Technical Documentation for Licensure and Workforce Survey Data Analysis

April 2010

Technical Documentation for Licensure and Workforce Survey Data Analysis

Background

The Florida Center for Nursing (Center) was established in 2001 to address issues related to nurse supply, demand, and shortage in this state. The nurse licensure database maintained by the Florida Board of Nursing (FBON) is one important source of information on the state's nurse supply. The licensure database contains the most complete information available in the state specific to the regulation of nurses. Included is information on the number of licensed nurses, their eligibility to practice, their demographic characteristics, and their distribution across the state of Florida.

Licensure data collected by the FBON do not include information about the work behavior of nurses, and this limits their usefulness for strategic labor force planning. The data do not indicate whether nurses are working, whether they work in or outside the field of nursing, how much they work, or in what setting. Because the Center is primarily interested in the amount of nursing *labor* provided in Florida, in contrast to the number of Florida nursing *licenses* that are held, licensure data are cleaned and subset to isolate nurses who could reasonably be practicing nursing within the state of Florida. We call this subset the *potential nurse workforce*.

Beginning in 2008, the Center worked with the Florida Board of Nursing and Medical Quality Assurance to integrate a voluntary Workforce Survey into the online renewal process for nurses. The Workforce Survey generates important data for workforce analysis, such as work status, hours worked, and highest degree held. Using license number as a unique identifier, workforce survey data are merged with licensure data so that members of the *potential* nurse workforce can be counted as *actually working* in nursing, if they indicate that they are.

The Workforce Survey reached more than 92 percent of renewing nurses in 2008 and 2009, but a number of cases still lack workforce data. In addition to those who did not complete the survey during renewal, nurses newly licensed in Florida during 2008 and 2009 will not be exposed to the Workforce Survey until they renew their licenses for the first time. These groups comprise about 16.5 percent of the potential nurse workforce. The Center uses information that we have about these non-respondents (such as their age and gender) to estimate whether and how much they work in the field of nursing. This document provides technical details about the process of merging, cleaning, and estimating values for some cases (called *imputing* values) using licensure and Workforce Survey data.

Data Extract

In 2006, an agreement between the Center and FBON was reached whereby licensure data is provided to the Center regularly as a data extract (a static file) drawn from the dynamically changing licensure database. The extract includes records for each nursing license held in Florida by Registered Nurses (RNs), Advanced Registered Nurse Practitioners (ARNPs), Clinical Nurse Specialists (CNSs), and Licensed Practical Nurses (LPNs). Each record contains information on license type (RN, ARNP, CNS, or LPN), license status (e.g., active, suspended),

date of birth, gender, race/ethnicity, year of initial licensure in Florida, and address. Beginning in 2007, application type (examination or endorsement) was incorporated into the extract.

The Center uses an extract drawn in late December of each year to represent the population of licensees as of January in the following year. In addition to being an intuitive measurement time since it is the start of a calendar year, this time point is also the beginning of license renewal cycles. Florida nurses renew their licenses every two years. In odd years, approximately one-third of RNs and ARNPs renew from January through April and all LPNs renew from March through July. In even years, the remaining two-thirds of RNs and ARNPs renew from January through July. Analysis of the late December extract has two advantages. First, it avoids the rapid changes to the licensure database that occur during renewal cycles. Second, it gives nurses who missed their renewal deadlines from the previous year an additional five or seven months to complete renewal.

The Workforce Survey implemented in January 2008 is housed separately from the licensure database, and an extract of Workforce Survey data is provided to the Center regularly along with the licensure data extract. The questionnaire used is presented in Appendix A. Because nurses renew biennially, it takes two years of renewals to amass complete survey data from all renewing nurses choosing to take the survey. The 2008-2009 Workforce Survey closed on December 31, 2009, completing the first two-year cycle, and a new set of survey data will be collected as nurses renew in 2010 and 2011.

The licensure data extract representing January 2010 was merged with Workforce Survey data collected during 2008 and 2009 using nursing license numbers as the unique identifier for joining records. Though survey data are collected over a two-year period, we treat survey data as representing the best possible estimate of a nurse's 2010 work status. Each license in the January 2010 extract, regardless of participation in the survey, is assigned an estimated work status and full-time equivalent (FTE) value. Nurses who completed a survey but are no longer in the licensure database as of January 2010 (N=435) are excluded from the combined dataset. Thus, the merged dataset contains the best possible estimate of the workforce as of January 2010.

Data Cleaning

Licensure data are first cleaned for implausible dates of birth and initial licensure. The data contain some records with clearly inaccurate birth years dating back to the 19th century. Nurses with birth years earlier than 1913 or with missing birth years are coded as missing for age. The upper limit for age is 97 using this procedure. Similarly, some records contain birth years that would render the nurse implausibly young. For RNs, age is coded as missing if the birth year dates them at less than 20 years of age. For LPNs, the lower limit for age is set to 18 years of age. These lower limits were selected based on the earliest typical age at graduation from nursing programs as well as a sharp drop in the number of records with younger ages.

Dates of initial licensure are also inspected for implausible dates and the proportion of records containing each date. Nurses recorded as being licensed before 1939, a date indicating the nurse has been continuously licensed in Florida for 71 years, are coded as missing for this field. An important note regarding this variable is that it may be "reset" if nurses allow their licenses to

expire but later become licensed again in the state. Since it is not possible to tell whether this has happened to nurses in the licensure extract we receive, the measure must be interpreted with caution. Average tenure as a nurse in Florida may be underestimated by these data.

Missing data generated by these cleaning procedures are typically minimal. Less than .5 percent of RN and ARNP cases were recoded to missing on the age variable in 2010, as were less than 1 percent of LPN cases. Roughly the same numbers of cases were recoded to missing on original license year. Naturally occurring missing data also exist for other variables in the licensure database. About 1-3 percent of records are typically missing data on gender and 4-6 percent are typically missing data on race and ethnicity. In analyses reported by the Center, percentages are based on cases with non-missing values for a variable.

Nurse Placement in Counties and Regions

Stakeholders depend on local data for nurse workforce planning, so the Center provides as much detail as possible at regional and county levels. We use both Workforce Survey data and address information from the licensure database to identify nurses who report living and/or working in the state of Florida. When present, Workforce Survey data is always the source of information used. Nurses completing the survey identify the Florida county in which they work or, if they are not working, the Florida county in which they live. Those working and living in other states select "Outside of Florida." We judge survey data to be superior to address information from the licensure database because it is unknown how recently the address fields in the licensure database have been updated.

The licensure database contains two sets of address fields used to place nurses who did not complete the Workforce Survey: a mailing address (where nurses wish to receive mail from the FBON) and a practice address. Analysis of the two fields indicates that more than half of nurses have the same zip code listed for both mailing and practice addresses. Discussion with FBON and Medical Quality Assurance staff revealed that "mailing" address may be a home residence or a work location – wherever the nurse wishes to receive mail from the FBON. In addition, mailing address may be substituted for practice address if the latter is left blank. In this case, the addresses may be identical either because the nurse receives mail at work *or* because the nurse is not practicing. As a result, the data do not allow a clean analysis of where nurses live and work, although they do allow identification of nurses who report working or receiving mail outside of Florida.

Nurses with Florida addresses are placed into counties, Regional Workforce Boards, and larger regions of the state consisting of multiple workforce boards. There are 67 counties and 24 Regional Workforce Boards, a classification used by Workforce Florida, Incorporated for workforce planning and outreach efforts. See Appendix B for a map of Florida showing the regions into which each county falls. As noted previously, precision is difficult to obtain for nurses not completing a Workforce Survey since many nurses do not have unique practice location information and it is unknown whether their mailing address represents a residence or work location. Since many nurses are likely to commute from one county to another in order to work, county placement is probably least accurate for describing the distribution of the nursing

workforce. More accuracy is probable in larger geographic areas since the areas are more likely to encompass both work and residence locations.

When nurses not completing a Workforce Survey give unique practice location information, they are placed according to the county in which they are employed. In all other cases, nurses are placed according to the county in which their mailing address – which may be home or work – is located. This procedure maximizes the accuracy of placement for describing workforce distribution as much as can be expected given the data limitations.

To investigate data quality and salvage missing data on county placement, each licensee's recorded zip code is compared against a SAS statistical software lookup table of zip codes which matches the extract date as closely as possible.[1] In the vast majority of cases, nurses with Florida placements give valid zip codes within the state of Florida. In some cases, however, nurses have a valid county placement via Workforce Survey or licensure data when the zip codes they give are invalid. The decision was made to use the pre-existing county information from survey or licensure data in all cases where it is given. The zip code lookup table is used to assign county and regional placement in cases where valid zip codes are available but pre-existing county placements are not. The process typically creates a county placement for several hundred Florida addresses which were previously unplaced.

Identifying the Potential Nurse Workforce

A three-step process is used to generate a subset from the total file representing the potential nurse workforce: those eligible to work as nurses and providing a Florida address. In the first step, nurses with an inactive license are dropped. In the second, nurses with status codes rendering them ineligible to practice, such as Delinquent, Suspended, or Retired, are dropped. In the third step, nurses who have been placed out of state are dropped. Table 1 details the number of licensees excluded in each step of the subsetting process for the years of licensure data presented in our analysis.

Table 1. Licensees Excluded From Analysis, By Year and Reason

Analysis Year	Inactive License	Ineligibility	Lives and/or Works Outside FL	Total Exclusions
2008	10,680	22,044	34,138	**66,862**
2010	5,941	27,630	31,678	**65,249**

Reasons for exclusion are roughly split between license status and non-Florida address in both years. In 2010, however, more were excluded due ineligibility and fewer reported an inactive license. Table 2 shows the number retained and excluded by license type for the 2010 analysis.

---

[1] Zip codes change over time but are reasonably steady in the short term. SAS makes new zip code lookup tables available quarterly. Zip code tables can be downloaded from
http://support.sas.com/rnd/datavisualization/mapsonline/html/misc.html

Table 2. Membership in the Potential Nurse Workforce, January 2010

	All Florida Licensees	Potential Nurse Workforce	Excluded
RN	238,936	187,093	51,843
ARNP	15,391	12,603	2,788
LPN	68,472	57,861	10,611
CNS	60	53	7
Total Licensed Nurses	**322,859**	**257,610**	**65,249**

Nurses who are excluded from the potential nurse workforce tend to be slightly older (one or two years, on average) and more often of White race. In general, excluded nurses are similar to those who are analyzed with the exception that they did not meet inclusion criteria. Table 3 shows the characteristics of licensees excluded from the 2010 analysis.

Table 3. Characteristics of Excluded Licensees, January 2010

	Percent of Licensees		Percent of Licensees
Race/Ethnicity	**%**	**Age**	**%**
White	77.7	21-30	6.7
Black	10.3	31-40	17.3
Hispanic	5.6	41-50	21.3
Asian	4.9	51-60	27.7
Native American	0.3	61 or older	27.0
Other	1.3	**Average Age**	**51.6**
Gender	**%**		
Female	89.8		
Male	10.2		

An additional data cleaning and analytic step was added in 2009 to identify nurses who have both an LPN and an RN record in the licensure database. While license upgrading to the ARNP is detectable because nurses retain their license numbers, LPNs who obtain an RN license receive new license numbers and are more difficult to track via licensure data. Some of the LPNs who have upgraded their license to RN have overlapping license expiration dates, whereby the LPN license expires after the RN license begins. This results in a person having two current nursing licenses and two "near duplicate" FBON records, one for the LPN license and one for the RN license. These "near duplicate" records can be identified through a combination of birth date and social security number. When this situation occurred, we assumed the nurse would be practicing as an RN and removed the LPN license from further analysis. In 2010, we found 3,381 "near duplicate" LPN records. If the duplicate record was generated between 2008 and 2010, we counted them as losses to the potential LPN workforce during this time frame. In the small number of cases where nurses are renewing both LPN and RN licenses year after year, we simply removed the LPN record from further analysis.

Survey Response Rates and Bias Analysis

The online Workforce Survey was available for completion between January 1, 2008 and December 31, 2009. Nurses who renewed their licenses during this time encountered the survey as part of the online renewal process. Though the survey was not included as part of the paper renewal process, the vast majority of nurses have renewed online since 2008, when the Board of Nursing began mailing postcard reminders instead of paper forms for license renewal. Paper renewal forms must now be specially requested.

Table 4. 2008-2009 Workforce Survey Response Rates

Renewal Bucket	# Renewed	# Completed Survey	Response Rate (%)
January-March 2008	50,344	45,378	90.4%
March-July 2008	70,883	64,934	91.6%
January-March 2009	84,899	81,004	95.4%
March-July 2009	57,949	53,890	93.0%
License Type	**# Renewed**	**# Completed Survey**	**Response Rate (%)**
RN	206,126	191,316	92.8%
ARNP	13,947	13,232	94.9%
LPN	57,949	53,890	93.0%
Overall Response Rate	**264,075**	**245,206**	**92.8%**

Table 4 shows the response rates obtained for the 2008-2009 Workforce Survey by renewal "bucket" (the period during which the nurse renewed) and license type. Overall, the survey achieved an impressive 92.8 percent response rate. That rate did not vary much by renewal bucket or license type. The highest response rates were observed for those renewing January-March 2009 (an RN/ARNP cohort) and for ARNPs across all renewal buckets.

Response rates this high suggest that we can feel very confident about the ability of responders to represent all renewing licensees (generalizability). In addition to survey non-respondents, however, nurses newly licensed in Florida during the survey would not have renewed their licenses during the survey period and have a low probability of exposure to the survey. A few newly licensed nurses encountered the survey before their first renewal when logging on to check their license status or provide updated contact information (N=173), but the vast majority would not have been aware of the survey. We retained and analyzed the survey information provided by all members of the potential workforce whether it was completed during renewal or at another time. Whenever a nurse completes the survey during the two year period, it becomes inaccessible to them for the duration of the two years.

To examine patterns of potential bias, we compared survey respondents with those missing survey data due to non-response or new licensure in Florida (Table 5). Those missing data are less likely to be white and non-Hispanic, they are about four years younger, and they have been licensed in Florida for five fewer years when compared with survey respondents. The tendency towards younger, more diverse nurses is consistent with the profile of new licensees. It also means that our survey data under-represent new licensees in ways that are likely important to the estimation of workforce size and characteristics.

Table 5. Comparison of Survey Respondents and Those Missing Data

	Survey Respondents	Missing Data
Race/Ethnicity	**%**	**%**
White	70.6	61.8
Black	14.5	18.7
Hispanic	7.7	11.6
Asian	5.6	5.6
Native American	0.2	0.2
Other	1.4	2.2
Gender	**%**	**%**
Female	90.5	89.4
Male	9.5	10.6
Age	**%**	**%**
21-30	6.7	24.2
31-40	19.0	20.1
41-50	25.8	17.4
51-60	30.5	17.6
61 or older	18.0	20.4
Average Age	**49.2**	**45.2**
Years Licensed in FL	**14.9**	**9.4**

Estimation of the Actual Nurse Workforce

An analysis of survey data was undertaken to discover what, if any, characteristics were associated with nurses' workforce participation and number of hours worked. Of the characteristics available for all licensees, regardless of survey participation, license type (RN/ARNP vs. LPN), age, and gender were most strongly related to workforce participation and number of hours worked. In general, older nurses are less likely to work in nursing, a finding that is particularly true for women. Men had higher participation rates for all ages with the exception of very young nurses. Finally, RNs and ARNPs had higher participation rates than did LPNs.

We used workforce participation rates for survey respondents to select some of the cases missing data for inclusion in our count of the nursing workforce. We first calculated workforce participation rates among survey respondents for each of 20 categories representing *age by gender by license type*. Next, we randomly selected missing cases in each of the 20 categories for assignment to the group "working in nursing." The proportion of missing cases assigned to the group depended on the proportion of survey respondents in that category who reported working in nursing. For an individual case the result may be inaccurate, but in the aggregate analysis an appropriate number of nurses in each age, gender, and license type are counted as members of the workforce. Since selection of nurses for inclusion was randomized, counts by any other licensure data variables (e.g., race, region of the state) are accurate under the assumption that workforce participation is similar across the categories of those variables.

A more rigorous imputation approach was used to assign an estimated FTE value to each member of the potential nurse workforce missing this information. Survey respondents provided information on the number of hours they worked, and this information was used to assign an FTE value with the following formula:

$$\textbf{FTE weight} = \frac{\textbf{(hours/week} \times \textbf{weeks/year)}}{\textbf{1,976}}$$

In this formula, the numerator represents the hours worked per year by the respondent, and the denominator represents the hours worked in a year if a nurse represents 1.0 FTE. A person working 38 hours per week (midpoint of 36 and 40, two typical full-time schedules for nurses) and 52 weeks per year (including paid time off) will work a total of 1,976 hours in one year. Nurses working more than 1,976 hours per year were capped at 1.0 FTEs, while those working less than the 1,976 hours per year threshold were assigned a fraction of an FTE. When a nurse reported he or she was not working in nursing, the FTE was assigned a value of 0. The typical number of hours per year used in computations like this is 2,080, which is based on 40 hours worked per week for 52 weeks. Our survey collected "hours worked per week" in categories, and we used category midpoints to assign a single value to each respondent. Thus, full-time employees working 36-40 hours per week were assigned the midpoint value of 38, which explains why our denominator is less than the value more typically used in computing FTEs.

Although the FTE value for members of the potential nurse workforce is unknown if they did not complete a Workforce Survey, we were able to use information we had about each non-responding nurse to make an "educated guess" about how much he or she worked. Nurses who did not have survey data were assigned the average FTE (for survey respondents) of that person's license type, age group, and gender. In cases where age group was missing, the assigned FTE was based on gender only (and vice versa). In cases where both age and gender were missing, the overall license type FTE was assigned. The sum of these FTEs (from survey respondents and non-respondents alike) resulted in our estimate of 143,538 FTEs in the nurse workforce.

Once FTE values have been imputed, the resulting dataset can be directly analyzed to produce estimates of FTEs by license type, region, gender, and any other variable that exists in the licensure database. Estimation of FTEs by employment setting, provision of direct care, or any other survey variables required some additional steps. For example, we were able to use the percentage distributions among survey respondents to distribute our total estimate of FTEs into different employment settings.

Conclusions and Limitations of the Data Sources

All analyses of the Florida nurse supply based on licensure and Workforce Survey data inevitably suffer from some degree of missing or inaccurate data. The Center's process for cleaning the data, assigning nurses into counties and regions, and imputing missing data attempts to correct some of the data problems which, if left unchecked, would distort our view of the nurse supply. The exclusion process we use to identify the potential nurse workforce generates our best estimate of nurses who *could* be working in Florida, including their location in a specific

region of Florida. However, it is important to reiterate that licensure data do not indicate whether nurses are working in the field of nursing or how much they work. If Workforce Survey data are available for a nurse, it is a straightforward process to determine whether and where a nurse practices nursing. For those missing Workforce Survey data, the "whether" must be imputed and the "where" must be gleaned from address fields in the licensure database, which are known to have problems.

The incorporation of the Workforce Survey beginning in 2008 has dramatically improved data quality and facilitated our efforts to accurately quantify the nursing workforce. While there will always be at least some missing data, possessing complete workforce information on more than 90 percent of renewing nurses is a huge step forward for nurse workforce analysis and planning in Florida. We look forward to continuing collaboration with the Florida Board of Nursing and Medical Quality Assurance during the 2010-2011 nurse renewal cycles. The Center has modified its Workforce Survey to adopt the National Forum of State Nursing Workforce Centers' National Nursing Workforce Minimum Dataset for nurse supply, available at http://www.nursingworkforcecenters.org/resources/files/Nurse_Supply_Dataset.pdf. The result will be a much richer and nuanced set of supply data for analysis in January 2012.

Appendix A: 2008-2009 Workforce Survey Questions

1. Highest education degree completed:
    ___ Certificate – Licensed Practical Nurse
    ___ Diploma – Registered Nurse
    ___ Associate Degree
    ___ Bachelor Degree in Nursing
    ___ Bachelor Degree in field other than nursing
    ___ Master Degree in Nursing
    ___ Master Degree in field other than nursing
    ___ Doctorate in Nursing
    ___ Doctorate in field other than nursing

2. Current employment situation:
    ___ Employed in nursing (nursing license required for job)
    ___ Employed in field other than nursing
    ___ Seeking nursing employment
    ___ Currently not working and not looking for a job
    ___ Retired or with no plans to return to work

3. County of primary employment setting (if you are not working, please indicate your county of residence):
    [This list is presented as a drop-down box online.]

___ Other than in Florida

___ Alachua	___ Collier	___ Glades	___ Jackson	___ Marion	___ Pasco	___ Suwanee
___ Baker	___ Columbia	___ Gulf	___ Jefferson	___ Martin	___ Pinellas	___ Taylor
___ Bay	___ De Soto	___ Hamilton	___ Lafayette	___ Miami-Dade	___ Polk	___ Union
___ Bradford	___ Dixie	___ Hardee	___ Lake	___ Monroe	___ Putnam	___ Volusia
___ Brevard	___ Duval	___ Hendry	___ Lee	___ Nassau	___ St Johns	___ Wakulla
___ Broward	___ Escambia	___ Hernando	___ Leon	___ Okaloosa	___ St Lucie	___ Walton
___ Calhoun	___ Flagler	___ Highlands	___ Levy	___ Okeechobee	___ Santa Rosa	___ Washington
___ Charlotte	___ Franklin	___ Hillsborough	___ Liberty	___ Orange	___ Sarasota	
___ Citrus	___ Gadsden	___ Holmes	___ Madison	___ Osceola	___ Seminole	
___ Clay	___ Gilchrist	___ Indian River	___ Manatee	___ Palm Beach	___ Sumter	

4. Present employment status at primary employment location: ___ FT   ___ PT ___ Per Diem/Agency

5. Do you work for more than one employer?   ___ Yes   ___ No

6. **TOTAL** number of hours worked in a typical **WEEK** at **ALL JOBS**:
    ___ **less than 10**
    ___ **10-15**
    ___ **16-20**
    ___ **21-25**
    ___ **26-30**
    ___ **31-35**
    ___ **36-40**
    ___ **41-45**
    ___ **46-50**
    ___ **more than 50**

7. Number of weeks per year that you work at **ALL JOBS,** including paid time off (year round employment = 52 weeks):

     ___ **0-10**
     ___ **11-20**
     ___ **21-30**
     ___ **31-40**
     ___ **41-50**
     ___ **51 or 52**

8. If you work in nursing, select one setting that best describes your primary nursing employer:

     ___ Hospital
     ___ Ambulatory Care
     ___ Public/Community Health
     ___ Occupational Health
     ___ Long Term Care
     ___ Home Health Care
     ___ Insurance Company
     ___ Nursing Education – Academic Setting
     ___ School Health
     ___ Physician or other Health Provider Office
     ___ Temporary Agency
     ___ Healthcare Consulting / Product Sales
     ___ Corrections Facility
     ___ Other

9. If you work in nursing, does your primary nursing position involve providing **DIRECT CARE SERVICES** to patients/families?

     ___ Yes   ___ No

Appendix B: County Composition of FCN Regions
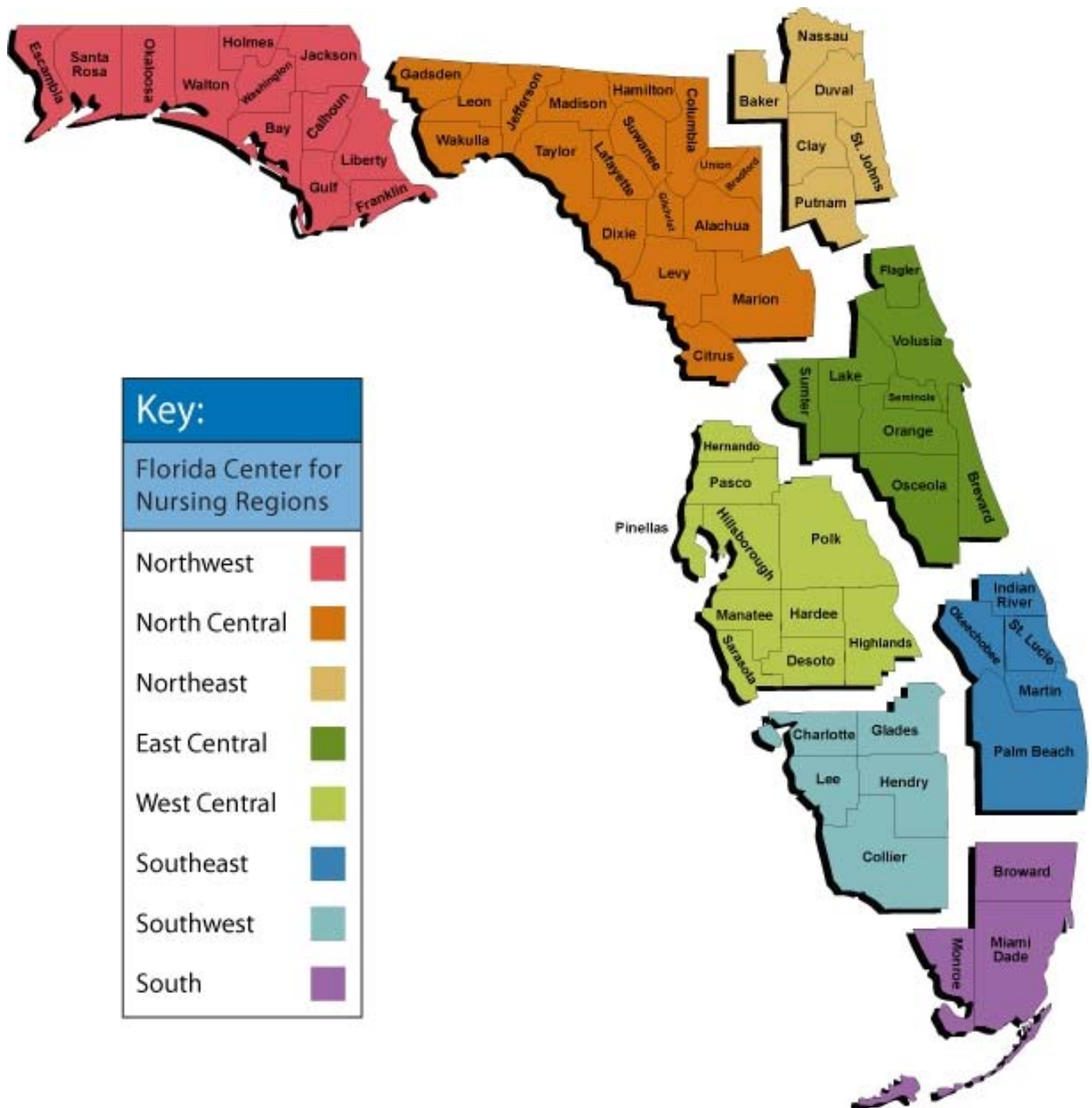
Key:

Florida Center for Nursing Regions

Northwest
North Central
Northeast
East Central
West Central
Southeast
Southwest
South