George Zangaro
David I. Auerbach
Douglas O. Staiger
Aprita Chattopadhyay
Peter I. Buerhaus

Improving Nursing Workforce Forecasts: Comparative Analysis of the Cohort Supply Model and the Health Workforce Simulation Model

Executive Summary

- Forecast models can help policymakers, educators, and others to anticipate and take action to forestall or mitigate, surpluses or shortages of key healthcare providers.
- Confusion and uncertainty over forecasts of registered nurse (RN) supply, demand, and shortages is particularly unsettling to organizations, educators, and public policymakers.
- Two of the leading forecasters of the RN workforce jointly convened a meeting of the nation's experts on modeling the supply and demand for nurses to better understand and reconcile underlying model differences, and discuss communication, messaging, and other strategies around workforce forecasts.
- Key details of two supply forecasting models – the Cohort supply model and the Health Workforce Simulation Model – are discussed.
- A comparative analysis, key factors that workforce analysts should monitor to anticipate changes in the future size of the nation's RN workforce, and recommendations for how the two forecast models can be modified are presented.

BALANCING THE SUPPLY and demand of healthcare professionals is critical to a well-functioning healthcare delivery system. Shortages of providers can have severe consequences for patients who need care, and can induce burnout and overwork among scarce providers. Surpluses of providers can also have harmful effects; healthcare professionals invest heavily in their careers and cannot always easily shift to other careers when no jobs are available.

Forecast models can help policymakers, educators, and others to anticipate and take action to forestall or mitigate surpluses or shortages of key providers. For example, forecast models of the nursing workforce in the early 2000s were projecting large nursing shortages by the 2010s and 2020s as Baby Boomers retired and smaller cohorts of registered nurses (RNs) followed them (Buerhaus, Staiger, & Auerbach, 2000; U.S. Health Resources and Services Administration (HRSA) (2002). These forecasts were instrumental in fostering public and private initiatives to boost interest in nursing and enrollment into nursing education programs, which doubled between 2002 and 2012 (Auerbach, Staiger, Muench, & Buerhaus, 2013).

Forecast models rely on a number of predictable factors, such as demographic trends and observed workforce patterns to make inferences about potential future imbalances. Not surprisingly, because these highly complex models use different underlying data sources,

DAVID I. AUERBACH, PhD, is External Adjunct Faculty, Montana State University, Bozeman, MT.

APRITA CHATTOPADHYAY, PhD, is Chief, Workforce Analysis Branch, National Center for Health Workforce Analysis, Bureau of Health Workforce, Health Resources Service Administration, Rockville, MD.

GEORGE ZANGARO, PhD, RN, FAAN, is Director, National Center for Health Workforce Analysis, Bureau of Health Workforce, Health Resources Service Administration, Rockville, MD.

DOUGLAS O. STAIGER, PhD, is John French Professor of Economics, Department of Economics, Dartmouth College, Hanover, NH; and Research Associate, National Bureau of Economic Research, Cambridge, MA.

PETER I. BUERHAUS, PhD, RN, FAAN, FAANP(h), is Professor of Nursing and Director, Center for Interdisciplinary Health Workforce Studies, College of Nursing, Montana State University, Bozeman, MT.

model structures and assumptions, and serve somewhat different purposes and audiences, their projections often do not agree. In the case of the RN workforce, this variation in forecasts leads to confusion among the many stakeholders who rely on these forecasts (such as federal and state governments, healthcare delivery organizations, educational institutions, and staffing firms) to best prepare their organizations for the future (Spetz, 2015).

Given evolving changes in delivery systems, implementation of payment reform, adoption of a population focus, retirement of Baby-Boomer RNs, and physician shortages, the confusion and uncertainty over forecasts of RN supply, demand, and shortages is particularly unsettling to organizations, educators, and public policymakers (Buerhaus, Skinner, Auerbach, & Staiger, 2017). As a researcher recently noted about these varying forecasts, "what is a nurse leader or educator to do?" (Spetz, 2015, p. 178).

Partly in response to this challenge, two of the leading forecasters of the RN workforce – HRSA Bureau of the Health Professions, and Montana State University's Interdisciplinary Center for Healthcare Workforce Studies – jointly convened a meeting of the nation's experts on modeling the supply and demand for nurses to better understand and reconcile underlying model differences, and discuss communication, messaging, and other strategies around workforce forecasts.

Following the July 2016 Montana meeting, key investigators of both models discussed similarities and differences in the respective supply forecasting models. To elucidate the areas that make a major difference in each model's output, it was decided to conduct a side-by-side empirical analysis of underlying assumptions, variable definitions, and determine their impact on supply forecasts made from each model over the same projection time period. This article reports on the results of this comparative analysis, identifies the key factors that workforce analysts should monitor to anticipate changes in the future size of the nation's RN workforce, and makes recommendations for how the two forecast models can be modified.

Model Descriptions

Key details of the two supply forecasting models are discussed, beginning with the Cohort supply model and followed by the Health Workforce Simulation Model (HWSM) developed by IHS Markit and used by HRSA.

Structure of the Cohort Supply Model

The model first used by Buerhaus and colleagues (2000) (Cohort model) is an age-cohort based model derived from the field of labor economics that decomposes the proportion of each birth cohort (defined by birth year) working as RNs in each year into the product of two components: (a) a cohort effect that estimates propensity of individuals born in any given year to work as an RN (which captures, for example, changes across birth cohorts in perceived attractiveness of a nursing career relative to other occupations); and (b) an age effect that captures propensity of RNs to be working at different ages over their career lifespan (which captures, for example, life cycle patterns such as retirement and tendency of female RNs to work less during their childbearing years) (Auerbach, Buerhaus, & Staiger, 2015; Buerhaus et al., 2000).

Intuitively, with respect to the age effect, the Cohort model seeks to estimate the trend in number of employed RNs from a given birth cohort as people in that cohort age. For example, of people born in 1950, the age effect estimates how many people are working in nursing as RNs in 1975 (at age 25); and later in 2000 (at age 50). These patterns naturally incorporate factors such as RNs leaving to work part-time in their 30s to care for children (which would manifest as fewer full-time equivalent [FTE] RNs for a given cohort when they are in their 30s compared to when they are in their 50s), or a decrease in work output or labor force participation when the cohort begins to retire (which would manifest as fewer observed RNs in nursing when the cohort is age 65 vs. 50). Because these factors are critical to the Cohort model, stability of the age patterns have been routinely and carefully examined over the past 15 years, and adjustments made where necessary. In the 1990s, for example, RNs began entering nursing education programs at later ages than their predecessors. Auerbach, Buerhaus, and Staiger (2007) described this phenomenon and adjusted the model accordingly to capture this effect.

In contrast to the age effect, the cohort effect is more straightforward. While age effects describe FTE RN production of the 1950 cohort at age 50 versus 25, for

NOTES: The study was funded by the U.S. Department of Health and Human Services, Health Resources and Services Administration under contract number HHSH250201600003A. The views expressed in this article are solely the opinions of the authors and do not necessarily reflect the official policies of the U.S. Department of Health and Human Services or the Health Resources and Services Administration, nor does mention of the department or agency names imply endorsement by the U.S. Government.

The July 2016 meeting of workforce experts was funded, in part, by a grant from the Gordon & Betty Moore Foundation.

ACKNOWLEDGMENT: The authors acknowledge the Health Resources and Services Administration and The Gordon & Betty Moore Foundation for supporting the 2016 Montana Supply/Demand Meeting and to all of the participants who attended the meeting.

A *Nursing Economic$* Podcast conversation with Peter Buerhaus regarding improving nursing workforce forecasts can be accessed at www.nursingeconomics.net

example, cohort effects describe FTE RN production of the 1975 cohort at age 50 versus the 1950 cohort at age 50 (or any given age). Factors such as alternative career opportunities for women or economic conditions at the time career decisions are made are incorporated into the cohort effects. Thus, all factors influencing cohort and age effects are subsumed naturally in the data and assumed to be permanent features of that cohort. If declines in manufacturing jobs and expanded opportunities for RNs due to the Affordable Care Act lead more people born in the 1990s to become RNs, the Cohort model implicitly assumes those cohorts will be permanently larger than other cohorts.

Internationally educated RNs who enter the workforce are also implicitly built into age and cohort effects rather than explicitly modeled. Roughly 5.6% of the RN workforce is educated in other countries and enters the United States to work as RNs (U.S. Department of Health and Human Services [DHHS], 2013). When these RNs enter the country and become employed, they are captured in underlying census data and added to whatever age and cohort they belong – increasing those age and cohort effects in whatever pattern these entrants display. Changes in immigration policy that might affect immigration of RNs en-masse might affect workforce numbers in a way that the model would not anticipate, but given the relatively small proportion of internationally educated RNs, this issue is likely to not be of significant concern when doing national projections (in certain states, such as California or New York, internationally educated RNs compose a larger portion of the workforce and state-level forecasts would want to include these RNs in projection models).

Despite these and other changes the model cannot account for, in prior work, the Cohort model accurately predicted number and age distribution of FTE RNs both in sample and in out-of-sample forecasts (Buerhaus et al., 2000). For example, using data through 1988, the model predicted 43% total workforce growth by 1998, near the 35% growth that actually occurred. The model also successfully captured a dramatic aging dynamic of the workforce by predicting the percentage of the workforce that was under age 40 would decrease from 59% in 1988 to 38% in 1998 (the actual percentage in 1998 was 42%).

Data used by the Cohort supply model. The Cohort model requires many years of repeated observations of the nursing workforce to generate estimates of age and cohort effects. It relies on data from the U.S. Census Bureau Current Population Survey (CPS) from 1979 to 2000, and the Census Bureau's American Community Survey (ACS) from 2001 to 2015. The CPS is a household-based, nationally representative survey of over 100,000 individuals administered monthly by the Census Bureau. The CPS has asked detailed questions about employment (including occupation and hours worked) in a consistent manner since 1979 and is used extensively by the U.S. Department of Labor to estimate current trends in unemployment, employment, and earnings. CPS data have been used in prior work to estimate employment trends for RNs and project age and supply of both RNs and physicians in the United States (Auerbach, Buerhaus, & Staiger, 2014; Staiger, Auerbach, & Buerhaus, 2009).

The ACS, which began reporting data in 2001, is modeled after the long form of the decennial census and, although it contains fewer questions than the CPS, it obtains a much larger sample size (U.S. Department of Commerce, 2017). While the CPS surveys roughly 3-4,000 RNs per year, the ACS surveyed approximately 12,000 RNs in each year from 2001 to 2004, and more than 30,000 RNs in each year starting in 2005 (when the sample was enlarged). Because these larger sample sizes allow for recent workforce trends to be analyzed with greater accuracy, ACS data were used rather than CPS data beginning in 2001. The occupation and employment questions in the ACS are similar to the CPS and generate similar estimates of total RN employment for overlapping years. HRSA has also begun using the ACS in its efforts to produce estimates of nursing supply and demand. To make estimates representative of the U.S. non-institutionalized population, observations are weighted by sampling weights provided by the CPS and ACS. Additional data on the U.S. population by year and by age between 1979 and 2013 were obtained from the Census Bureau. Forecasts of the U.S. population by age through 2030 were obtained from the "middle series" projections prepared by the Census Bureau.

The Cohort model includes all individuals age 23-69 who reported being employed as an RN during the week of the survey. Final workforce estimates are adjusted to represent RNs under age 23 and over age 69, who collectively contribute less than 5% of workforce output and whose numbers are too small to generate reliable age-cohort coefficients. In work to date, we have assigned RNs reporting fewer than 30 hours worked in a typical week as a 0.5 FTE and those working more than 30 hours at 1.0 FTE. All individuals reporting their occupation as an RN are included in the sample.

Cohort model estimation and projections. The coefficients of the Cohort model are estimated in log form, using ANOVA to estimate age and cohort effects. The dependent variable in the model is the logarithm of the number of FTE RNs of every age between 23 and 69 for every year between 1979 and 2015 (47 years of age x

37 years = 1,739 total observations) divided by the total U.S. population in that given year-age cell. The ANOVA model estimates main effects for cohort (birth year) and age, as well as interaction effects that allow for a different set of age effects below age 30 for cohorts born after 1964, and a different set of age effects above age 50 for cohorts born after 1940 as described previously. Predictions from this model are exponentiated and multiplied by the U.S. population in that cohort-age cell to yield predictions of the number of FTE RNs. All statistical analyses were performed using Stata 13.1.

Estimates of age and cohort effects are used to project numbers of FTE RNs through 2030. It is assumed age effects in future years and for future cohorts will be the same as those estimated for cohorts born after 1964. For cohorts that have already entered the labor market (age 23 or older in 2015, born 1992 or before), age effects are used in combination with the estimated cohort effect for each birth year to project FTE RNs supplied by each cohort as RNs in that cohort grow older. For cohorts that will enter the workforce in the future (born after 1992), it is assumed their cohort effect will equal the average of the five most recent RN cohorts observed (1988 through 1992 birth cohorts, who were observed at age 23 in 2011-2015). In other words, it is assumed the propensity of future cohorts to enter nursing will be similar to the most recently observed cohorts.

Structure of the Health Workforce Simulation Model

In contrast to the Cohort model that is based on age and cohort effects and aggregate rates, the HWSM used by HRSA to make projections is a microsimulation model using a simple accounting process to move the workforce forward in time. Starting with the number of RNs active in the workforce in 2012, the HWSM tracks nurses by age, sex, state, and highest education level attained to produce annual projections of RN supply through 2025. For each year, the model adds the estimated number of newly licensed RNs, subtracts the estimated number of separations (deaths and RNs who retire), and calculates an end-of-year estimate of licensed RNs. The end-of-year estimate becomes the starting value for the next year's projections. To estimate the number of RNs active in the health workforce and the number of FTE RNs employed in healthcare, the model projects the number of licensed RNs and then applies workforce participation rates and estimated number of hours worked to each active RN.

Data used by the HWSM. Primary data used by the current HWSM are pooled 2006-2011 ACS files. Standard Occupational Codes are used to identify RNs in the ACS sample. To determine current or base year supply of RNs, HWSM recalibrated the 5-year ACS (2006-2011) sample weights to sum to the national totals in 2012 ACS. The weighted count of all RNs who were less than 75 years of age and active in the workforce are taken as the total base year workforce. HWSM combines multiple years of data to obtain stable estimates of the number and characteristics of RNs by age, sex, state, and education level in the base year, as well as generate robust estimates of participation rates and hours worked by future nurses. Additional information on labor market characteristics that may influence RNs' work effort, such as state unemployment rate and average wages, are obtained from the Bureau of Labor Statistics' Local Area Unemployment Statistics and Occupational and Employment Statistics and attached to individual records. HWSM uses these variables in regression equations that predict number of hours each RN will work. Estimates of survival probabilities required to "age" the current workforce forward are derived from the age-sex specific mortality rates from the Centers for Disease Control and Prevention's (CDC) mortality tables.

HWSM uses candidates educated in the United States taking the NCLEX-RN for their first time to estimate the number of new entrants to the nursing workforce. NCLEX does not publish information on age and sex or educational attainment of NCLEX passers. Therefore, HWSM used age-sex distribution from a 2012 National League for Nursing (NLN) survey of students enrolled in entry-level nursing programs during the 2008 academic year to impute age and sex of new RNs (NLN, 2013). Information on educational level and state are taken from data on first-time NCLEX-RN takers (National Council of State Boards of Nursing [NCSBN], 2014).

Forecast methods. The HWSM starts with a micro data file that mirrors the universe of the current (base year) RN workforce derived from pooled 2006-2011 ACS data. This is done by using ACS weights to replicate individual records of the sample such that the final file is inflated to contain unique records representing each individual in the current (base year) RN workforce. Once the population dataset is created from the ACS sample, a unique workforce outcome can be simulated for each individual based on the individual's specific characteristics. In states with small sample sizes, creating multiple records helps "smooth" the impact of individual characteristics on labor supply decisions (e.g., retirement). The base year file is aggregated to determine the current active RN workforce, as well as to estimate regression equations that are used to simulate future workforce decisions.

The mechanism for simulating new entrants into the workforce is accomplished by creation of a "synthetic" cohort of RNs, which is based on the number and characteristics of recent National

Council Licensure Examination (NCLEX-RN) test takers as mentioned earlier. HWSM creates one record for each NCLEX-RN test taker, with information on age, sex, state, and educational level. Information on these variables is used to simulate labor market status (active/inactive) and number of hours worked in current and future cohorts of RNs.

In addition to stimulating new entrants, there are three other components that determine the future size of the RN workforce: (a) survival probabilities of the current cohort of RNs, (b) labor force participation rates, and (c) number of hours worked by future RN cohorts. In Step 1 of the HWSM, individuals in the base year micro data are aged forward according to their age-sex specific survival rates. Step 2 of HWSM simulates activity status of RNs projected to be alive, based on age and sex specific workforce participation rates. In other words, in Step 2 workforce participation rates are applied to RNs who are alive to determine which RNs would be projected to be working and which would not. Steps 1 and 2 are repeated until the end of the projection period is reached. Step 3 estimates the expected number of hours each RN in future cohorts will work. The HWSM uses predicted values obtained from a regression of hours worked on age, sex, education level, state unemployment rate, and average wage in the profession. Simulated numbers of work hours at the individual level from the "aged" data and new entrants files are aggregated to obtain the projected workforce in person-hours. This result is divided by average number of hours worked by the baseline RN population to obtain projections of FTE RNs.

HWSM estimation and projection. To "age" the base year supply data forward over time and simulate future workforce outcomes at the individual RN level, the HWSM moves from a sample file to a "census" type file that includes a record for each RN. To develop the latter file, records of survey participants are replicated according to their sample weight. For example, if a nurse's record in the recalibrated ACS has a sample weight of 100, 100 records with the same characteristics are created in the HWSM micro data file. Survival probabilities (1 minus the probability of dying at any given age) from CDC are adapted to account for the fact age-adjusted mortality rates through age 65 for professional and technical occupations are approximately 25% lower than overall national rates for men and 15% lower for women.

The estimated survival rates are applied to individual records in the micro data file through a random number generation process using a uniform (0, 1) distribution. Depending upon the value of the random number and estimated probability of dying, records in the HWSM micro data file are deleted. The resultant file is survivors of the previous year's cohort of RNs. To these survivors, cohorts of new RNs are added for every year in the projection period.

For example, in 2012, there were 150,266 U.S.-educated first-time test takers of the NCLEX-RN. Of these, 62,535 RNs had completed a baccalaureate degree and 87,731 had completed a diploma or an associate degree (NCSBN, 2014). HWSM assumes these numbers of 2012 graduates would continue annually throughout the forecast window (the nation would produce 62,500 new RNs [baccalaureate level] and 87,700 new RNs [less than the baccalaureate level]). The new RN supply includes the estimated 16,000 licensed practical nurses who further their education to become RNs each year. HWSM also assumes all RNs educated in a given state became new entrants to the nursing workforce in that same state. Estimates on the number and characteristics of future entrants to the RN workforce are made under the assumption current patterns continue throughout the projection period. HWSM applies the age-sex distribution derived from the NLN survey on first-time NCLEX-RN test takers (which includes foreign-educated nurses). A record for each new entrant is created and HWSM generates a series of random numbers. Depending upon the value of random numbers and probability of having a particular characteristic, the individual record is assigned that characteristic.

Estimates of workforce participation rates by age and sex come from 2006-2011 ACS data. For ages below 50 years, HWSM calculates the age-sex specific participation rates directly. ACS does not capture the profession for individuals out of the workforce for 5 years or more but does capture education, so activity rates based on the highest educational attainment are used for ages over 50 years. ACS data are used to determine the highest level of education (less than baccalaureate degree, baccalaureate degree, and graduate degree) and to calculate labor force participation rates of each education group. RNs over age 50 with baccalaureate education are assumed to have similar labor force participation rates as other women with a baccalaureate degree. For RNs educated at the associate-degree level, the HWSM uses participation rates of women educated with an associate degree in nursing. HWSM estimates approximately 95% of RNs aged below 50 to be active in nursing. However, from age 50 onwards, labor force participation rates decline quite precipitously. RNs with graduate degrees have slightly higher activity rates than RNs of similar age who have a baccalaureate or less education. HWSM assumes age-specific workforce participation rates would remain unchanged in future years.

HWSM determines the expected number of hours worked for RNs projected to be active in the workforce using a two-step regression model. First, expected hourly

Table 1.
Impact of Changing FTE RNs by Age Groups in 2015 Using the Cohort Supply Forecasting Model

	Age 21-34	Age 35-49	Age 50 and Over	Total
Old method	850,123	1,086,945	1,105,247	3,042,316
New method	875,795	1,145,887	1,165,990	3,187,672
Percent difference	3.0%	5.4%	5.5%	4.8%

SOURCE: Cohort forecast model
FTE = full-time equivalent; RNs = registered nurses

wage is estimated from 2006-2011 ACS data on RNs working at least 20 hours per week by regressing hourly wage on age, sex, state unemployment rate, and mean U.S. wage for RNs. State unemployment rates and U.S. mean wages are introduced as time-varying covariates in the model. In the second step, number of hours RNs work per week is estimated using the following regression specification:

Log (hrs) = $\alpha + \beta_1$ (unemp) + β_2 (age) + β_3 (sex) + β_4 (log predicted wage) + β_5 (educ) + ε

Predicted hours worked are aggregated across the simulated micro file for each forecast year to obtain the total person hours of RN supply. HWSM converts person-hours of RN supply into FTEs by dividing average number of hours worked by RNs in 2012. HWSM used 35.8 hours per week as a measure for 1 FTE.

Key Assumptions Underpinning the Cohort and HWSM Models

This section describes four key assumptions and aspects of the models that were identified during the Montana meeting that appear to drive differences in output when projections from the two models are compared.

Defining and counting FTE RNs. A difference between models is in how they count FTEs, which creates a divergence in workforce estimates. The Cohort model counts workforce output as FTE RNs, where an FTE is defined as 0.5 FTE for an RN reporting usual weekly hours worked between 1 and 30 hours, and as 1.0 FTE for an RN reporting usual weekly hours greater than 30. The HWSM calculates an FTE differently by computing the fraction of full-time hours worked by each RN, where full-time hours are defined as average hours worked among RNs working more than 20 hours per week. A virtue of this method is it gives greater weight to RNs working 45 hours versus those working 35 hours, whereas such RNs are treated equivalently under the Cohort model.

Based on discussions at the Montana meeting, both teams decided to align their assumptions: (a) the Montana Cohort Team would switch its method to that used by the HWSM model, and (b) both models would, rather than using a definition of the number of hours worked that is deemed to be 1.0 FTE based on observing hours worked in the actual data, use a standard definition of a 40-hour workweek. This definition is grounded in labor economics and aligns with those used by the Bureau of Labor Statistics, and also avoids having to change the definition each year based how actual hours worked may change in the data.

This change has implications for both models – for the Cohort model, the net change in the overall workforce size is small – but as noted previously, internal shifts in weighting will occur among all RNs in the model, with those working more hours generally weighted more heavily. With respect to the HWSM model, since the prior model had used a full-time workweek definition of roughly 36 hours, the net effect of this change will be to reduce the effective size of the workforce by roughly 10%, although each RN within the model will have the same weight relative to each other as before.

Applying the HWSM method of counting an FTE to the Cohort method (using 40 hours as the base estimate of a full-time RN) results in a 0.6% higher net FTE count in the entire workforce in 2015 and increases average age of the workforce by one-tenth of 1 year, from 43.9 years to 44.0 years. As observed in Table 1, the new method reduces slightly the FTE count among younger RNs and increases it slightly among older RNs because older RNs tend to work longer hours and the new method assigns a relatively higher FTE to those with more hours worked. (There are some exceptions; e.g., the old method used by the Cohort model assigned 1.0 FTE for an RN working 31 hours and 0.5 for an RN working 29 hours, and in the new method, each RN is assigned roughly 0.75 FTE.)

The number of hours worked per week used to define a 1.0 FTE in the HWSM model was 35.8 hours in previous model versions, but increased to 40 in its most recent version. Comparisons of total workforce size in tables later in this article will make use of this aligned assumption between the models.

Inclusion of advanced practice RNs (APRNs). The HWSM model distinguishes between RNs and APRNs (nurse practitioners, nurse midwives, nurse anesthetists, and clinical nurse spe-

Table 2.
Impact of Removing APRNs from the Cohort Model by FTE Age Group in 2015

	Age 21-34	Age 35-49	Age 50 and Over	Total FTEs
With APRNs	840,763	1,100,052	1,119,350	3,060,165
Without APRNs	799,771	1,029,647	1,044,562	2,873,980
Percent difference	-4.9%	-6.4%	-6.7%	-6.1%

SOURCE: Cohort forecast model
APRN = advanced practice registered nurses, FTE = full-time equivalent

Table 3.
Net Change in Total FTE RNs Estimated by the Cohort Model After Adjusting FTE Definition and Eliminating APRNs (2015)

	Age 21-34	Age 35-49	Age 50 and Over	Total FTE RNs
Net change from both adjustments	-5.9%	-5.3%	-5.5%	(168,335)

SOURCE: Cohort forecast model
APRNs = advanced practice regististerd nurses, FTE = full-time equivalent, RN = registered nurse

cialists). The HWSM excludes APRNs from its base RN model. APRNs typically perform different roles in the nursing workforce and their clinical activities often substitute for physicians; therefore, the Montana Cohort team agreed that it is reasonable to exclude APRNs when making projections of the general nursing workforce. In 2011 the ACS started identifying APRNs in its survey (with the exception of clinical nurse specialists), which does not present a problem for the HWSM as this model does not rely on older data.

In general, the Cohort model included APRNs in the model without distinction from other RNs (however, some analyses excluded APRNS in sensitivity analysis). Following discussions at the Montana meeting, the Cohort model team agreed with the logic of excluding APRNs from the main Cohort model. To exclude APRNs from the Cohort model, the Montana team uses a regression-imputation method to identify factors associated with APRNs in the years for which they are identified in the data (2011-2015) and impute APRN status based on those same factors and coefficients for earlier years (1979-2010). This is necessary because this model relies on data going back to 1979, unlike the HWSM. Significant coefficients were obtained for measures of nursing degree (education), age, hours worked, industry sector, and earnings ($r^2 = 0.3$).

The exclusion of APRNs from the model has an impact on projections that is captured naturally in age and cohort effects. For example, to the extent RNs leave to become APRNs at certain ages, this change will be manifested in fewer RNs working at those ages and, thus, a diminished age effect at those ages. To the extent this phenomenon is larger in some cohorts than others, it will be manifested as a difference in the cohort effect for those cohorts.

As seen in Table 2, APRNs represent roughly 6.1% of workforce FTEs in 2015. The impact of their removal on workforce size by age reveals older RNs are more likely to be APRNs (primarily because the additional education takes time, and many RNs embark on this additional education later in their careers).

On net, combining the new FTE definition change with the elimination of APRNs from results in a decrease of roughly 170,000 FTEs in 2015 for the Cohort model (see Table 3).

Entry of new RNs into the workforce: Cohort model. As discussed earlier, the Cohort model's assumption of new entry into the RN workforce is based on the estimated sizes of the last five cohorts of RNs observed in the ACS data. That is, in 2015, the most recent cohort (because the model begins with RNs who are 23 years of age) comprises RNs born in 1992 who are observed in the ACS data only once (in 2015). The second most recent cohort are RNs born in 1991 who are observed twice in the data: at age 23 in 2014 and age 24 in 2015. The forecast also allows for possibility of a recent trend in entry based on observations of recent changes in NCLEX test takers; where such a trend is present, an adjustment is made to the model in corresponding direction. Since NCLEX trends have been relatively flat in recent years, this adjustment has essentially no effect as of 2015.

These assumptions are reflected in the number of RNs present in the workforce (from 2012-2015) and who are forecast to be in the workforce over the years 2016-2025 (see Table 4). The last several cohorts of RNs (those born in the late 1980s and early 1990s) are, in fact, larger than any cohorts of RNs observed previously, with a likelihood to be working as an RN roughly 60% greater than the previously largest RN cohort (those born in 1955).

The Cohort model finds that the number of RNs under age 35 will grow from 681,000 in 2012 to 950,000 in 2025 (39.4%), reflecting increased rates of entry over the course of the past decade.

Table 4.
Number of Younger RNs Estimated by the Cohort Model, 2012-2025

Year	RNs Age <35
2012	681,722
2013	692,139
2014	750,474
2015	800,918
2016	849,688
2017	879,032
2018	905,285
2019	925,311
2020	938,249
2021	947,613
2022	955,495
2023	954,850
2024	955,548
2025	950,296
Percent change (2012-2025)	39.4%

SOURCE: Cohort forecast model
RNs = registered nurses

Entry of new RNs into the workforce: HWSM model. For each forecast year, the HWSM simulates a fixed number of 152,000 new RNs with demographic characteristics mirroring the NLN (2013) survey to enter the workforce (see Table 5). This number is based on average number of NCLEX first-time domestic test takers over the last few years. The HWSM also converts counts of new RNs to FTEs by first estimating the workforce participation rates and then the number of hours new RNs work using the same activity patterns as in the base year. The implication of this assumption is the contribution of new RNs in the workforce is driven by their age and educational distribution as they enter the workforce.

Comparison of the Two Models

The ways that entry into the RN workforce is handled in each model highlights different model structures. In the HWSM, new waves of nursing school graduates are explicitly added to the model in each successive year. In the Cohort model, each cohort of RNs follows a labor force contribution trajectory estimated by the model's internal data from past years of data. Thus, the concept of "new entry" is implicit in the Cohort model but explicit in the HWSM model and, consequently, comparisons are not straightforward.

As an exercise, to approximate a comparison between the two models' estimates of how many new RNs were entering the workforce in recent years, the Montana team approximated new entry by (a) adding RNs observed in the model in 1 year (say, 2012) under age 40, and (b) then adding RNs observed in the model in the following year under age 41, with the addition of the new group of 23-year-old RNs who had not been present in the model the previous year. In the years 2011-2015, this method yields an estimate of 125-130,000 RNs entering the workforce annually. These figures are slightly below the HWSM figures, but after attrition is accounted for in the HWSM, the amount of overall net entry is broadly similar.

To assess whether these numbers of entering RNs are reasonable, the team investigated entry trends obtained from the 2008 National Sample Survey of Registered Nurses in the United States (NSSRN) (DHHS, 2010) along with data from the NCLEX first-time domestic test takers (used by HWSM in its explicit entry assumptions). We identified RNs in the sample in 2008 who stated their year of graduation from their basic nursing education in any of the years between 2001 and 2008. The number of NCLEX first-time test takers in a given year nearly exactly matched the number of RNs who reported in the NSSRN they graduated in that given year. This result confirms number of NCLEX first-time test takers is a good proxy for number of eventual nurses who obtain a RN license (see Table 6). However, number of FTEs reported by that group of RNs was roughly 85%-90% as high as number of licensed RNs because some worked part-time, some worked in areas other than nursing, and some were not working at all (note the percentage of NCLEX test takers is lower in the earlier years,

Table 5.
Age and Sex Distribution of New RNs Using the HWSM

Number of RN Graduates	Female (%)	Age Distribution (%)			
		<25	26-30	31-40	>41
152,000	85%	42%	21%	24%	14%

SOURCE: NLN, 2013.
HWSM = Health Workforce Simulation Model, RNs = registered nurses

Table 6.
Comparison of NCLEX Test Takers to FTE RNs Identified in the 2008 NSSRN

Graduation Year	NCLEX Domestic First-Time Test Takers	Total RNs in the 2008 NSSRN who Graduated in the Given Year	Nursing FTE of those RNs	# FTE/# NCLEX Test Takers
2001	66,399	70,594	59,663	84.5%
2002	68,203	65,294	55,685	85.3%
2003	74,047	73,668	63,544	86.3%
2004	83,903	86,460	75,761	87.6%
2005	95,540	99,288	88,017	88.6%
2006	106,738	109,152	100,197	91.8%
2007	115,675	115,195	106,241	92.2%

FTE = full-time equivalent, NSSRN = National Sample Survey of Registered Nurses, RNs = registered nurses

Column 5 in Table 6). Further, this pattern suggests that as more time passes since the year of graduation, fewer RNs are working on a FTE basis in nursing. That is, starting in 2008 and looking back to the cohort that graduated in 2001, the number of FTEs from that year was only 84.5% as high as the number of NCLEX test takers versus more than 90% for more recent cohorts. Thus, the Cohort model's estimate of new entry into the RN workforce of 130,000 RNs per year is consistent with roughly 150,000 NCLEX domestic first-time test takers in recent years.

Thus, although the models are different in how they estimate new entry, they arrive at similar estimates of growth of new entry, as reflected in growth in the number of younger RNs from 2012 to 2025 as shown in Table 7.

Both models forecast the same degree of growth in young RNs over the 2012-2025 period (39.4% vs. 39.5%), although the HWSM estimates a slightly larger number of FTE RNs at any given point in time.

Table 7.
Number of RNs Under 35 Years of Age, by Year, Estimated by the HWSM and Cohort Models

Year	RNs Age <35, HWSM Model	RNs Age <35, Cohort Model
2012	692,699	681,722
2013		692,139
2014		750,474
2015		800,918
2016		849,688
2017		879,032
2018		905,285
2019		925,311
2020		938,249
2021		947,613
2022		955,495
2023		954,850
2024		955,548
2025	966,110	950,296
Percent change (2012-2025)	39.5%	39.4%

NOTE: Both models use the same definition of an FTE (40 hours) and exclude APRNs
APRNs = advanced practice registered nurses, FTE = full-time equivalent, HWSM = Health Workforce Simulation Model, RNs = registered nurses

Exit of RNs from the workforce: Cohort model. Similar to entry, the Cohort model does not explicitly model retirement from the workforce; it is captured in the age effects, which begin to diminish strongly after age 60. As with entry, these estimates are based on observations from 1979-2015, and indicate the number of RNs remaining in the workforce at older ages compared to the number observed when that same cohort of RNs was younger. These estimates were adjusted for more recent cohorts after an observed shift toward later retirement over the 2000s (Auerbach et al., 2014). As an example of the effective retirement rates, for any given cohort, about three-fourths of RNs observed working at age 50 are no longer in the workforce by age 69.

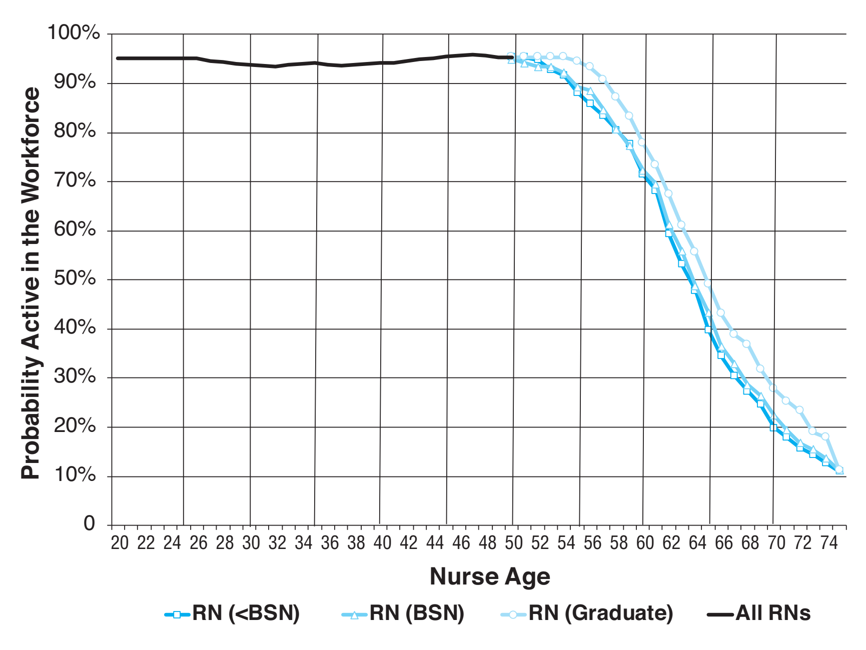
To illustrate the retirement rate implicit in the Cohort model,

Table 8.
Number of FTE RNs Remaining in the Workforce from the Baby Boom Generation (RNs Born Between 1946-1964)

Year	Baby Boomers in the RN Workforce
2012	1,106,936
2013	1,028,938
2014	1,002,976
2015	948,571
2016	900,712
2017	839,016
2018	769,935
2019	713,061
2020	637,818
2021	577,735
2022	515,905
2023	447,789
2024	391,509
2025	340,836
Percent change (2012-2025)	-69.2%

FTE = full-time equivalent, RNs = registered nurses

Figure 1.
Estimated Percent of RNs Remaining Active in the Workforce by Type of Nursing Education, 2006-2011



SOURCE: DHHS, 2015

the number of RN FTEs produced by the Baby-Boomer cohort in each year from 2012 (when RNs were between age 48-66) to 2025 (when they will be between age 61-79) is shown in Table 8. Overall, the cohort loses 69.2% of its overall size in the workforce over this period.

Exit of RNs from the workforce: HWSM model. The HWSM estimates a stable workforce participation rate of approximately 95% through age 50. It is likely the 5% of RNs under age 50 who do not participate in the workforce may have changed their careers. For some of the younger RNs this may reflect a temporary departure from the workforce for raising children. However, from age 50 onwards, the decrease in labor force participation is precipitous and likely reflects retirement (see Figure 1). HWSM estimates also reveal RNs with a graduate degree tend to retire later than RNs with a baccalaureate or associate degree level of education.

Comparison of estimates produced by the Cohort and HWSM. Although the models take different approaches to retirement (or exit from the workforce for other reasons such as becoming an APRN) due to their inherent structures, the estimated number of RNs who will retire through 2025 are similar, as shown in Table 9.

That the estimates generated by both models are so close is not surprising because, ultimately, all RNs will eventually retire and it is only the rate of retirement each year that is uncertain. Yet, it is nevertheless comforting given the differences in the way both models handle retirement.

Total registered nurse workforce supply forecast from the Cohort model and HWSM. Despite different approaches and structures, the two models produce fairly similar workforce forecasts and estimates. Table 10 shows full forecasts of the FTE RN workforce resulting from each model.

The most substantial differences in the approaches taken in each model concerns how each models the new entry of RNs into the nursing workforce, as noted previously. The HWSM model, in relying on the number of NCLEX test takers in the baseline model year, is sensitive to possibilities that year may be atypical or unrepresentative of future trends. It is also possible age distribution of test takers may differ from the year in which age distribution is borrowed, which would impact future sizes of cohorts of new graduates. Similarly, the Cohort model is sensitive to the possibility future entry trends will diverge from those of previous 5 years. There is inherent uncertainty surrounding entry where future cohorts' attraction to nursing as a career must be extrapolated from recent trends, and sharp changes are always possible, unpredictable, and have large and far-reaching impacts on size and age composition of the future workforce. In this case, partly because new entry has been relatively stable in the last several years, the models produce similar estimates of growth in entry.

Of course, there are additional factors that differ between the models not covered in this discussion (e.g., inclusion of internationally educated RNs, which is modeled explicitly in the HWSM and is implicit in the Cohort supply model and reflected in age and cohort estimates). Yet, overall, modeling of workforce entry and exit are the largest drivers of both estimates of supply and changes in supply over time.

Discussion and Recommendations

The results of a comparative analysis of the two leading models used to forecast the future supply of RNs in the United States are described. The aim of the analysis is to elucidate the factors that make a meaningful difference in each model's output so workforce analysts know which factors require close monitoring and

Table 9.
(Comparison of the Cohort and HWSM Estimates of the Retirement of the Baby-Boomer Generation of RNs (RNs Born Between 1946-1954)

Year	Number of Baby Boomers Estimated by Cohort Model	Number of Baby Boomers Estimated by the HWSM*
2012	1,106,936	1,053,610
2013	1,028,938	
2014	1,002,976	
2015	948,571	
2016	900,712	
2017	839,016	
2018	769,935	
2019	713,061	
2020	637,818	
2021	577,735	
2022	515,905	
2023	447,789	
2024	391,509	
2025	340,836	340,347
Percent change (2012-2025)	-69.2%	-67.7%

NOTES: Both models use the same definition of an FTE (40 hours) and exclude APRNs. * Estimates for 2013-2024 are unavailable.
HWSM = Health Workforce Simulation Model, RNs = registered nurses

Table 10.
Total RN Workforce Supply Forecast from the Cohort Model and HWSM, 2012-2025

Year	Total FTE RNs, Cohort Model	Total FTE RNs, HWSM
2012	2,661,041	2,579,381
2013	2,650,664	2,642,596
2014	2,795,167	2,728,071
2015	2,873,979	2,809,094
2016	2,938,096	2,886,556
2017	3,000,601	2,959,566
2018	3,058,993	3,028,124
2019	3,127,453	3,094,010
2020	3,181,539	3,156,336
2021	3,247,218	3,215,100
2022	3,310,850	3,271,193
2023	3,369,623	3,324,614
2024	3,437,999	3,376,255
2025	3,514,456	3,427,006
Percent change (2012-2025)	32.1%	32.9%

SOURCE: Full workforce forecasts from Cohort model and HWSM.
NOTE: Both models use the same definition of an FTE (40 hours) and exclude APRNs.
APRNs = advanced practice registered nurses, FTE = full-time equivalent, HWSM = Health Workforce Simulation Model, RNs = registered nurses

which factors are less important in anticipating changes in the future supply of RNs.

The comparative analysis found differences in the way the two models defined an FTE RN exerted an important impact on the total number of RNs (both at baseline and in projections of future size of the RN workforce) and on the age composition of the workforce. To eliminate variance between the two models arising from use of different FTE definitions, the authors concluded both Cohort supply model and HWSM should adopt the same definition: a 40-hour work week to define an FTE RN. This definition is common in labor economics and will not be affected by temporal changes in RN work patterns.

The Montana team agreed with the logic of the HWSM's exclusion of APRNs when making forecasts of the general nursing workforce. A regression-based strategy was developed to extricate RNs from CPS data from 1979-2000, and from ACS data from 2001-2010, data needed by the Cohort model to develop estimates of age and cohort effects that drive the model. The exclusion of APRNs from the Cohort model led to a decrease of just over 6% (nearly 200,000 RNs) in estimates of number of RNs in 2015.

Adoption of a single definition of an FTE and exclusion of APRNs from supply forecasts eliminates two sources in which the two models differed; consequently, assuming these changes are adopted by both models, analysts should not be concerned with these factors when examining sources of differences when the models estimate future size and age composition of the general RN workforce.

While the two models differ in how they estimate number of RNs who will retire in the future, estimates were very close to each other. Based on this outcome, analysts behind the Cohort supply model and HRSA agreed there would be little gained in changing each model's approach to estimating RN retirements.

In contrast, the area where the two models differ significantly, and where there is the greatest impact on projecting the future supply of RNs, involves the entry of new RNs into the workforce. Estimating the flow into the nursing workforce is critical to ascertaining whether the workforce will grow in the future, particularly when large numbers of RNs born in the Baby-Boom generation are retiring and demand for nurses is expected to increase. The comparative analysis clarified how each model implicitly or explicitly accounted for entry, and identified concerns associated with using the number of NCLEX-RN test takers to estimate future entrants (used by the HWSM) and, similarly, with using the average size of recent cohorts to estimate future entrants (Cohort model). However, there was no difference in opinion among the authors that monitoring changes in RN entry is the single most important factor that affects each model and hence accuracy of its projections. Consequently, going forward, when estimates of future supply of RNs are published based on either model, analysts will be able to focus on how each model handled entry in judging whether the model's output is credible. Along these lines, it would be worthwhile for HRSA to consider recent numbers and trends in the numbers of test takers observed in the NCLEX-RN, and to continue to monitor closely any available data on the number of RNs who leave nursing for careers in other fields or who never ultimately become an RN to refine its attrition assumptions. The new edition of the NSSRN, when data become available, should be helpful in this regard in that it asks RNs when they earned their basic nursing degree, therefore allowing for insight into their career trajectory from that date forward.

While the Montana meeting and ensuing comparative analysis of the Cohort supply model and the HWSM have been beneficial in improving future supply forecasts, there is still much that is not understood about the many factors that impact the nursing workforce directly, indirectly, subtly, and suddenly. Understanding and anticipating these factors is particularly important at the state level where health reforms are likely to be felt most, especially reforms involving insurance coverage and the role of Medicaid. Toward this end, it is recommended HRSA continue to explore any systematic differences in state-based administrative data on RNs and numbers derived from the ACS (e.g., totals, age distributions, and industry) that may help with state-based forecasting and modeling.

Finally, looking ahead to 2025, the nation's nursing workforce will be challenged by the aging of 76 million Baby Boomers, most of whom have multiple chronic diseases, by physician shortages and the uneven geographic distribution of primary care and specialist physicians, by retirement of the RN workforce and the annual loss of millions of years of nursing experience, and by implementation of payment and delivery system reforms (Buerhaus et al., 2017). Given these challenges, efforts should continue to bring workforce modelers and analysts together on a routine basis to monitor and study these changes, develop information to help inform stakeholders on what is happening to the nursing workforce, examine what to expect in the near and long-term future, and offer advice on how employers, educators, and policymakers can ensure an adequately sized and well-prepared RN workforce in the United States. **$**

REFERENCES

Auerbach, D.I., Buerhaus, P.I., & Staiger, D.O. (2007). Better late than never: Workforce supply implications of later entry into nursing. *Health Affairs, 26*(1),178-185.

continued on page 326

Improving Nursing Workforce Forecasts

continued from page 294

Auerbach, D.I., Staiger, D.O., Muench, U., & Buerhaus, P.I. (2013). The nursing workforce in an era of health care reform. *New England Journal of Medicine, 368*(16), 1470-1472.

Auerbach, D.I., Buerhaus, P.I., & Staiger, D.O. (2014). Registered nurses are delaying retirement, a shift that has contributed to recent growth in the nurse workforce. *Health Affairs, 33*(8),1474-1480.

Auerbach, D.I., Buerhaus, P.I., & Staiger, D.O. (2015). Will the RN workforce weather the retirement of the baby boomers? *Medical Care, 53*(10), 850-856.

Buerhaus P.I., Staiger, D.O., & Auerbach, D.I. (2000). Implications of an aging registered nurse workforce. *Journal of the American Medical Association, 283*(22), 2948-2954.

Buerhaus, P., Skinner, L., Auerbach, D., & Staiger, D. (2017). Four challenges facing the nursing workforce in the United States. *Journal of Nursing Regulation, 8*(2), 40-46.

National Council of State Boards of Nursing, Inc. (NCSBN). (2014). *2012 and 2013 nurse licensee volume and NCLEX® examination statistics*. Retrieved from https://www.ncsbn.org/14_2012_2013_NCLEXExamStats_vol61.pdf

National League for Nursing (NCN). (2013). *Annual survey of schools of nursing, fall 2012*. Washington, DC: Author.

Spetz, J. (2015). Too many, too few, or just right? Making sense of conflicting RN supply and demand forecasts. *Nursing Economic$, 33*(3),176-78,185.

Staiger, D.O., Auerbach, D.I., & Buerhaus, P.I. (2009). Comparison of physician workforce estimates and supply projections. *The Journal of the American Medical Association, 302*(15),1674-1680.

U.S. Department of Commerce. Census Bureau. (2017). *The American community survey*. Retrieved from https://www.census.gov/programs-surveys/acs/

U.S. Department of Health and Human Services, Health Resources and Services Administration (DHHS). (2010). *The registered nurse population: Findings from the 2008 national sample survey of registered nurses*. Retrieved from https://bhw.hrsa.gov/sites/default/files/bhw/nchwa/rnsurveyfinal.pdf

U.S. Department of Health and Human Services, Health Resources and Services Administration, National Center for Health Workforce Analyses (DHHS). (2013). *The U.S. nursing workforce: Trends in supply and education*. Rockville, MD: Author.

U.S. Department of Health and Human Services, Health Resources and Services Administration, National Center for Health Workforce Analysis (DHHS). (2015). *Technical documentation for HRSA's health workforce simulation model.* Rockville, MD: Author.

U.S. Health Resources and Services Administration (HRSA). Bureau of the Health Professions, U.S. Department of Health and Human Services, National Center for Workforce Analysis. (2002). *Projected supply, demand, and shortages of registered nurses:* 2000-2020. Washington, DC: Author.