counts and observed number of personal doctor counts for each observation to calculate the number of concordant and discordant pairs of obervations in the data set. A concordant pair is a pair of observations where one observation has both a high predicted personal doctor count and observed personal doctor count than the other. A discordant pair is when one observation has a higher predicted personal doctor count but lower observed personal doctor count, or vice versa.

The Goodman and Kruskal's gamma statistic is calculated as the difference between the number of concordant and discordant pairs divided by the total number of concordant or discordant pairs:

$$\widehat{\gamma} = \frac{C - D}{C + D}$$

Gamma can range from -1 to 1, with negative values indicating more discordant than concordant pairs and positive values indicating more concordant pairs. In our case, a gamma value greater than 1 indicates a degree of predictive power, the closer to one the better. An approximate standard error for gamma is given by:

$$\widehat{SE}[\gamma] = \sqrt{\frac{n(1-\widehat{\gamma}^2)}{C+D}} \text{where} n = \text{total number of observations}$$

We will use this approximate standard error to calculate approximate 95% confidence intervals for gamma. We created the R function below to calculate gamma and it's confidence interval for each model. Table 11 reports these figures. The confidence intervals are relatively large but none of them includes zero, indicating that our models do have a degree of predictive power.

```
# Function takes observed values, fitted values, and desired alpha level as arguments

GK.gamma <- function(obs, fit, alpha = 0.05){

  z <- qnorm(alpha/2, lower.tail = F)
  n <- length(obs)
  data <- rbind(obs, fit)

  C <- 0
  D <- 0
```

```r
  for(i in 1:(n-1)){
    o <- data[1,i]
    f <- data[2,i]
    for(j in (i+1):n){
      if(o < data[1,j] & f < data[2,j]){
        C <- C + 1
      }
      else if(o > data[1,j] & f > data[2,j]){
        C <- C + 1
      }
      else if(o > data[1,j] & f < data[2,j]){
        D <- D + 1
      }
      else if(o < data[1,j] & f > data[2,j]){
        D <- D + 1
      }
    }
  }

  gama <- (C - D) / (C + D)

  SE.gama <- sqrt( n*(1-gama^2)/(C+D) )

  G.LB <- gama - z*SE.gama
  G.UB <- gama + z*SE.gama

  out <- c(G.LB, gama, G.UB)

  out
}
```

Table 13: **Table 11:** Estimated gamma values and 95% confidence intervals for each model.

	Lower Bound	Point Estimate	Upper Bound
2011	0.2455	0.5631	0.8808
2012	0.2364	0.5559	0.8754
2013	0.1787	0.5094	0.8402
2014	0.1328	0.4717	0.8106
2015	0.1415	0.479	0.8164
Five Year Average	0.22	0.5428	0.8656

Model Interpretation

Now that we have verified that our models satisfactorily fit the data, let us look more closely at our models. Figure 5 shows a plot including each model so they can be seen together for comparison. While the slope of each model appears to similar, as time goes on the models appear to shift to the right, indicating that the intecept changes.

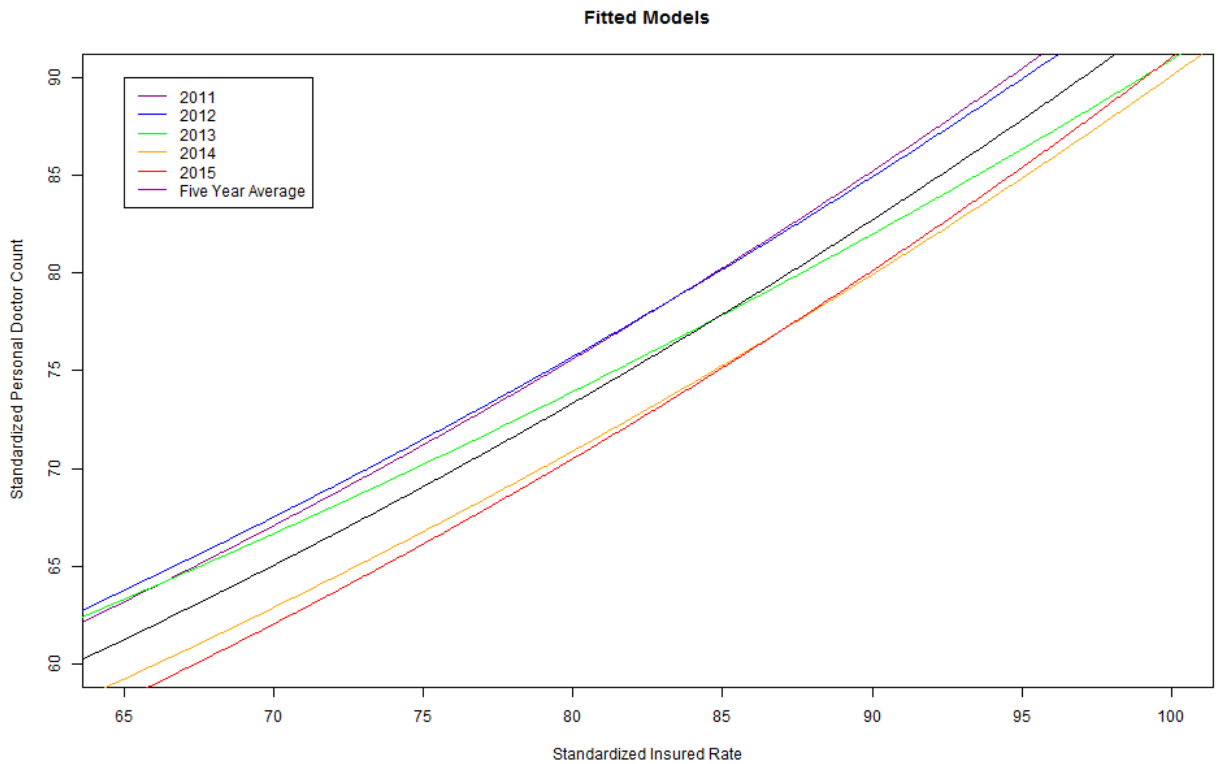

Figure 6: **Figure 5:** Fitted log-linear models. The slopes of each model appear relatively similar but as time goes on the models appear to shift to the right, indicating that their intercepts differ.

Figure 6 shows the 95% confidence intervals for the estimated intercept and slope for each model. Although the apparent shift seen in the log-linear models in figure 5 seems to indicate that the intercepts for the models differ, we can see from figure 6 that the confidence intervals for each intercept includes the point estimates for the intercepts of the other model. This negates the possibility that the intercepts differ significantly. Additionally, as expected from looking at figure 5, the slopes do not differ significantly either.

Table 12 reports the exponentiated estimated slope and the corresponding 95% confidence interval for each model. The exponentiated slope represents the factor by which the predicted mean percentage of individuals with a personal doctor should increase by for a one percentage point increase in the insured rate. For example, the for 2015 estimate indicates that for a 1 percentage increase in the number of insured individuals we should expect the mean of the percentage of the number of individuals with a personal doctor to increase by a factor of 1.013. Or, interpreting the confidence interval, for a 1 percentage point increase in the number of individuals with insurance at the state level, we are 95% confident that the mean percentage of adaults with insurance should increase by a factor in the range of 1.009 to 1.017. While these may seem like small changes, on the population level they can be quite significant.

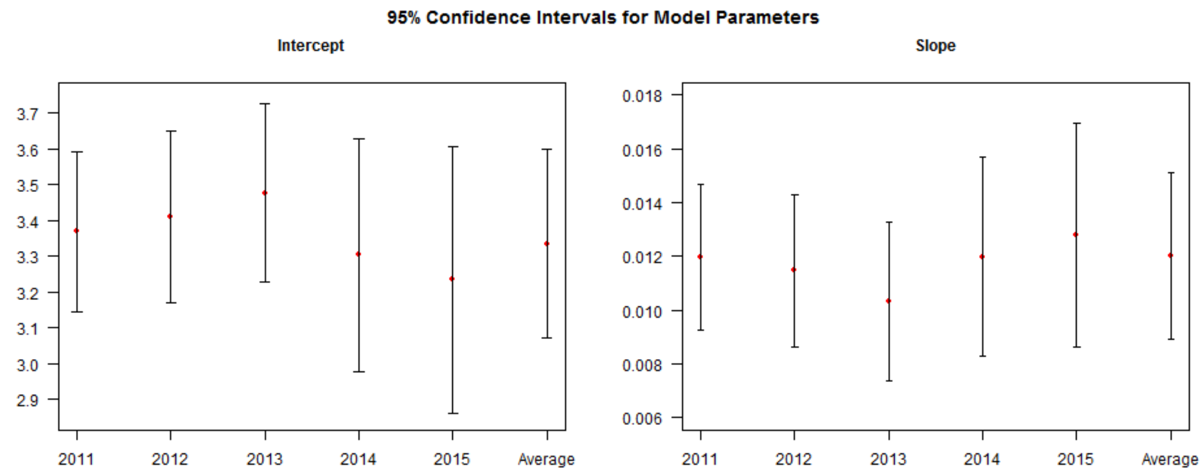

Figure 7: **Figure 6:** 95% confidence intervals for the model parameters. Since the confidence intervals for each intercept include the point estimates for every other intercept we can conclude that the intercepts do not differ significantly between models. The same is true for the slope parameters.

Table 14: **Table 12:** 95% confidence intervals for the exponentiated slope parameter of each fitted model.

	Lower Bound	Point Estimate	Upper Bound
2011	1.009	1.012	1.015
2012	1.009	1.012	1.014
2013	1.007	1.01	1.013
2014	1.008	1.012	1.016
2015	1.009	1.013	1.017
Five Year Average	1.009	1.012	1.015

Discussion

Our analysis shows that the number of individuals with a personal doctor is indeed correlated with the number of individuals with the number of individuals with insurance at the state level. The nature of this relation did not change significantly from 2011 through 2015 even though insured rates did change significantly. Thus it is reasonable to expect that this trend will continue in the future. Looking to the model which used data from all five years, our model predicts with 95% confidence that for a 1 percentage point increase in the number of insured individuals at the state level the mean percentage of individuals with a personal doctor should rise somewhere between 1.009 and 1.015 fold. This may seem like a small increase, but when viewed at the state population level it can be quite significant. Even for a small state such as Maine, a 1.009 fold increase in the percentage of individuals with a personal doctor corresponds to tens of thousands of people. When considering a large state such as California, it corresponds to hundreds of thousands of more people having a personal doctor.

This trend is strong supportive evidence of the relationship between primary care access and the insured rate in a state. States that have not opted for the Medicaid expansions would do well to reconsider there position. Overall, insured rates in states which adopted the Medicaid expansions are about 5% than those which did not (see tables 3 and 4 in the insured rates section of the methods and results). If our models are correct, if those states who have not already opted into the Medicaid expansion finally did so, there should be a corresponding increase in the number of individuals with a personal as the insured rate increases.

Conclusion

NEW ENGLAND JOURNAL OF MEDICINE {2014};**371**:275–81.

3. Barnett J, Vornovitsky M. Health Insurance Coverage in the United States: 2015. *United States Census Bureau* {2016};**P60-257**.

4. CDC. Behavioral risk factor surveillance system annual survey. 2011–2015.