JNM

Journal of Nursing Measurement

SPRINGER PUBLISHING COMPANY

www.springerpub.com/jnm

Journal of Nursing Measurement, Volume 24, Number 1, 2016

Test–Retest Reliability of the Genetics and Genomics in Nursing Practice Survey Instrument

Kathleen A. Calzone, PhD, RN, APNG, FAAN
National Cancer Institute, National Institutes of Health, Maryland

Stacey Culp, PhD
West Virginia University Department of Statistics

Jean Jenkins, PhD, RN, FAAN
National Human Genome Research Institute, National Institutes of Health, Maryland

Sarah Caskey, MS
West Virginia University School of Nursing

Pamela B. Edwards, EdD, MSN, RN-BC, FABC, CNE
Mary Ann Fuchs, DNP, RN, FAAN, NEA-BC
Duke University Health System, North Carolina

Amber Reints, RN-BC
Brita Stange, RN
Janice Questad, RN-BC, BAN
Avera Behavioral Health Center, South Dakota

Laurie Badzek, LLM, JD, MS, RN, FAAN
University of North Carolina Wilmington School of Nursing

Background and Purpose: Assessment of nursing genomic competency is critical given increasing genomic applications to health care. The study aims were to determine the test–retest reliability of the Genetics and Genomics in Nursing Practice Survey (GGNPS), which measures this competency, and to revise the survey accordingly. Methods: Registered nurses (n = 232) working at 2 Magnet-designated hospitals participating in a multiinstitutional genomic competency study completed the GGNPS. Cohen's kappa and weighted kappa were used to measure the agreement of item responses between Time 1 and Time 2. Survey items were revised based on the results. Results: Mean agreement for the instrument was 0.407 (range = 0.150–1.000). Moderate agreement or higher was achieved in 39% of the items. Conclusions: GGNPS test–retest reliability was not optimal, and the instrument was refined based on the study findings. Further testing of the revised instrument is planned to assess the instrument performance.

Keywords: nursing; genetics; genomics; survey; nursing practice; competency

http://dx.doi.org/10.1891/1061-3749.24.1.54

Genomics, defined as all genetic variation coupled with personal, environmental, and lifestyle influences, has significant health implications (Green & Guyer, 2011). Yet, integrating genomics into health care practice presents several challenges, including one pivotal issue: health care provider genomic competency. Low genomic competency affects health care provider capacity to appropriately use genomics in practice. Launching any genomic competency educational initiative requires a detailed understanding of the attitudes and receptivity, confidence, knowledge, practices, and influences of the practice social system to design programs that address specific competency deficits and measure outcomes. The Genetics and Genomics in Nursing Practice Survey (GGNPS) instrument was designed to measure these constructs of genomic nursing competency; however, the instrument is long with varying per question response rates (Calzone et al., 2012). The primary purpose of this article is to provide the results of a study aimed at determining the test–retest reliability of the GGNPS and the resulting instrument refinement.

BACKGROUND AND CONCEPTUAL FRAMEWORK

Nursing is especially challenged when attempting to achieve genomic competency given the enormity of the profession, more than 4,100,000 registered nurses (RNs) as of 2013 of which 82% are actively practicing (Budden, Zhong, Moulton, & Cimiotti, 2013). Varying levels of nursing education and practice further complicate genomic competency achievement. Only baccalaureate and higher degree programs require some genomics as part of the curriculum as defined by the more recent American Association of Colleges of Nursing Essential Series. The baccalaureate and higher degree-specific essentials developed and updated through an iterative consensus process provide expected curricula and continuing education content and competency expectations for each degree level (American Association of Colleges of Nursing, 2014). Currently, 57% of RNs hold degrees less than baccalaureate with 18% diploma and 39% associate degree, levels in which no genomic content is required in the curriculum (Budden et al., 2013). In addition, the profession continues to age, with a mean age of 50 years, up from 44.6 approximately 5 years ago (Budden et al., 2013; Health Resources and Services Administration, 2010). Consequently, most practicing nurses completed their academic preparation before any genomics was required in nursing curricula. As reported in a study of more than 7,000 practicing RNs, more than 50% indicated they had genomics content in their nursing curriculum, however, significant genomic competency deficits persisted indicating that what is now required in nursing curricula may not be adequate (Calzone, Jenkins, Culp, Caskey, & Badzek, 2014). Therefore, assessment of nursing genomic competency to inform curricular and continuing education design and measure program outcomes for this and other genomic competency initiatives remains a critical priority. These baseline data were components of a study of 23 Magnet-designated hospitals across the United States participating in a 12-month intervention led by study trained institutional administrator/educator champion dyads aimed at increasing practicing nurse genomic competency (Calzone et al., 2014).

Rogers's diffusion of innovations (DOI) is the theoretical framework used to guide development of the GGNPS (Rogers, 2003). The DOI has five dimensions: (a) knowledge; (b) persuasion, encompassing innovation advantage, compatibility, complexity, trialability, and observability; (c) adoption decision-making, consisting of adoption or rejection

and, if adopted, whether this continues versus later adoption; (d) implementation; and (e) confirmation. Rogers (2003) explains that the decision about adopting an innovation has multiple influences, including communication channels, time, and the social system, which in this case is the health care system.

PROCEDURES FOR INSTRUMENT DEVELOPMENT

The GGNPS was derived from an instrument designed specifically to assess family physicians (Jenkins, Woolford, Stevens, Kahn, & McBride, 2010). The family physician (FP) instrument, intended to be administered online, was constructed to measure essentials of Rogers's DOI (Rogers, 2003). The FP instrument development team was interdisciplinary, consisting of family physicians, a behavioral scientist, online survey designers, and genetic/genomic experts. Items developed by the team underwent evaluation for content validity by external content experts. A pilot survey using a random sample of family physicians who were members of the American Academy of Family Physicians was followed by a factor analysis of the resulting data, which led to instrument refinement. The subsequent FP instrument measures the following DOI domains: attitudes, receptivity, confidence, knowledge and competency, and decision on adoption, as well as the social system that influences those domains. Family history use is used as a measure of adoption because it is the most widely available genetic test and is not cost- or technology-dependent (Berg et al., 2009). The FP instrument was then validated using structural equation modeling (SEM) to ascertain whether the items aligned with the DOI domains and the direction of the association. SEM revealed the following:

- Comparative fit index (CFI = .93, >.9)
- Tucker-Lewis index (TLI = .92, >.9)
- Root mean square error of approximation (RMSEA = .039; 90% confidence interval, <.08)
- Standardized root mean square residual (SRMR = .056, <.1).

These findings indicated that the FP survey items adequately aligned with the DOI domains. All standardized path weights were statistically significant ($p < .001$; Jenkins et al., 2010).

The GGNPS is a discipline-specific derivative of this validated FP survey instrument. To level the items for nursing practice, two different types of experts were used; nursing experts and genomic nursing experts. All items were reviewed for content validity, applicability to nursing's scope of practice, consistency with the genomic competencies for nurses (Consensus Panel on Genetic/Genomic Nursing Competencies, 2006, 2009), and leveling for nursing practice at entry level. Initial usability pilot testing of the nursing specific instrument was performed by a convenience sample of nurses representing the target population (Health Resources and Services Administration, 2010). This was followed by a larger pilot test of the instrument in a representative nursing population (Calzone et al., 2012). Data from this larger pilot study were used to further refine the instrument. One identified weakness included a lack of detailed knowledge assessment. To address this deficit, two questions assessing detailed knowledge of the genomics of common disease were added with permission from the Genomic Variation Knowledge Assessment Index (Bonham, Sellers, & Woolford, 2014). The resulting survey served as the final GGNPS instrument used in this reliability study.

DESCRIPTION, ADMINISTRATION, AND SCORING OF THE INSTRUMENT

The GGNPS includes eight sections: (a) broad attitudes on genomics, (b) confidence and family history specific attitudes, (c) adoption of family history restricted to those actively seeing patients, (d) general genomic knowledge, (e) detailed genomic knowledge, (f) personal genomic competency assessment, (g) social system, and (h) demographics. Instrument question types include select all that apply pick lists, multiple choice, yes/no, true/false, and Likert scale (Calzone et al., 2012).

The GGNPS can be administered online or in paper format. Both administration methods have been used in studies using this instrument (Badzek, Calzone, Jenkins, Culp, & Bonham, 2013; Calzone, Jenkins, Culp, Bonham, & Badzek, 2013; Calzone et al., 2014).

Items corresponding to the attitudes, receptivity, confidence, and adoption domains, as well as the social system influence items, are analyzed individually and are not combined to form scores. The responses to 12 items measuring genomic knowledge are combined to form a knowledge score. Reponses to each of the 12 items are first graded as correct or incorrect, and a total knowledge score is calculated as the number of correct responses out of 12, with a minimum possible score of 0 and a maximum possible score of 12. Calculation of the total knowledge score is restricted to individuals responding to all 12 items.

METHOD

Sample

The test–retest sample consisted of RNs working at one of two American Nurses Credentialing Center Magnet-designated hospitals participating in a multiinstitutional genomic nursing competency study, A *M*ethod for *I*ntroducing a *N*ew *C*ompetency into Nursing Practice (MINC). MINC was a longitudinal study of 23 Magnet hospitals consisting of 2 hospitals as controls and the remaining 21 in the intervention arm. This study was aimed at increasing institutional genomic nursing competency. MINC used the GGNPS as one of the outcome measures.

Procedures

The GGNPS was administered at Time 1 (test) and Time 2 (retest) to RNs employed at the two participating MINC hospitals at the conclusion of the year-long MINC study intervention period. Administering the test–retest at the conclusion of MINC was done to minimize guessing at question answers which could influence reliability. The Time 1 survey was open for completion for 28 days 1 year after both institutions' baseline surveys were completed as part of the parent MINC study. To achieve the 1-year interval from the baseline survey administration, Institution 2 opened the Time 1 survey 4 days following Institution 1. One week from the close of the Time 1 survey, the Time 2 survey was opened for 14 days. Survey recruitment was performed via e-mail notification at both time points consisting of a minimum of three e-mails for both Time 1 and 2; baseline, midpoint in survey collection period and 2 days prior to survey closure.

RNs were required to enter a hospital specific personal study number, generally used for all other institutional surveys. The study number was used to match Time 1 and Time 2 results. A list of valid study numbers from each institution was provided to the

investigators. Upon completion of the Time 2 survey period, Time 1 and Time 2 reliability data sets were merged based only on study number. If a study number was duplicated at a single time point, matches were validated using gender and year of birth variables.

Eligibility. To be eligible for participation in the study, employment as an RN at one of the two participating Magnet hospitals was required. RNs at all levels of academic preparation, clinical roles, and specialties were eligible to participate.

Exclusion Criteria. Licensed practical nurses and non-RNs were excluded because the GGNPS was leveled for registered nursing practice. Additional exclusions after the Time 1 and Time 2 data merger consisted of the following:

- Study number missing
- Invalid study number
- All Time 1 data missing
- All Time 2 data missing
- Duplicate use of study numbers that could not be resolved through matching on gender and year of birth

Regulatory Approval

The Institutional Review Board (IRB) of West Virginia University (WVU) approved the study. A reliance agreement between WVU and the Office of Human Subjects Research at the National Institutes of Health was also established. Both of the participating institutions received approval to rely on the WVU IRB approval.

Data Analysis for Reliability Assessment

Cohen's kappa (κ) was used to measure the agreement of item responses between Time 1 and Time 2 (Cohen, 1960). For questions measured on an ordinal scale, weighted kappa (κ_w) values were calculated (Cohen, 1968). Landis and Koch (1977) provides the framework used for interpreting the kappa values. The irr package (Gamer, Lemon, Fellows, & Singh, 2012) in R (R Core Team, 2013) was used to calculate the weighted kappa values. SPSS Version 21 was used for all other analysis (IBM, 2012). A power analysis showed that a sample size of 85 was required to detect large effects with power equal to .95 after adjusting for multiple testing.

RESULTS

Reliability Sample for Assessment

There were 993 RNs who completed the GGNPS at Time 1 and 564 at Time 2. After accounting for the exclusions described earlier, consisting of 296 with all Time 1 data missing, 664 with all Time 2 data missing, 23 with invalid study numbers, and 6 duplicate study numbers that could not be resolved through matching on gender and year of birth, there were 232 cases for analysis. This consisted of 11 from Institution 1 and 221 from Institution 2.

Reliability

Table 1 summarizes all the instrument items and associated kappa and weighted kappa coefficients. The agreement mean across all instrument items was 0.407 (range = 0.150–1.000).

TABLE 1. Instrument Item Kappa Coefficients

Item Number	Item Format	Question Text	Kappa	Weighted Kappa
P1-1	Likert scale	How important do you think it is for the nurse to become more educated about the genetics of common diseases?		0.415
P1-2	Select all that apply pick list	To what extent do you think each of the following would be a potential advantage of integrating genetics of common diseases into your practice? (Check all that apply)	0.249	
		Advantage—better decisions about recommendations for preventative services		
		Advantage—better treatment decisions	0.302	
		Advantage—improved services to patients	0.326	
		Advantage—better adherence to clinical recommendations	0.314	
		Advantage—genetic risk triaging could make better use of visit time	0.249	
		What potential disadvantages do you see for integrating genetics of common diseases into your practice? (Check all that apply)	0.442	
		Disadvantage—would take too much time		
		Disadvantage—not reimbursable/too costly	0.316	
		Disadvantage—need to "retool" professionally	0.383	
		Disadvantage—would not improve patient care	0.165	
		Disadvantage—increase patient anxiety about risk	0.397	
		Disadvantage—would increase insurance discrimination	0.432	
		Disadvantage—medical-legal problems for nurses related to testing	0.231	
		Disadvantage—greater burden of responsibility on nurses	0.304	
		Disadvantage—need to educate nurses on genetics	0.212	
		Disadvantage—no disadvantages	0.301	

(Continued)

TABLE 1. Instrument Item Kappa Coefficients *(Continued)*

Item Number	Item Format	Question Text	Kappa	Weighted Kappa
P2-1	Likert scale	Decide what family history information is needed to tell something about a patient's genetic susceptibility to common diseases.		0.349
		Use family history information to support treatment decisions for patients who have common diseases.		0.353
		Counsel patients about inherited risk for common diseases.		0.382
		Assure that patient's family history information relating to common diseases remains confidential.		0.409
		Provide emotional follow-up support after family history assessment.		0.426
		Discuss how family history affects recommended screening intervals.		0.404
		Decide which patients would benefit from a referral for genetic counseling and possible testing for susceptibility to common diseases.		0.433
		Access reliable and current information about genetics and common diseases.		0.366
		Provide information about the AVAILABILITY of genetic testing for common diseases.		0.403
		Give patients information about the RISKS of genetic testing for common diseases.		0.424
		Give patients information about the BENEFITS of genetic testing for common diseases.		0.415
		Give patients information about the LIMITATIONS of genetic testing for common diseases.		0.426
		Facilitate referrals for genetic services for common diseases.		0.489
		Provide emotional follow-up support after genetic testing.		0.455

(Continued)

TABLE 1. Instrument Item Kappa Coefficients *(Continued)*

Item Number	Item Format	Question Text	Kappa	Weighted Kappa
P2-2	Select all that apply pick list	To what extent do each of the following limit your ability to discuss the genetics of common diseases with individuals? (Choose all that apply)	0.209	
		Limitation—my difficulty finding info on genetics and common diseases		
		Limitation—lack of time to get complete family history	0.291	
		Limitation—no reimbursement for counseling or testing	0.394	
		Limitation—no place to send patients for counseling or testing	0.150	
		Limitation—increased nursing liability	0.307	
		Limitation—not in my scope of practice	0.224	
		Limitation—my limited expertise in genetics	0.227	
		Limitation—lack of use of genetics in my clinical area	0.293	
P2-3	Likert scale	A family history that includes only 1st degree relatives such as parents, siblings, and children should be taken for every new patient.		0.265
		A family history that includes 2nd and 3rd degree relatives such as grandparents, aunts, uncles, and cousins should be taken for every new patient.		0.372
		Family history taking should be a key component of nursing care.		0.392
		There is a role for nurses in counseling patients about genetic risks.		0.393
P3-1	Dichotomous yes/no	Are you actively seeing patients?	0.753	
P3-2	Multiple choice	In the past 3 months, how often have you collected a complete family history from a patient that includes the following components: information on disorders from three generations, and age at diagnosis and death for each affected family member?		0.592
P3-3	Dichotomous yes/no	In the past 3 months, has any patient initiated a discussion with you about genetics?	0.441	

(Continued)

TABLE 1. Instrument Item Kappa Coefficients *(Continued)*

Item Number	Item Format	Question Text	Kappa	Weighted Kappa
P3-5	Likert scale	In the past 3 months, how often have you used family history information when facilitating clinical decisions or recommendations for your patients?		0.439
		In the past 3 months, how often have you facilitated referrals to genetic services?		0.472
P4-1	Likert scale	Do you think that genetic risk (e.g., as indicated by family history) has clinical relevance for the following?		
		Breast cancer		0.308
		Colon cancer		0.475
		Coronary heart disease		0.404
		Diabetes		0.383
		Lung cancer		0.493
		Ovarian cancer		0.427
P4-2	Likert scale	When patients indicate a disorder in the family, which of the following piece of information do you collect in your standard family history assessment?		
		Age at diagnosis of condition		0.397
		Relationship to the patient		0.541
		Race or ethnic background		0.302
		Age at death from condition		0.397
		Both sides of the family (maternal/paternal)		0.535
P4-3	Multiple choice	Which of the following best describes how you collect family history information?	0.254	
P4-4	Likert scale	Thinking about how you support clinical decisions (such as administering drugs prescribed), to what extent do you think each of the following is important to consider?		
		Gender		0.375
		Race/ethnicity		0.327
		Genes		0.375
		Family history		0.306
		Age		0.216
		Insurance status		0.342
P5-1	True/false	The DNA sequences of two randomly selected healthy individuals of the same sex are 90%–95% identical.	0.565	

(Continued)

TABLE 1. Instrument Item Kappa Coefficients *(Continued)*

Item Number	Item Format	Question Text	Kappa	Weighted Kappa
P5-2	True/false	Most common diseases such as diabetes and heart disease are caused by a single gene variant.	0.409	
P6-1	Dichotomous yes/no	The Essential Competencies and Curricula Guidelines for Nurses in Genetics and Genomics are endorsed as being a standard part of nursing practice. Have you heard or read about these Competencies?	0.520	
P6-2	Likert scale	Rate your understanding of the genetics of common diseases.		0.604
		In describing your genetic/genomic knowledge, would you consider it to be		0.597
P7-1	Dichotomous yes/no	Did your nursing curriculum include genetics content?	0.496	
		Since licensure, have you attended any courses that included genetics as a major component?	0.583	
		Do you intend to learn more about genetics?	0.524	
		Would you be able to attend a course during work hours?	0.480	
		Would you attend a course on your own time?	0.476	
		Do you think your senior staff members see genetics as an important part of YOUR role?	0.566	
		Do you think your senior staff members see genetics as an important part of THEIR role?	0.511	
P7-2	Likert scale	Books and printed information for self study		0.397
		Electronic and web-based information for self study		0.468
		A mix of both		0.290
		A mix of self study and face-to-face meetings		0.446
		Workshops (mixture of presentations and group activities)		0.418
		Lectures		0.398
P8-1	Numeric text	What year were you born?	0.993	
P8-2	Multiple choice	What is your gender?	1.000	

Moderate agreement or higher as defined by Landis and Koch (1977) was achieved in 39% (n = 36/95) of the instrument individual items.

Select all that apply pick lists performed the poorest, mean $\kappa = 0.308$ (range = 0.165–0.432) for attitude questions, and mean $\kappa = 0.262$ (range = 0.150–0.394) for personal assessment of ability to discuss genetics of common diseases with individuals. Of the Likert scale questions, the items assessing importance of genetics and other factors on medication administration have seven options, 1 = *not at all* through, 7 = *essential*, and these items performed fairly, mean $\kappa_w = 0.324$ (range = 0.216–0.375). Likert scale items with fewer options (i.e., four, three) performed slightly better. In addition, Likert scale questions focused on personal assessment of genomic competency (i.e., Table 1, P6-2: $\kappa_w = 0.604$ and $\kappa_w = 0.597$) outperformed similarly formatted questions requiring some level of genomics competency (i.e., Table 1, P4-1: κ_w range = 0.308–0.493). The 12 items included in the total knowledge score ranged from a low of $\kappa_w = 0.265$ for one question about the scope of family history collection to a high of $\kappa = 0.565$ for a specific knowledge true/false question.

With the exception of demographic questions, no items achieved almost perfect agreement. One dichotomous yes/no question, P3-1, asked, "Are you actively seeing patients*?*" and achieved substantial agreement ($\kappa = 0.753$) in this study. However, even this value was lower than expected. Review of this question with nurses representative of the population studied indicated that the term *seeing patients* is sometimes interpreted as having scheduled visits with patients in an outpatient clinic and could be confusing for those who see patients in an acute care or inpatient setting.

To revise the instrument, an instrument modification team was convened which included the study investigators and GGNPS developers. Kappa values were reviewed for each item. The weakest items measuring a domain in which there was another better performing item were eliminated from the instrument. Poorly performing question formats, including the select all that apply pick lists and large Likert scale questions, were targeted for reformatting. Pick lists were reduced to the best performing items and reformatted to dichotomous questions (e.g., P1-2, advantage/no advantage). Large Likert scales underwent scale reduction to no more than three selections consisting of one middle anchor and a positive and negative response. All 12 items used to compute the knowledge score were retained, but the Likert scale questions underwent scale reduction. However, this scale reduction does not impact the ability to score each item as correct or incorrect and thus will not alter the calculation of the knowledge score. Last, P3-1 ("Are you actively seeing patients?") was revised to address the terminology confusion. The revised instrument, which continues to be open access, is accessible at http://www.genome.gov/27527636 under Research Tools.

DISCUSSION

These data represent the first reliability testing of the GGNPS, an instrument designed to measure the domains of Rogers's DOI as applied to nursing and genomics. The findings demonstrate that the mean agreement for all items was moderate agreement (mean κ = 0.41, range = 0.150–1.000), and less than half (39%) of the items achieved moderate or greater agreement. Building on the psychometric validity of the instrument and alignment to Rogers's DOI as measured by SEM, modifications to the instrument to increase reliability are worthwhile because no other similar instrument currently exists.

The poorest performing questions included Likert scale questions with five or more options. These scales included positive and negative endpoints with a corresponding definition and a middle anchor. In general, Likert scales have been shown to be reliable in test–retest reliability evaluation, including scales with greater than five response options (Weng, 2004). In this study, similarly designed Likert scale questions performed poorer if genetic knowledge was needed when responding in contrast to questions limited to self-assessment. For example, P6-2 ("Rate your understanding of the genetics of common diseases"), a Likert scale item with five response options, was at the highest end of moderate agreement with $\kappa_w = 0.604$. In contrast, a knowledge-based question with a five-response option Likert scale, P2-3 ("A family history that includes 2nd and 3rd degree relatives such as grandparents, aunts, uncles, and cousins should be taken for every new patient"), attained only fair agreement ($\kappa_w = 0.372$). The study design had attempted to minimize this effect by conducting the test–retest study at the conclusion of the parent MINC intervention when genomic knowledge would have been expected to be the greatest. However, preliminary MINC findings reveal that the intervention predominately increased awareness, relevance, and attitudes about genomics, the first steps in DOI, but significant knowledge deficits persist and additional intervention was still needed. The degree to which these knowledge deficits contributed to lower kappa coefficients because of guessing when responding to questions is unknown. Carefully selecting the population for the next reliability analysis is needed to further minimize this potential impact; one possibility would be to study nurses who have completed a more robust genomics education intervention.

The GGNPS as a self-report assessment has all the inherent limitations of such an approach. This instrument evaluation highlights the need to establish a more reliable, objective measure of nursing genomic competency. The National Database of Nursing Quality Indicators (NDNQI) established by the American Nurses Association (ANA) in 1998 consists of a standardized list of indicators used to collect nursing unit-level structure, process, and outcome data to evaluate and build on nursing's body of knowledge and demonstrate the value of nurses to improving the quality and safety of patient care (Montalvo, 2007). With the establishment of the NDNQI, the convention has been to measure nursing quality through safety outcomes such as nursing hours per patient day (structure), falls or falls with injury (process and outcome), pressure ulcer prevalence (process and outcome), and nosocomial infections (outcome; Izumi, 2012; Montalvo, 2007). Safety outcomes represent the minimum standards for nursing quality and do not necessarily rely on nurses' expertise, professional judgment, and clinical reasoning to provide quality of care (Izumi, 2012). As nursing continues to define what quality nursing care is and how can it be measured, nursing sensitive indicators will need to focus on the values and knowledge underlying practice, the process of how the health care is delivered, and health care outcomes, not just the safety outcomes (Izumi, 2012). In the area of genomics, the *Essentials* define what the nurse (all academic levels, roles, and clinical specialties) is required to know about genomics to achieve competency (Consensus Panel on Genetic/Genomic Nursing Competencies, 2009). Nursing sensitive quality indicators with a focus on genomics should be designed to evaluate nurses' use of professional judgment and clinical reasoning, for example, prompting nurses to integrate family history knowledge into their clinical assessment, which will allow for better assessment, planning, education, and referral of patients and their families. Quality indicators would therefore enable objective competency assessment, serve as a mechanism for measuring competency attainment, and/or indicate the need for further remediation thereby improving care quality and safety.

Nursing Implications

The translation of genomic discoveries into patient care represents one of the most promising developments in health care in decades. Achieving optimal health benefits from the use of genomic information depends largely on health care provider genomic competency. And yet, most practicing nurses and other health care providers have little to no education in this science or its clinical applicability. This presents a daunting challenge for all health care disciplines considering that scientific discoveries are occurring rapidly with decreasing costs and the health implications span the entire health care continuum encompassing almost all aspects of health care (Calzone, Jenkins, Nicol, et al., 2013). Furthermore, the gap in competency includes lack of recognition of the relevancy of genomics to nursing practice which may impact the uptake of continuing education in this area in the absence of a regulatory requirement. To overcome these challenges, robust interventions are needed with reliable measures that can adequately assess the outcomes of these strategies until such time as objective measures of competency in practice can be implemented. Achieving genomic competency is an interprofessional challenge that nursing has led by being the first health care profession to establish competencies applicable to the entire profession. With reliable measures to inform interventions needed and outcomes associated with their implementation, nursing will continue to serve as the model for the interprofessional health care community on how to achieve genomic competency with the resulting health benefits associated with the appropriate use of genomic information.

Limitations

The participants in this study were recruited from Magnet hospitals participating in a larger genomic competency research project (MINC); therefore, they may not be representative of the general population of nurses. Considerable attrition occurred from Time 1 to Time 2. Despite more than 990 nurses completing the survey at Time 1, 664 did not complete Time 2, and after accounting for exclusion criteria, only 232 were eligible for this analysis, which is less than 25% of the original Time 1 sample. Although the sample size still exceeded the minimum participant number required for this analysis, attrition bias may have influenced the strength of our reliability assessment. In addition, the study was weighted heavily toward Institution 2 because only 11 matched cases were available from Institution 1 which also contributes potential bias. Finally, the study recruitment strategy, which targeted an entire institution's registered nursing population, resulted in differing numbers of days between the completion of the Time 1 and Time 2 surveys among participants. On average, the time between survey responses was about 24 days, with a minimum of 8 days and a maximum of 44 days.

CONCLUSION

Genomics is a science that has substantial implications for health care applications, but most health care providers including nurses have little to no educational preparation or knowledge needed for genomic competency which can hinder use in practice (Calzone, Jenkins, Nicol, et al., 2013). The GGNPS is designed to assess domains of Rogers's DOI but test–retest reliability is not optimal. Refinement of the instrument was completed based on this test–retest reliability analysis. Evaluation of the modified instrument is needed to

assess whether revisions made as a result of this study have improved the instrument's performance. Further work is needed to develop observable, reliable measures that focus on outcomes representative of underlying genomic competency that do not have self-report limitations.

REFERENCES

American Association of Colleges of Nursing. (2014). *Essentials series*. Retrieved from http://www.aacn.nche.edu/education-resources/essential-series

Badzek, L., Calzone, K. A., Jenkins, J., Culp, S., & Bonham, V. L., Jr. (2013). Nursing leaders hold the key to translating genomics into practice. *American Nurse Today*, *8*(12). Retrieved from http://www.americannursetoday.com/article.aspx?id=11030&fid=10986

Berg, A. O., Baird, M. A., Botkin, J. R., Driscoll, D. A., Fishman, P. A., Guarino, P. D., . . . Williams, J. K. (2009). National Institutes of Health State-of-the-Science Conference statement: Family History and improving health. *Annals of Internal Medicine*, *151*(12), 872–877.

Bonham, V. L., Jr., Sellers, S. L., & Woolford, S. (2014). Physicians' knowledge, beliefs, and use of race and human genetic variation: New measures and insights. *BMC Health Services Research*, *14*, 456.

Budden, J. S., Zhong, E. H., Moulton, P., & Cimiotti, J. P. (2013). Highlights of the National Workforce Survey of Registered Nurses. *Journal of Nursing Regulation*, *4*(2), 5–14.

Calzone, K. A., Jenkins, J., Culp, S., Bonham, V. L., Jr., & Badzek, L. (2013). National nursing workforce survey of nursing attitudes, knowledge and practice in genomics. *Personalized Medicine*, *10*(7), 719–728.

Calzone, K. A., Jenkins, J., Culp, S., Caskey, S., & Badzek, L. (2014). Introducing a new competency into nursing practice. *Journal of Nursing Regulation*, *5*(1), 40–47.

Calzone, K. A., Jenkins, J., Nicol, N., Skirton, H., Feero, W. G., & Green, E. D. (2013). Relevance of genomics to healthcare and nursing practice. *Journal of Nursing Scholarship*, *45*(1), 1–2.

Calzone, K. A., Jenkins, J., Yates, J., Cusack, G., Wallen, G. R., Liewehr, D. J., . . . McBride, C. (2012). Survey of nursing integration of genomics into nursing practice. *Journal of Nursing Scholarship*, *44*(4), 428–436.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, *70*(4), 213–220.

Consensus Panel on Genetic/Genomic Nursing Competencies. (2006). *Essential nursing competencies and curricula guidelines for genetics and genomics*. Silver Spring, MD: American Nurses Association.

Consensus Panel on Genetic/Genomic Nursing Competencies. (2009). *Essentials of genetic and genomic nursing: Competencies, curricula guidelines, and outcome indicators* (2nd ed.). Silver Spring, MD: American Nurses Association.

Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2012). *irr: Various coefficients of interrater reliability and agreement*. Retrieved from http://CRAN.R-project.org/package=irr

Green, E. D., & Guyer, M. S. (2011). Charting a course for genomic medicine from base pairs to bedside. *Nature*, *470*(7333), 204–213.

Health Resources and Services Administration. (2010). *The registered nurse population: Findings from the 2008 National Sample Survey of Registered Nurses*. Washington, DC: U.S. Department of Health and Human Services. Retrieved from http://bhpr.hrsa.gov/healthworkforce/rnsurveys/rnsurveyfinal.pdf

IBM. (2012). *GPL reference guide for IBM SPSS statistics*. Retrieved from http://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/21.0/en/client/Manuals/GPL_Reference_Guide_for_IBM_SPSS_Statistics.pdf

Izumi, S. (2012). Quality improvement in nursing: Administrative mandate or professional responsibility. *Nursing Forum*, *47*(4), 260–267.

Jenkins, J., Woolford, S., Stevens, N., Kahn, N., & McBride, C. M. (2010). Family physicians' likely adoption of genomic-related innovations. *Case Studies in Business, Industry and Government Statistics*, *3*(2). Retrieved from http://www.bentley.edu/csbigs

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Montalvo, I. (2007). *The National Database of Nursing Quality Indicators® (NDNQI®)*. Retrieved from http://www.nursingworld.org/MainMenuCategories/ANAMarketplace/ANAPeriodicals/OJIN/TableofContents/Volume122007/No3Sept07/NursingQualityIndicators.aspx?%3E

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Rogers, E. (2003). *Diffusion of innovations* (5th ed.). New York, NY: Free Press.

Weng, L.-J. (2004). Impact of the number of response categories and anchor labels on coefficient alpha and test-retest reliability. *Educational and Psychological Measurement*, *64*(6), 956–972.

Acknowledgments. This research was funded through a grant from the National Council of State Boards of Nursing and supported by West Virginia University and the Intramural Research Programs of the National Institutes of Health, National Cancer Institute, and National Human Genome Research Institute.

Correspondence regarding this article should be directed to Kathleen A. Calzone, PhD, RN, APNG, FAAN, National Institutes of Health, National Cancer Institute, Center for Cancer Research, Genetics Branch, 37 Convent Drive, Building 37, RM 3039, MSC 4256, Bethesda, MD 20892. E-mail: calzonek@mail.nih.gov