Winter 12-15-2017

Baseline Prevalence Study of Hendrich II Fall Risk Assessment Tool at a Local Community Hospital

Cyrus Rivera
crivera@myself.com

Baseline Prevalence Study of Hendrich II Fall Risk Assessment Tool at a Local Community Hospital

Cyrus A. Rivera, MSN, RN, EMT

University of San Francisco

Abstract

A baseline prevalence study on Hendrich II Fall Risk Assessment Tool was conducted per the request of the fall committee of a community hospital in the San Francisco Bay Area. Hendrich II assessments were completed on 106 patients (n = 106) between September 2017 and November 2017. In addition, assessments by registered nurses on those same patients were obtained via retrospective chart review. Analysis via two-sample Z-test revealed students and nurses scored the dizziness/vertigo part of Hendrich II differently ($p = 5.51^{-10}$). Similarly, students and nurses scored the Get-Up-And-Go part differently as well ($p = 0.05$). A literature review of other fall risk assessment tools was also conducted to identify alternatives to Hendrich II. No alternative can be recommended because results on sensitivity & specificity for any given tool are mixed. If an alternative is desired, it must be quicker to complete than Hendrich II. Otherwise, nurses may not be inclined to complete parts such as dizziness/vertigo and Get-Up-And-Go accurately because of the excess time they take away from higher priority tasks.

Keywords: baseline prevalence study, fall risk assessment tool, sensitivity, specificity, Hendrich II Fall Risk Model, Morse Fall Scale, St. Thomas's Risk Assessment Tool in Falling Elderly Inpatients (STRATIFY), Innes, Johns Hopkins Fall Risk Assessment Tool, Schmid Fall Risk Assessment Tool

Introduction

Patient falls have become endemic in practically all hospitals throughout the US. Falls are defined as unwanted contact with the ground that occurs in inpatient units and may or may not result in injury. According to Bouldin et al. (2013), the fall rate is as high as 4.03 per 1,000 patient days (pd) in medical units across all US hospitals. Given that this data was collected between July 2006 and September 2008, it is also probable that that rate is now higher when considering that elderly persons continue to make up a larger proportion of the total population.

Reducing this fall rate continues to be a primary aim of many hospitals because falls lead to adverse patient outcomes and are also very costly. According to the Center for Disease Control (CDC, 2017), one in every five falls results in significant injury. The best example is head injury, which in severe cases is fatal. In 2015, total annual cost of falls was $31 billion and cost per fall was $30,000. These expenditures cover in-hospital care, rehabilitation, and medications, to name a few. The psychosocial well being of patients may also be adversely affected. For example, patients with histories of previous falls may be less inclined to participate in various activities out of fear of falling again.

Objectives

The fall committee of a community hospital in the San Francisco Bay Area had two objectives. First, to identify better alternatives to Hendrich II by reviewing the most recent literature on fall risk assessment tools. Hendrich II is the tool currently in use at this hospital, and the committee expressed dissatisfaction with it. The second objective was to identify how registered nurses use Hendrich II tool.

Methods

For the initial review, a literature review of six peer-reviewed research articles comprising six different fall risk assessments tools used in acute care settings was undertaken. Articles were qualitative and quantitative in nature, including systematic reviews with meta-analysis, pilot studies with randomized control trials, retrospective observational studies, and cross-sectional studies. Databases utilized were Cumulative Index to Nursing and Allied Health Literature (CINAHL) and PubMed. Example search terms were "fall risk assessment tools" and "review." With few exceptions, articles were restricted to those published in the past five years in order to ensure the most up-to-date information.

The second objective entailed four steps. In step one, Hendrich II assessments were completed on a total of 106 patients (n = 106) from six different units between September 2017 and November 2017. Units represented were medical-surgical, progressive care, telemetry/stroke, and critical care. Assessments with the hospital's designed Patient Mobility Assessment Tool (PMAT) were also completed at the same time. Step two was to identify how registered nurses assessed those same patients on Hendrich II and PMAT via retrospective chart review. In step three, all data were pooled and analyzed via statistical analysis, two-sample Z-test in particular. In the fourth and final step, three themes from the analysis were identified and are discussed later alongside findings of the literature review.

Literature Review

To date, the literature on sensitivity & specificity of fall risk assessment tools is very limited. The offering is even lower when filtering out tools used in outpatient and/or community-dwelling settings. Only tools used for adult patients in in-hospital/acute care settings similar to the hospital at the center of these objectives were sought. High heterogeneity +90% is another challenge, determined by Park (2017) via a systematic review with meta-analysis of 33 studies. Thus, what might be useful in one hospital might be marginally effective in a different hospital because patient populations and/or study settings can be very different. For instance, a tool that has high sensitivity & specificity with patients undergoing orthopedic surgeries will not necessarily be effective for patients with infections because admit diagnoses are different. Given these limitations, it is imperative to be cautious when inferring the utility of any tool.

Oliver, Daly, Martin, & McMurdo (2004) provide a strong basis nonetheless. In a systematic review with meta-analysis of 13 studies, six individual risk factors consistently had high sensitivity & specificity +70%, regardless the tool. Those were gait instability, lower limb weakness, urinary incontinence/need for assistance to toilet, past history of falls, agitation/confusion/impaired judgment, and medications. This contrasts with other studies and reviews that examine tools in their entirety. Moving forward, this review assesses tools in their entirety and the risk factors they comprise alongside these six individual risk factors. The greater the number of these risk factors comprising any given tool, the more likely that tool will have high sensitivity & specificity in its entirety.

Of the six tools reviewed, the literature on Innes, Johns Hopkins, and Schmid is especially limited. Thus, these are not recommended for pilot testing. Each is reviewed nonetheless to be comprehensive, beginning with Innes, which is perhaps the most limited in terms of amount of literature available. Innes assesses seven risk factors, yet no information on how to score each risk factor or score range could be found. At best, Oliver et al. (2004) listed those risk factors. Those are trauma, disorientation, impaired judgment, sensory disorientation, muscle weakness, multiple diagnoses, and language barrier. Innes in its entirety had high sensitivity 89.3% and high specificity 73.5%. Paradoxically however, only three of its seven risk factors had high sensitivity & specificity per Oliver et al. (2004). Those are disorientation, impaired judgment, and muscle weakness.

Like Innes, John Hopkins also assesses seven risk factors. Those are age, fall history, elimination, medications, patient care equipment, mobility, and cognition. Scoring is cumbersome mainly because each risk factor comprises line items with varying points. With age for instance, 80+ years old warrants 3 points, 70 - 79 years old warrants 2 points, and 60 - 69 years old warrants 1 point. Medications by comparison are more nuanced and warrant as much as 7 points if sedating agents were administered within the past 24 hours. Total score of 13+ is high risk while a score between 6 - 13 is moderate risk. In it entirety, Johns Hopkins had low sensitivity 27.1% yet high specificity 89.6%. This was determined via retrospective chart review of 13,574 patients (*n* = 13,574) at a large academic medical center in the Midwest (Klinkenberg & Potter, 2017). In other words, Johns Hopkins assesses low to moderate risk more accurately than high risk. Paradoxically however, five of its seven risk factors have high sensitivity & specificity per Oliver et al. (2004). Those are fall history, elimination, medications, mobility, and cognition. Of note, Johns Hopkins is unique in that it is one of few tools that suggests fall prevention interventions after a score is determined. For instance, it might advise that a gait belt be placed at the bedside of a patient who is moderate risk.

Schmid is the last tool reviewed that is not recommended for testing. It assesses five risk factors. Those are mobility, mentation, elimination, prior fall history, and current medications. Scoring each line item in any of its risk factors is easier than in Johns Hopkins because each is worth just 1 point. With mentation for instance, confusion warrants 1 point regardless if periodic or persistent. Prior fall history is the only exception in which 2 points may be warranted if a patient fell during admission. A total score of 3+ is high risk. In its entirety, Schmid had high sensitivity 92.5% and high specificity 78.2%. Moreover, all five of its risk factors had high sensitivity & specificity as well (Oliver et al., 2004). Such strong marks would make it an intriguing prospect for testing were it not for the limited amount of literature.

Hendrich II, Morse, and STRATIFY are three tools reviewed substantially through the literature. Thus, either of these is recommended for testing depending on which is to the liking of the committee. Hendrich II assesses seven risk factors. Those are confusion/disorientation/impulsivity, symptomatic depression, altered elimination, dizziness/vertigo, male gender, antiepileptics, and benzodiazepines. It is also unique in that it comprises a functional component requiring patients to physically get up, hence the name of the subtest "Get-Up-And-Go." While number of points varies with each risk factor, scoring is straightforward because those points are fixed. Confusion/disorientation/impulsivity for instance warrants only 4 points. Benzodiazepines warrant only 1 point by comparison. A total score of 5+ is high risk. Results on sensitivity & specificity are mixed. In Park (2017), Hendrich II had high sensitivity 76% yet low specificity 60%. A separate systematic review of 8 studies (*n* = 8) comprising 10,479 patients across several acute care hospitals had similar results, high sensitivity 70% yet low specificity 61.5% (Callis, 2016). Then in a separate study on 1,815 patients (*n* = 1,815) in an acute care hospital in Lebanon, the opposite was true for Hendrich II, low sensitivity 55% yet high specificity 90% (Nassar & Madi, 2014). Per Oliver et al. (2004), four of its seven risk factors have high sensitivity & specificity. Those are confusion/disorientation/impulsivity, altered elimination, antiepileptics, and benzodiazepines.

Morse assesses six risk factors. Those are history of falling within the past three months, secondary diagnosis, ambulatory aid, IV/Heparin lock, gait/transferring, and mental status. Scoring is similar to Hendrich II because number of points varies with each risk factor. For example, ambulatory aid warrants 15 points if a patient uses a walker or 0 points if on bedrest. IV/Heparin lock by comparison warrants 20 points if present or 0 points if not present. A score of 51+ is high risk, 25 - 50 low risk, and 0 - 24 no risk. Results on sensitivity & specificity of Morse are also mixed as with Hendrich II. In Oliver et al. (2004), Morse had high sensitivity 73.2% and high specificity 75.1%. Callis (2016) determined conflicting results however, high sensitivity 88.3% yet low specificity 48.3%. At worst, Morse had very low sensitivity 36.9% and low specificity 54% (Nassar & Madi, 2014). Per Oliver et al. (2004), four of its six risk factor have high sensitivity & specificity. Those are history of falling, ambulatory aid, gait/transferring, and mental status.

Last is STRATIFY, which is perhaps the simplest tool of all six reviewed. It assesses the least number of risk factors at five. Those are recent history of fall, agitated, visually impaired, frequent toileting, and transfer & mobility. Scoring is simple. Each risk factor is a yes or no question, which corresponds with 1 or 0 points, respectively. For example, a patient who does in fact require frequent toileting warrants 1 point. Transfer & mobility is the only exception because it comprises two parts. Transfer for instance warrants 3 points if a patient independently transfers from bed to chair. By comparison, mobility warrants 3 points if that patient independently ambulates. If the combined score of these two parts is 3+ points, then it warrants 1 point for transfer & mobility. A total score of 2+ is high risk. Results on sensitivity & specificity are mixed once again, following the same narrative seen in Hendrich II and Morse. In Oliver et al. (2004), sensitivity & specificity were the highest of any tool, 93% and 87.7%, respectively. Park (2016) confirmed high sensitivity 89%, yet specificity was low at 67%. Then the opposite was determined in Callis (2016), low sensitivity 55% yet high specificity 75.3%. A retrospective observational study on 365 patients (*n* = 365) across 40 units in an acute care hospital in Italy certainly did not bolster these results. In that study, STRATIFY had very low sensitivity 27% and low specificity 50% (Castellini, Demarchi, Lanzoni, & Castaldi, 2017). Circling back to Oliver et al. (2004), four of five risk factors in STRATIFY have high sensitivity & specificity. Those are history of fall, agitated, frequent toileting, and transfer & mobility.

Results of Statistical Analysis

There were two findings from the two-sample *Z*-test specifically that were statistically significant. First, there was a difference between student and nurse scores on the dizziness/vertigo risk factor in Hendrich II ($p = 5.51^{-10}$). Second and similarly, there was a difference on the Get-Up-And-Go subtest ($p = 0.05$). The high frequency of differences rules out chance variation. Of note, both differences also had confidence intervals 95% (see Appendix).

There were also two general findings aside from those from the *Z*-test. First, students were more likely than nurses to score patients as high risk. Specifically, students scored 29% of the entire sample as high risk, while nurses by comparison scored 25% as high risk. This difference equates to about 4.24 patients that students and nurses disagree on fall risk status. Second, nurses did not complete PMAT on 40% of patients.

Discussion

Pertaining to the first objective, no explicit recommendations for an alternative to Hendrich II can be made at this time. This is solely because of conflicting results on sensitivity & specificity of any given tool. For instance, STRATIFY would be suitable given the highest recorded sensitivity & specificity compared to other tools. Moreover, four of its five individual risk factors have high sensitivity & specificity (Oliver et al., 2004). It is also helpful that it is perhaps the easiest and quickest tool to use, requiring only yes or no responses for the most part. Nurses would certainly appreciate the reduced amount of time needed to complete a fall assessment with STRATIFY than with Hendrich II. Ultimately, this would increase compliance and the quality of assessments altogether. Such is wishful thinking however, since Castellini et al. (2017) determined the lowest sensitivity & specificity for STRATIFY, contending with the findings of Oliver et al. (2004). At best, Park (2017) recommended using two tools at once in order to leverage the strengths of both. This aligns with this hospital's simultaneous use of Hendrich II and PMAT.

Pertaining to the second objective, all three themes are generally due to the disinclination of nurses to accurately complete the labor-intensive parts of Hendrich II, especially dizziness/vertigo and Get-Up-And-Go. The first theme is that these parts require nurses to devote excess time to aid patients with ambulation that would otherwise be devoted to higher priority tasks. To be more specific, ambulating a patient takes about ten minutes, perhaps even more considering the patient population is predominately geriatric. This might not seem like much, yet it is important to note that nurses are required to do this for four patients. That summates to approximately 40 minutes of time otherwise devoted to those tasks. Quite honestly, this is not a realistic expectation given tasks such as physical assessment related to medical diagnoses, and medication passing. Rightfully so, those sorts of tasks are higher priority because they are more impactful on patient outcomes in general. Thus, if an alternative tool is desired, it must be quicker to complete than Hendrich II so that time is not taken away from such tasks.

The second theme is that that disinclination also explains the underestimation on fall risk status on about 4.24 patients. Nurses tend to higher priority tasks, and it simply saves time to assess those patients as low risk. This is potentially catastrophic because those are patients that otherwise should be subject to aggressive fall risk prevention.

The third and final theme is that that disinclination explains why PMAT was not completed on 40% of patients. Again, it simply takes away time otherwise devoted to higher priority tasks. Additionally, it is important to note that there is no obligatory process for PMAT assessment currently. This only reinforces that disinclination. Objectively, given PMAT is a very new implementation, it is also encouraging that compliance is reasonably high at 60%. In any case, it is the responsibility of the committee to foster buy-in. Also of note, compliance with PMAT varied depending on patient population of any given unit. A unit dedicated to orthopedic surgeries for instance had relatively higher compliance than other units probably because mobility is more highly regarded in the rehabilitation process.

In sum, an alternative tool to Hendrich II cannot be explicitly recommended at this time because of mixed results on sensitivity & specificity in the literature. STRATIFY would be the most suitable because it recorded the highest sensitivity & specificity (Oliver et al., 2004). In addition, it is a potential solution to the disinclination of nurses to complete fall assessments because it can be completed on patients more quickly than with Hendrich II, ultimately saving time. Yet to reiterate, even STRATIFY had low sensitivity & specificity when studied elsewhere (Castellini et al., 2017).

Implementation

A pilot test of any tool to the liking of the committee will utilize a two-group posttest only design. This is advantageous because it seamlessly works into this hospital's existing workflow. In particular, one medical-surgical unit will be the control group, while a different medical-surgical unit with a similar patient population will be the experimental group. Additionally, data can be collected during patient admission. Turnaround time from data collection to analysis will be fast as well. A fall assessment for one patient can be completed in a matter of minutes, meaning it is feasible to complete assessments for a 30-patient census in a day. The tentative timeframe is March 2018 when students from University of San Francisco (USF) entry-Master's in Nursing Cohort 22 will be assigned to this hospital for their internship.

The test entails four steps. In step one, day shift nurses in the control group will complete Hendrich II assessments on their patients as they normally would. In the same manner, day shift nurses in the experimental group will complete STRATIFY assessments. Students will then retrieve these assessment scores via retrospective chart review for later analysis.

Step two will be the most labor intensive as it requires students to complete the same fall assessments on the same patients in both control and experimental groups. Immediately afterwards, students will compare their scores alongside that of the nurses in order to tally four distinct values for both control and experimental groups: True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). These values are necessary to compute sensitivity & specificity of Hendrich II and the other tool being tested.

TP is total number of patients correctly identified as high risk for falls. For this to be the case, both students and nurses will have assessed patients as high risk. FP is total number of patients incorrectly identified as high risk. In this instance, nurses will have assessed patients as high risk, while students confirm patients are low risk. TN is total number of patients correctly identified as low risk. In this case, both students and nurses will have assessed patients as low risk. Finally, FN is total number of patients incorrectly identified as low risk. For this to be the case, nurses will have assessed patients as low risk, while students confirm high risk. Of note, the tools define the threshold for high vs. low risk. A score of 5 or greater is high risk in Hendrich II. Determination of high risk in the tool being tested is the responsibility of students.

In step three, students will calculate sensitivity of Hendrich II and the tool being tested by inputting TP and FN into the following formula: $\frac{TP}{TP+FN}$. Specificity will be calculated by inputting FP and TN into a similar formula: $\frac{TN}{TN+FP}$ (Center for Evidence-Based Medicine, 2017). Whichever tool yields higher sensitivity & specificity is the better tool in terms of correctly identifying patients who are truly high risk.

Step four will evaluate sensitivity & specificity by conducting repeat tests on different units. Number of repeat tests is at the discretion of the committee. Repeat tests are encouraged because they ensure that higher sensitivity & specificity of either tool are not due to chance.

This test design is not without its weaknesses. Lack of random sampling is the most obvious. Convenience sampling by way of patient census on medical-surgical units will be used in place of random sampling because convenience sampling is much easier to obtain. It is not ideal, yet it expedites the time to experimentation. Moreover, it expedites calculation of sensitivity & specificity, which are critical in determining the better tool for this hospital's needs.

Another weakness is that the overall quality of data is dependent on both students and nurses completing fall assessments accurately. This is not always the case in normal daily workflow, exemplified anecdotally in instances where assessments are completed with minimal effort. Thus, it will be imperative to reinforce the importance of accurate assessment, perhaps during the morning huddle prior to start of shift.

Cost Analysis

To clarify, it was never the intent to reduce the fall rate at this hospital. Such an initiative is large enough to warrant separate research. The objectives at hand are far more modest. Nonetheless, since it is likely that the fall committee will want to reduce the fall rate in the future, it is worth briefly discussing cost saving potential.

The committee did not disclose their specific fall rate, so a fall rate of 4.03 per 1,000 pd is assumed (Bouldin et al., 2013). Important to note that pd is not the same as days in a calendar year. Rather, it measures number of inpatients receiving care in the hospital for one day only. Thus, 1,000 pd is more easily understood as 1,000 inpatients receiving hospital care in a day. Taking this fall rate in conjunction with cost per fall of $30,000 (CDC, 2017), savings can be as high as $120,000 per 1,000 pd: $\frac{4.03\ falls}{1{,}000\ pd}\ x\ \frac{\$30{,}000}{fall} \approx \frac{\$120{,}000}{1{,}000\ pd}$.

As for pilot testing, there are only two main cost items. First is a one-month subscription to the alternative fall risk assessment tool. Admittedly, no dollar figure could be found on this. Thus, $75,000 per month is a rough estimate based only on the general knowledge that custom changes in an electronic health record are quite costly. Second is hourly wages that are paid to employees in the information technology department who will input the tool into the hospital's electronic health record for trial use. Assuming two employees, an hourly wage of $75, and 40 hours in a week devoted to input, this summates to $6,000. In total, cost of pilot testing in a month is $81,000. While this is a costly venture, it is important to note that cost per fall of $120,000 per 1,000 pd is far more costly. Also of note is that implementation of pilot testing is completely cost-free because cohort 22 students will oversee it for their internship.

Future Directions

It is worth mentioning three ways that could have improved workflow and, ultimately, the quality of the data. First, it would be ideal for the fall committee to explain in detail how nurses at their hospital complete Hendrich II assessments. Not until the last day of student assessments was it determined that students were scoring Get-Up-And-Go differently. To be specific, students were unaware that the hospital assigns a score of zero for patients identified as "Physically Unable to Get Up." Students had been scoring a four for these patients instead, which in Hendrich II coincides with "Unable to rise without assistance during test." Moreover, it was unknown whether or not these patients were supposed to be excluded from the data as with patients on formal bedrest. Thus, to be transparent, this may have skewed the results slightly.

Second, it would have been helpful if staff or students from USF reported on the previous student's findings. To provide more context, several cohorts of students have worked towards these same objectives in the recent past. Their findings therefore might have provided a more solid basis in regards to identifying a better alternative to Hendrich II. Instead, minimal information was provided, so it was as if the entire process started from the very beginning.

Third, given the large workload of fall assessments on practically the entire hospital, it was imperative to have at least four students to make the process quicker, safer, and more manageable. Instead, there were only three students, making assessments slower and unsafe at times. For instance, there were a few cases in which there were not enough persons present to spot a patient who had an unsteady gait.

Concluding Remarks

In sum, it is worth reiterating that no alternative fall risk assessment tool can be recommended at this time because the literature is very limited. At best, STRATIFY may be suitable because it had the highest recorded sensitivity & specificity. Comprising only five risk factors, it is also the easiest and quickest tool to use, and nurses will surely appreciate how much time it saves. Regardless the tool however, effectiveness is highly dependent on compliance. Even the best tools are meaningless unless nurses internalize the importance of fall assessments and, subsequently, assess patients accurately. To that end, there is also tremendous responsibility on the hospital's leadership to create a sense of urgency about falls. Currently, it seems that falls are regarded as less important than other tasks such as medication passing. As a starting point, it may be beneficial for leadership to ask unit managers and charge nurses to discuss the importance of fall assessments during huddles with staff nurses.

References

Bouldin, E. D., Andresen, E. M., Dunton, N. E., Simon, M., Waters, T. M., Liu, M., & Shorr, R. I. (2013). Falls among adult patients hospitalized in the United States: prevalence and trends. *Journal Of Patient Safety*, *9*(1), 13-17. doi:10.1097/PTS.0b013e3182699b64

Callis, N. (2016). Falls prevention: identification of predictive fall risk factors. *Applied Nursing Research, 29*(1), 53-58. doi:10.1016/j.apnr.2015.05.007

Castellini, G., Demarchi, A., Lanzoni, M., & Castaldi, S. (2017). Fall prevention: is the STRATIFY tool the right instrument in Italian Hospital inpatient? A retrospective observational study. *BMC Health Services Research*, 171-7. doi:10.1186/s12913-017-2583-7

Center for Disease Control & Prevention (CDC, 2017). Costs of falls among older adults. Retrieved from https://www.cdc.gov/homeandrecreationalsafety/falls/fallcost.html

Center for Disease Control and Prevention (CDC, 2017). Important facts about falls. Retrieved from https://www.cdc.gov/homeandrecreationalsafety/falls/adultfalls.html

Center for Evidence-Based Medicine (2017). SpPin and SnOut. Retrieved from http://www.cebm.net/blog/2014/03/03/sppin-and-snnout/

Klinkenberg, W. D., & Potter, P. (2017). Validity of the Johns Hopkins fall risk assessment tool for predicting falls on inpatient medicine services. *Journal Of Nursing Care Quality, 32*(2), 108-113. doi:10.1097/NCQ.0000000000000210

Nassar, N., Helou, N., & Madi, C. (2014). Predicting falls using two instruments (the Hendrich Fall Risk Model and the Morse Fall Scale) in an acute care setting in Lebanon. *Journal Of Clinical Nursing, 23*(11/12), 1620-1629. doi:10.1111/jocn.12278

Oliver, D., Daly, F., Martin, F., & McMurdo, M. (2004). Risk factors and risk assessment tools for falls in hospital in-patients: a systematic review. *Age & Ageing*, *33*(2), 122-130.

Park, SH. (2017). Tools for assessing fall risk in the elderly: a systematic review and meta-analysis. *Aging Clinical and Experimental Research.* doi: 10.1007/s40520-017-0749-0

Appendix

z-Test: Two Sample for Means	All Units (106 Patients)	
	Student	RN
	Dizziness, Vertigo	*Dizziness, Vertigo*
Mean	0.330188679	0.028301887
Known Variance	0.223	0.028
Observations	106	106
Hypothesized Mean Difference	0	
z	6.203834234	
P(Z<=z) one-tail	2.75519E-10	
z Critical one-tail	1.644853627	
P(Z<=z) two-tail	5.51038E-10	
z Critical two-tail	1.959963985	

Reject Null Hypothesis because p=5.52479173521192E-10 is lower than alpha=0.05
There is a difference in the means of each Dizziness, Vertigo sample

z-Test: Two Sample for Means	All Units (106 Patients)	
	Student	RN
	Get Up & Go Test	*Get Up & Go Test*
Mean	1.29245283	1.764150943
Known Variance	2.476	2.239
Observations	106	106
Hypothesized Mean Difference	0	
z	-2.2365377	
P(Z<=z) one-tail	0.012658285	
z Critical one-tail	1.644853627	
P(Z<=z) two-tail	0.025316571	
z Critical two-tail	1.959963985	

Reject Null Hypothesis because p=0.0251764997007813 is lower than alpha=0.05
There is a difference in the means of each Get Up & Go sample