Appendix A
Estimation Methods

As has been documented throughout the report, available evidence on the prevalence of CKD in NC is limited, primarily because it is underdiagnosed, awareness is low, and it is difficult to obtain serum creatinine measures on a generalizable sample of the NC population. In order to inform the Task Force on the scope and magnitude of the problem, and to help develop cost estimates for some of the recommendations, North Carolina specific prevalence estimates were developed. The general method for developing these estimates is outlined below.

Step 1: Classify NHANES respondents by CKD Stage

National Health and Nutrition Examination Survey data from the 2001-2002 and 2003-2004 interview periods were used to generate national CKD prevalence estimates. The approach follows those used in published analyses.[1,2] Estimated GFR was calculated using the MDRD equation. Also following that work, persistent albuminuria was operationalized by assuming specified fractions of the microalbunuria cases were persistent (50.9% for those with microalbuminuria and an eGFR>90 ml/min/1.73 m^2; 75.0% for those with microalbuminuria and an eGFR of 60 to 89 ml/min/1.73 m^2; 100% of all macroalbuminuria cases).[2] This was accomplished by randomly selecting 50.9% of cases in which eGFR was greater than 90 ml/min/1.73 m^2 and microalbuminuria was present and labeling these cases "Stage 1." The remaining 49.1% of eGFR>90 ml/min/1.73 m^2 and microalbuminuria cases were not classified as having CKD. This process was repeated for the eGFR between 60 and 89 ml/min/1.73m^2 and microalbuminuria cases. All macroalbuminuria cases were classified as CKD with stage depending on the eGFR.

Prevalence estimates by stage are presented below in Table A-1. Note that Coresh et al ignore NHANES respondents classified as Stage 5 due to an eGFR<15. The identical approach is taken here.

Step 2: Develop prediction model for CKD Stage

With the Stage assignment in hand, prediction models were performed to ascertain the association between CKD Stage and commonly observed characteristics. The final

prediction model included age, gender, race (African-American non-Hispanic or otherwise), self-reported diabetes status, and self-reported hypertension status.

Multiple models were considered, including one logistic model predicting CKD Stage 1-4, another predicting CKD stage 3-4, and an ordered probit predicting stage of CKD. All models used the NHANES survey weights.

After model estimation, these parameter estimates were set aside.

Step 3: Assign CKD probabilities to BRFSS respondents

The Centers for Disease Control and Prevention annually performs a survey of adults that includes questions on factors, such as high blood pressure and diabetes, known to be associated with CKD. This survey, the Behavioral Risk Factor Surveillance System (BRFSS) is designed to be representative at the state level; North Carolina has roughly 17,000 respondents to the survey in any given year. Thus, we can apply the national prevalence estimates for CKD from NHANES to the NC population based on state specific health characteristics. For the final model used here, age, gender, history of hypertension, history of diabetes, and race of African-American were used as predictors.

The data from the 2005 BRFSS were formatted to correspond with the NHANES data. (High blood pressure awareness was not asked in the 2006 data, so 2005 was the most recent data available). With the data formatted in the same manner, predictions can be generated using the average associations between observed factors and CKD stage at a national level. A simplified example may be helpful in illustrating this approach. For purposes of the example, assume that a survey reveals that 20 percent of males and 60 percent of females have a certain characteristic. If a group of similar people is 50 percent male and 50 percent female, one estimate is that .5 x 20% + .5 x 60% = 40% of the group has the characteristic. If a second group is 80 percent male and 20 percent female, an estimate would be .8 x 20% + .2 x 60% = 28%. Assuming that the relationships in the development (here, NHANES) dataset is similar to the relationships in the *estimation* (here, NC using BRFSS) population, then we are able to estimate a valid prevalence rate of CKD in NC using this approach.[3]

Table A-1 presents 5 different sets of estimates. Column A is the prevalence as presented in Coresh using 1999-2004 data. Column B is the NC IOM's analysis based on 1999-2004. Both these are standardized to the 2000 standard population. Overall, the predictions are quite similar between the Coresh et al model analysis and the replication by the Task Force. Column C uses the weights in the NHANES data, meaning the predictions apply to a 1999-2004 population. This slightly lowers the estimated prevalence of CKD. Column D applies the estimated relationship in NHANES and looks at the in-sample predictive power. The in-sample predictive power, as expected, is overall quite good – the predicted prevalence is very similar to the estimated

prevalence. Column E applies the prediction model to the national BRFSS data (out-of-sample prediction). Applying the model to the BRFSS data increases the prevalence; one reason is the much higher prevalence of diabetes (7.7 in BRFSS vs. 6.8 in NHANES). Finally, Column F limits the prediction to only the North Carolina BRFSS data. Note the highly similar prevalence estimates for the US and North Carolina. Although North Carolinians have a higher estimated prevalence of CKD due to higher rates of diabetes and hypertension, our age profile is younger and this tends to lower the estimated prevalence.

Table A.1.
Comparison of Prevalence Estimates

Column	A	B	C	D	E	F
Area:	*National*	*National*	*National*	*National*	*National*	*North Carolina*
Population	*2000*	*2000*	*1999-2004*	*1999-2004*	*2005*	*2005*
Analysts:	*Coresh et al*	*NC IOM*	*NC IOM*	*NC IOM*	*NC IOM*	*NC IOM*
Data:	*NHANES*	*NHANES*	*NHANES*	*NHANES*	*BRFSS*	*BRFSS*
Type:	*Actual*	*Actual*	*Actual*	*Predicted*	*Predicted*	*Predicted*
No CKD	86.94%	86.68%	87.62%	87.48%	86.40%	86.32%
Stage 1	1.78%	1.64%	1.68%	1.69%	1.76%	1.77%
Stage 2	3.24%	3.16%	3.05%	3.18%	3.34%	3.36%
Stage 3	7.69%	8.13%	7.31%	7.30%	8.07%	8.11%
Stage 4	0.35%	0.38%	0.34%	0.35%	0.42%	0.43%

Step 4: Estimate prevalence for subcategories of the population

With these prevalence estimations in hand, other estimates can be derived in a relatively straightforward manner. The prevalence estimation yields a probability of the respondent having each of the 5 CKD stages (0-4) considered here. BRFSS asks other questions such as insurance status and whether the respondent has a usual source of care. Using the answers to these questions, the number of CKD patients who are uninsured, or who are uninsured and do not have a usual source of care, or the number of people with CKD or at high risk for CKD who are uninsured and have no usual source of care, can be estimated. The estimated number of people with CKD is estimated by summing the individual probabilities. Likewise, other estimates – such as the number of people with diabetes without CKD – are computed similarly by summing probabilities for individual respondents.

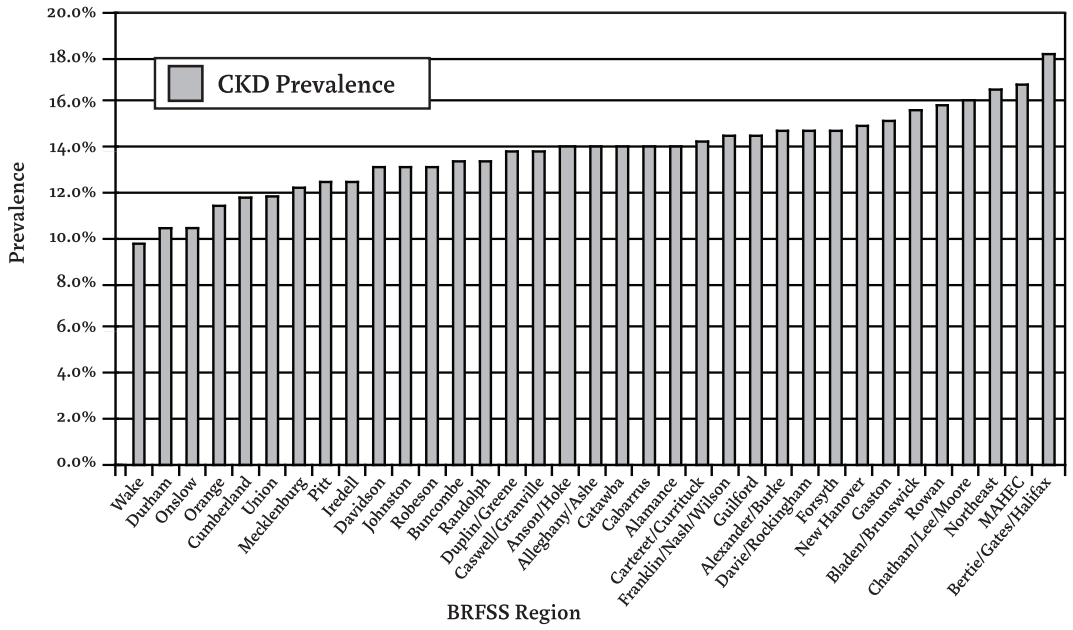
Step 5: Compute estimated costs

With subpopulation estimates in hand, estimated statewide costs can be derived by including per person costs from other sources. For example, for the Recommendation on purchase of medical care for nephrologists, multiplying the number of North Carolinians who are (a) uninsured (b) have no usual source of care and (c) have CKD by the average cost of a nephrologist visit yields the recommended appropriation.

Table A.2.
Estimated CKD Prevalence by Cardiovascular Disease (CVD), Diabetes, & Hypertension (HTN) Status

	Without CVD			With CVD			Total		
	With Diabetes	*Without Diabetes*	*Total*	*With Diabetes*	*Without Diabetes*	*Total*	*With Diabetes*	*Without Diabetes*	*Total*
Estimated North Carolinian Adults with CKD Stage 3-4									
No HTN	145,713	20,857	166,570	17,183	8,314	25,497	162,896	29,171	192,067
HTN	168,312	73,229	241,540	55,253	40,521	95,774	223,565	113,750	337,315
Total	314,025	94,086	408,110	72,436	48,835	121,271	386,461	142,921	529,382
Total North Carolinian Adults									
No HTN	4,015,397	135,574	4,150,971	157,160	31,652	188,812	4,172,557	167,226	4,339,783
HTN	1,231,064	257,322	1,488,386	246,155	122,782	368,937	1,477,219	380,104	1,857,323
Total	5,246,461	392,896	5,639,357	403,315	154,434	557,749	5,649,776	547,330	6,197,106
Percent with CKD Stage 3-4									
No HTN	3.6%	15.4%	4.0%	10.9%%	26.3%	13.5%	3.9%	17.4%	4.4%
HTN	13.7%	28.5%	16.2%	22.4%	33.0%	26.0%	15.1%	29.9%	18.2%
Total	6.0%	23.9%	7.2%	18.0%	31.6%	21.7%	6.8%	26.1%	8.5%

Chart A.3.
Estimated CKD Prevalence



REFERENCES

1. Coresh J, Byrd-Holt D, Astor BC, et al. Chronic kidney disease awareness, prevalence, and trends among US adults, 1999 to 2000. *J Am Soc Nephrol*. 2005;16:180-188.
2. Coresh J, Selvin E, Stevens LA, et al. Prevalence of chronic kidney disease in the United States. *JAMA*. 2007;298:2038-2047.
3. Rao JNK. Small Area Estimation. Hoboken, N.J.: John Wiley; 2003.