Research Article

# Use of the rasch model on nursing treatment of sleep disorders for cardiovascular post-surgery patients

## Abstract

This work follows the actual nursing practice for cardiovascular post-chirurgical care of patients, measuring their evolution at 24, 48 hours and after the patient leaves the intensive care area (around 72 to 96 hours). A 44 items questionnaire developed in Colombia and administered to about 250 patients in a single year, includes seven items for variables that affect sleep patterns. Measures must help to discriminate between traits resulting from the natural improvement of the health status of the patient, and other traits that need some special nursing care or a medical treatment, in particular to detect sleep disorder. Items are rated in two or three categories (from low to high health status) and analyzed in a two parts framework: (1) The improvement of health and sleep pattern of the patients in the period from 0 to 96 hours (the trait measurement must indicate better health condition at the end of the period) and (2) The identification of critical variables where positive evolution does not occur or where some abnormal condition is found, needing a nursing intervention to help the patient to reach better patterns in health and sleep. The analysis uses item characteristic curves looking if the observed outcomes of the traits fit the Rasch model and have the highest scores at the end of the period, otherwise a different nursing medical intervention should be proposed.

**Keywords:** rasch analysis, rasch model, sleep disorder, nursing treatment, cardiovascular post surgery patients

Agustín Tristán López,[1] Nancy Y Pedraza Corpus,[2] Claudia Ariza Olarte,[3] María Mercedes Durán de Villalobos[4]

[1]Direction, Instituto de Evaluación e Ingeniería Avanzada, México
[2]Psycho-Pedagogical Department, Instituto de Evaluación e Ingeniería Avanzada, México
[3]Universidad Pontifica Javeriana, Colombia
[4]Graduate Nursing School, Universidad Nacional, Colombia

**Correspondence:** Agustín Tristán-López, Direction, Instituto de Evaluación e Ingeniería Avanzada, Cordillera Occidental 635, Lomas 4ª sección, 78216 San Luis Potosí, México, Tel (52)4448255077, Email atristan@ieia.com.mx

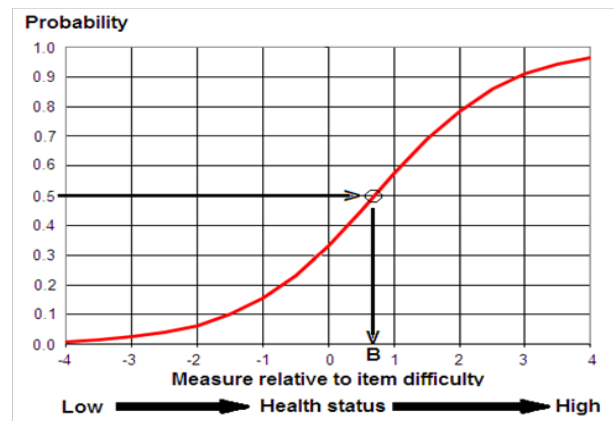**Abbreviations:** ICU, intensive care unit; RSA, rating scale analysis

## Introduction

Coronary disease is one of the first **morbidity c**auses worldwide where myocardial revascularization surgery is a usual medical intervention. Surgical treatment and patient's recovery process imply intensive care where the nurse participates actively and dynamically. This work shows the use of a questionnaire following a minimally invasive nursing procedure[1] including the most relevant clinical events that arise after surgery,[2] in order to identify those where the nurse's intervention is vital for an optimum recovery process, so that the patient may continue with a "normal" lifestyle as soon as possible.[3]

The main interest of this work is to show how Rasch analysis is a convenient tool to identify the most relevant traits that require nursing intervention or medical treatment, using the calibration of the test and the items. The Rasch model is suitable to evaluate the relationship of symptoms and signs associated with sleep disorders and the process of postoperative recovery on this kind of patients.

A complete mathematical formulation is out of the purpose of this paper, but it is important to say that the Rasch model is a function that provides the probability to answer an item by a person, knowing the measure of the stimulus and the person's measure.[4–6] The graph of the Rasch function has a logistic shape where horizontal axis represents the measured health status trait (from $-\infty$ to $+\infty$) in units called logits (for log odd ratio units), and the vertical axis is the probability (from 0 to 1) of the expected score. If data fit the Rasch model then it is possible to say that the patient's evolution is satisfactory, any

departure (or misfit) to the model may indicate a problem with the health status of a specific item.[7] Figure 1 shows the shape of the Rasch model (called item characteristic curve or ICC); the item measure B (or item difficulty) is located at the inflection point, where p=0.5; in this example B=0.7logits.
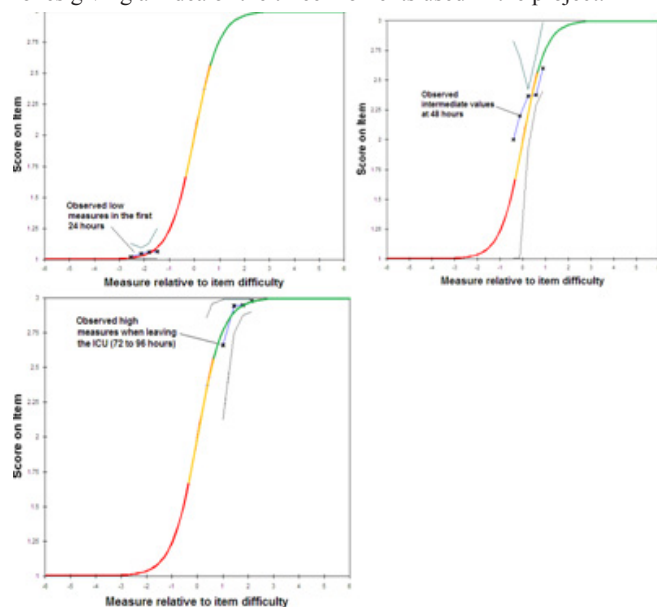


**Figure 1** Rasch model of expected score.

Item and person measures share the same scale. Suppose the nurse will rate the trait from 1 (bad condition) to 3 (good condition) on the item of Figure 1. A person with a health status of 2logits, has a probability of 0.8 to get high outcomes (probably rated as 3) in the trait represented by this ICC, while a person with a measure of -1.5 logits, has a probability of 0.1 to be in good condition (probably rated in category 1).

## Several clinical and practical implications come from the Rasch model

**Interpretation of the rasch measures:** The measures in different moments for a single trait show the evolution of the health status of the patients during the 24 and 48 post-surgical recovery period and after the patient leaves the intensive care area (around 72 to 96 hours). This evolution may be positive, negative or stable, depending on the evolution of each patient; a positive evolution is shown in Figure 2, where observed measures are compared to the Rasch ICC in three zones giving an idea of the three moments used in the project.
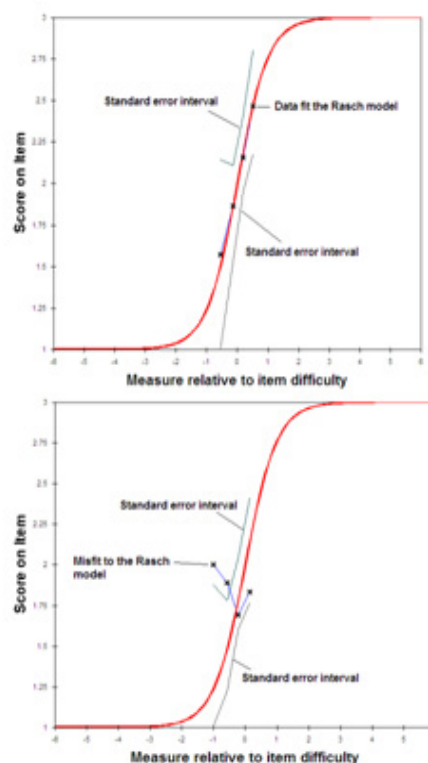


**Figure 2** Evolution on time of the Rasch measures.

**Interpretation of fit:** Real data from the nurse's observation may fit or not the Rasch model. If the item fits the Rasch model, then the observation corresponds to the expected theoretical measures explained in the previous pint. On the contrary, if an item shows a misfit, then the measures cannot be correctly modeled and the nurse has a powerful tool to identify a critical variable where intervention is needed to improve the patient's recovery. The fit parameter is explained later.

In the educational field, item misfit is regularly used to discard the item as it may show guessing, errors in the options, and other kind of defects on item design, administration or students' answers.[7] But in the cardiovascular post-surgery application, misfit has a useful interpretation for nursing practice: for instance blood pressure of patients may show misfit but it is evident that the item cannot be discarded because it is an important symptom to be cared by the nurse. What do misfit means in this case? Bad manometer's functioning? Bad application of the procedure? Guessing is out of focus! Therefore misfit is not related to a wrong functioning of the item but to a different kind of problem, it can be produced by a wrong rating of the item but mainly it has to do with nursing intervention, clinical attention, inconvenient administration of a drug, unexpected patient's reaction to the treatment, and so forth (Figure 3).

As a consequence of those implications, it is possible to see that the Rasch model may provide a useful source of information regarding the positive evolution of the patient and the problems, effects, limitations or improvements of nursing care.[8]



**Figure 3** Fit and misfit to the Rasch model.

The items used in the questionnaire follow the rating scale analysis (RSA), that according to the Rasch-Andrich model it is possible to construct measures from observations based on an ordinal rating scale according to this expression:[9] $log\ (P_{nik}\ /\ P_{ni(k-1)})\ =\ B_n - D_i - F_k$.

Where:

$P_{nik}$ is the probability to respond category k on item i by person n

$P_{ni(k-1)}$ is the probability to respond in category k-1

$B_n$ is the trait measure of person n

$D_i$ is the difficulty of item i

$F_k$ is the impediment to respond category k relative to category k-1.

The categories of the RSA must follow some guidelines for Andrich rating scales,[10,11] in particular they must show that thresholds (points with equal probability for two adjacent categories) are ordered, distinctly located and that no categories are missing or exceed the construct needs. In this analysis we used the Andrich rating scale model as usual in similar projects,[12,13] instead of the Masters Partial Credit used by other authors,[14] considering that it is a more parsimonious model and simpler to use by nurses because the categories are more evident for the complete instrument. In addition, the Andrich rating scale makes easier to assess unidimensionality of the scale and the structure of the data, in consequence each category has a real probability to be selected by the nurse for the patient's evaluation.

Finally, it is important to remember the concept of objectivity,[15]

an attribute that is a must for this type of questionnaires, in order to ensure fairness of the questionnaire, especially in RSA where the nurse is rating the evolution of the patients. Objectivity clarifies the interpretation of the outcomes and reduces the sources of bias that may affect the ethical and clinical use of the test.[16] In particular it is important to sketch that the Rasch model concerns the local objectivity of the construct or trait that is measured as a consequence of the set of data fitting model.[17]

## Materials and methods

A questionnaire of 44 health status traits, signs or symptoms, grouped by system, was reviewed by nurses and physicians (called judges for this research). The groups of topics considered included in the questionnaire are: neurological, cardiovascular, respiratory, gastrointestinal, elimination and skin, focusing on medical aspects (for instance Oxygen Saturation, Respiratory rate, Work of breathing, Chest expansion, Arterial oxygen saturation or Skin coloration, among others),[18,19] easily linked to the NANDA Domain 4 (Activity&/Rest) with the four Class 1 diagnosis for Sleep/Rest: Insomnia, Sleep deprivation, Readiness for enhanced sleep and Disturbed sleep pattern.[20]

The questionnaire included 44 rating scale items for health status traits, to be observed by the nurse who is in contact with the patient. Seven items are related to sleep pattern, considered an important factor for the patient's health recovery and to prevent complications in the postoperative period.[21] Each rating scale item has two or three categories (from low to high status of the trait, sign or clinical event). Judges considered that the clinical events evolve quickly from day to day, justifying the need to observe patients after surgery 24 and 48 hours, as well as between 72 and 96 hours, time when released if they do not experience complications.

**The seven items of interest for this paper are (Numbering corresponds to the questionnaire position)**

i.  2. Sleep and rest. Criteria: [1: Altered sleep pattern] [2: not used] [3: Not altered sleep pattern]

ii.  11. Respiratory rate. Criteria: [1: <80 or >100 per minute] [2: not used] [3: 80 to 100 per minute]

iii.  12. Respiratory Work. Criteria: [1: Altered] [2: not used] [3: Normal]

iv.  13. Thorax expansion. Criteria: [1: Asymmetric] [2: not used] [3: Symmetric]

v.  14. Arterial saturation. Criteria: [1: <90% or > 95%] [2: not used] [3: 90% to 95%]

vi.  15. Skin coloring. Criteria: [1: Altered: pale; yellowish] [2: not used] [3: Normal]

vii.  42. O2 saturation. Criteria: [1: <90% or > 95%] [2: not used] [3: 90% to 95%].

The questionnaire was administered to 250 patients in several hospitals in Bogota (Colombia) by nurses from Intensive Care Units (ICU), once a day during three to four days, describing the evolution of each patient. For this work, the clinical events that are associated with the sleep apnea-hypopnea syndrome were selected because the prevalence of these disorders is related to cardiovascular disease.[22,23]

The presence of this syndrome (or some signs or symptoms associated with) can increase the difficulty in the recovery of patients who are in the ICU.

Winsteps®[24] is the commercial software used for Rasch analysis in this project. The program provides the item measure B (in log odd units or logits) and fit parameters, in particular the standardized residual ZSTD. The couple {B, ZSTD} is reported at three moments: 24, 48 hours and after leaving the ICU (from 72 to 96 hours). A positive improvement in the health status is expected from the beginning to the end of the period (B should increase monotonically), and also a reduction of misfit should indicate a better behavior of the patient according to the measured trait (ZSTD is acceptable if less than 2.0). A misfit higher that 2.0 are related to some problems in the measure of the variable and nursing or clinical intervention would be needed. In the following discussion, the set of measures are indicated in the form $B_{24}$, $B_{48}$ … $ZSTD72$.

It is possible to assign items to two rating scale groupings using the command ISGROUPS of Winsteps, and get the plot the curves for the categories. Figure 4 displays the category probability curves for the 40 dichotomous and the four items with three categories.
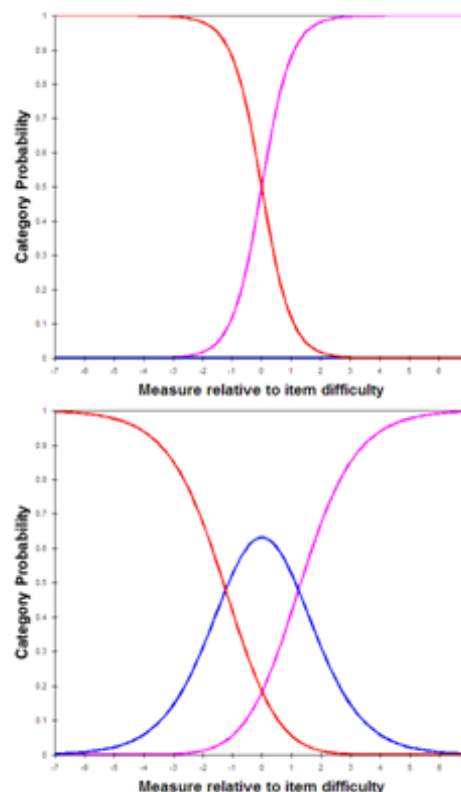


**Figure 4** Categories of the rating scale items.

Dichotomous items have a single threshold and the three categories have two Andrich thresholds, that appear ordered and well defined, satisfying the eight guidelines used to evaluate the items' functioning:[10,11,25]

i.  Frequency responses of each category are more than the minimum suggested of 10 observations. No need to increase the sample size to calibrate the items and categories.

ii. The categories are distributed in a regular pattern. There is no need to redistribute adjacent categories.

iii. The average category measures advance monotonically as expected from low to high. No need to redefine the interpretation of the categories.

iv. Outfit of categories is not extreme, and no categories need to be combined or dropped from the statistical point of view.

v. The Andrich thresholds increase monotonically in the scale, corresponding to the expected evolution of the construct, from low to high health status.

vi. Rating and measures are positively correlated. The categories are not nominal; instead they have a position in the scale, useful to construct measures.

vii. The distance between thresholds is significant, in general higher than 1.4 logits. No need to collapse categories.

viii. The distance between thresholds is less than 4 logits. There is no evidence that intermediate categories should be included.

For the theoretical ICC the software plots the curve relative to item difficulty, that is the same for all the items but the empirical points and the confidence limits are different for the three different moments of the trait. The positive expected evolution described in Figure 2 follows three curves: (A), (B) and (C) corresponding to each of the three moments already defined.

## Results and Discussion

Wright map is a useful tool to show the distribution of frequencies of persons´ measures and item´s difficulties distribution in a single scale.[26] Each item appears three times in this map (Figure 5), for the three measures described previously. Patients are located in the left part of the map (from low to high health status), while items are located to the right. Higher categories bring negative measures for "better health condition". For instance, item 2 has its first measure 2A close to 3.2logits, where category 1 is predominant; at 48 hours good results are more frequent and 2B is close to 1.2 logits and finally 2C has a measure near to 0.5 logits that implies the best condition of the moments. It is important to note that these measures appear in the ICC with a reversed sign, that is $B_{24}$=-3.2, $B_{48}$=-1.2 and $B_{72}$=-0.5.



**Figure 5** Wright map for the complete questionnaire.

It is interesting to see the evolution of some particular items in the Wright map, for instance:

i. Item 15 (Skin coloring) has the largest range on the evolution of health status: it shows the worst condition at the beginning with $B_{24}$=4.98logits, but increases above the mean in $B_{48}$=-0.89 and finishes in a very good condition at $B_{72}$=-2.48logits. The improvement for patients is higher than 7.5 logits, a very relevant amount for the health condition of the patients when leaving the ICU.

ii. Item 2 (Sleep and rest) shows a positive pattern from $B_{24}$=3.22logits and $B_{48}$=1.20logits, but the maximum outcome attains a mean value $B_{72}$=0.51logits. This trait could be improved trying a different nursing intervention or another clinical treatment.

iii. Item 13 (Thorax expansion) has an intermediate value $B_{24}$=-1.57logits for the first measure and ALL patients get the higher category in the second and final moments. In this case the Rasch model cannot provide a measure but Winsteps gives an estimate of $B_{48}$=$B_{72}$=3.96logits.

Similar descriptions of the behavior and evolution of other items are used to specify interventions on time for each trait.[3] For instance, it can be seen that item 5 (Heart rate) has an unexpected behavior. The range between extreme measures is about 1.15 logits that is significant. It initiates with $B_{24}$=0.77 and improves to the best condition in $B_{48}$=0.39hours but slightly decrease when leaving the ICU with $B_{72}$=0.45. This can indicate a problem with the patient's health status at the end of the period.

The complete set of items was validated from the initial steps of the project, showing two main behaviors or clinical groups:

i. The first clinical group comprises the variables showing misfit, for instance pain and physical activity in the neurological system or blood pressure in the cardiovascular system.

ii. The second clinical group includes those variables showing acceptable fit to the Rasch model and a positive evolution, but do not reach a satisfactory health status at the end of the period. Variables included in this group are: sleep and rest (neurological system); arterial saturation and pulmonary auscultation (respiratory system); vein punctures (skin) electrolytes $Na^+$ and $K^+$; complete blood count, packed cell volume, and hemoglobin (prognosis); in all those traits a nursing committee may define new interventions to improve the patient's evolution.

There is a third group of traits with a positive evolution and a good item fit to the Rasch model, where nursing intervention does not need to change and may continue according to the protocol. The seven items associated by judges to sleep pattern and related symptoms show the following characteristics:

**Item 2: Sleep and rest:** According to the Rasch Model, the couple {measurement-fit} has these results: $B_{24}$=- 3.22, $B_{48}$=- 1.20, $B_{72}$=- 0.51 logits. The evolution of the measure (from -3.22 to -0.51) means that the patients progress in their recovery process, there is a favorable evolution with regard to sleep and rest, denoted by a considerable increase of 2.71 logits. However, a significant percentage of patients (67%) continue with this altered variable (Figure 6).

i. Regarding fit, $ZSTD_{24}$=0.4, $ZSTD_{48}$=1.0, $ZSTD_{72}$=1.5. This means that sleep and rest shows a good fit to the Rasch model in

the period, a reasonable but not sufficient evolution is observed, because sleep and rest recovery is present only in 49% of the patients. It is probable that nursing intervention from 24 to 48 hours does not need to be modified, but something has to be done between 72 and 96 hours.

ii. **Item 11: Respiratory rate:** Measures of this trait are $B_{24}=0.00$, $B_{48}=0.07$, $B_{72}=0.47$; while $ZSTD_{24}=1.00$, $ZSTD_{48}=1.60$ and $ZSTD_{72}=0.50$ (Figure 7).

iii. **Item 12: Work of breathing:** This trait has a clear positive evolution with $B_{24}=0.37$, $B_{48}=0.62$ and $B_{72}=1.01$, with a good fit in the three moments, with $ZSTD_{24}=0.10$, $ZSTD_{48}=-0.10$ and $ZSTD_{72}=0.4$ (Figure 8).

iv. **Item 13: Chest expansion:** This variable provides measures only for moment A with the couple of values $\{B_{24}, ZSTD24\}=\{1.57, 0.6\}$. Winsteps® could not obtain measures for the next two days because all the patients are in highest status and despite the lack of calculated measure, chest expansion evolves as expected (Figure 9).

v. **Item 14: Arterial oxygen saturation:** It is an interesting trait because it shows a monotonic growth with misfit only in the first two days $\{B_{24}, ZSTD_{24}\}=\{-0.58, 3.10\}$, $\{B_{48}, ZSTD_{48}\}=\{-0.04, 2.00\}$ and $\{B_{72}, ZSTD_{72}\}=\{0.41, 0.90\}$. This is an example of a trait that should receive a better intervention at the initial days (Figure 10).

vi. **Item 15: Skin coloration:** This trait has the widest range of measures close to 7.5logits: $B_{24}=-4.98$, $B_{48}=0.89$, $B_{72}=2.48$, but $B_{24}$ is only an estimate of the program (no Rasch curve is obtained by the software) because all the patients had a poor status of their skin coloration at the beginning and improved from the second day. This condition can be related to the type of procedure, since the loss of blood and oxygenation causes cyanosis in some cases (Figure 11).

vii. **Item 42: Oxygen saturation:** Values show again a monotonic improvement and misfit indicates how the patients' pattern does not fit the Rasch model on the first two days $\{B_{24}, ZSTD_{24}\}=\{-0.95, 2.10\}$, $\{B_{48}, ZSTD_{48}\}=\{-0.60, 3.20\}$ and $\{B_{72}, ZSTD_{72}\}=\{0.67, 0.30\}$ (Figure 12).
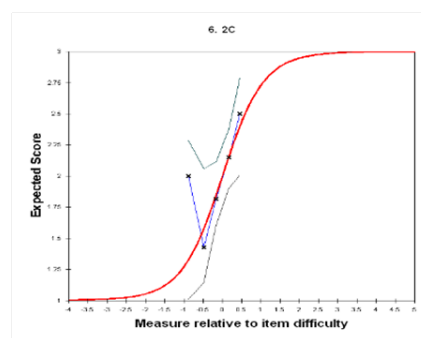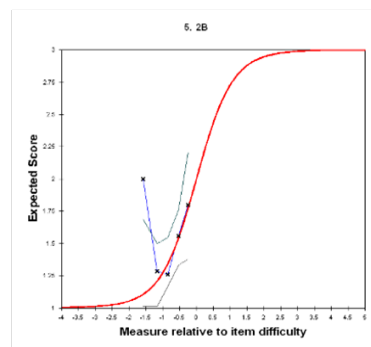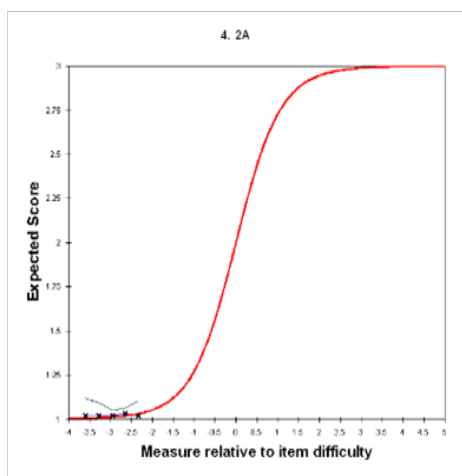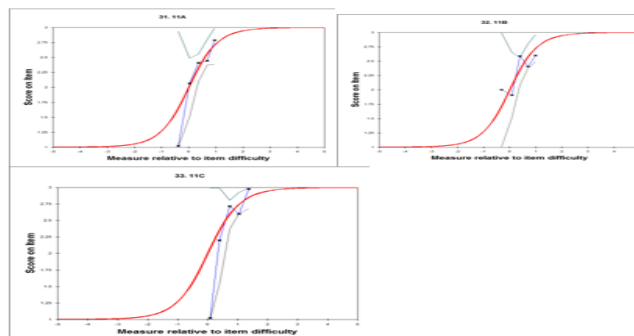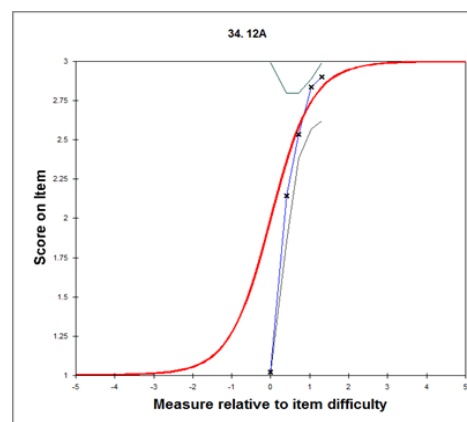




**Figure 6** Item 2. Sleep and rest.



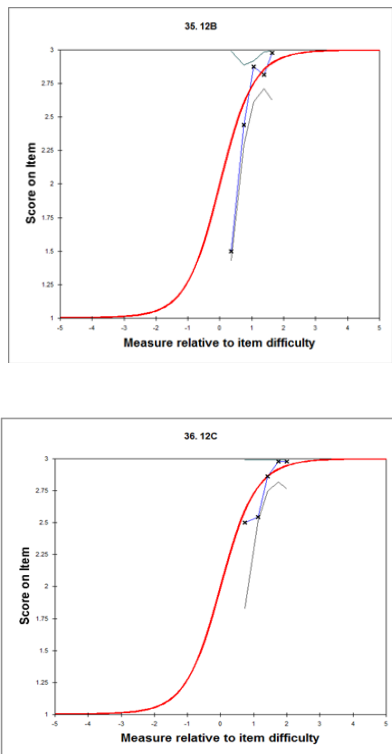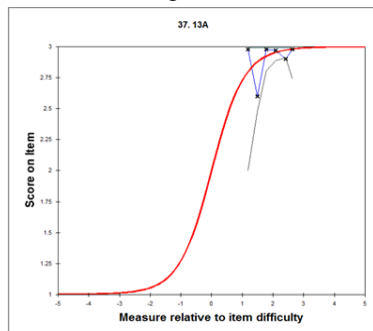**Figure 7** Item 11. Respiratory rate.

Figure 8 Item 12. Work of breathing.



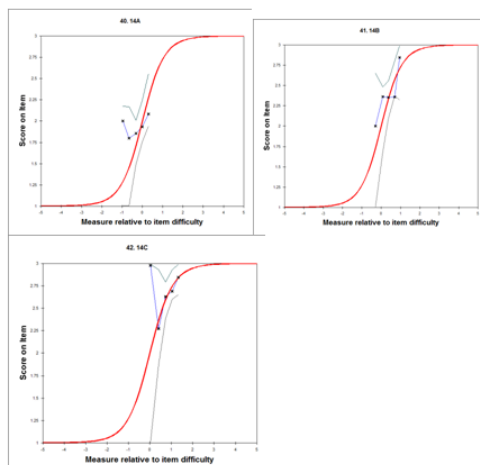**Figure 9** Item 13. Chest expansion.



**Figure 10** Item 14. Arterial oxygen saturation.
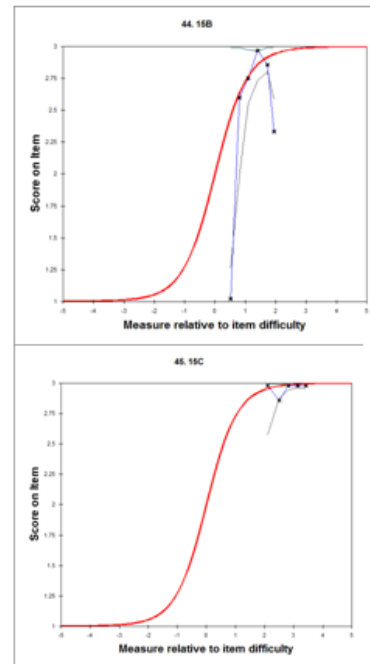


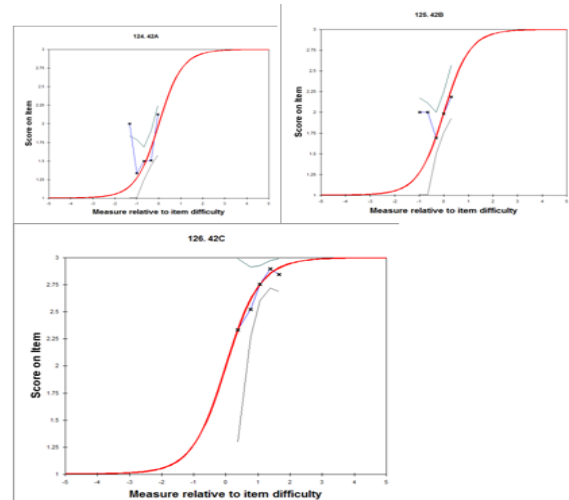**Figure 11:** Item 15. Skin coloration.



**Figure 12:** Item 42. Oxygen Saturation.

## Conclusion

This study shows how the combination of measure and fit provide relevant insight concerning the patient's evolution (in particular for sleep disorder) during a time period and also the feasibility to detect the critical variables needing a closer supervision of the nursing and medical team.

### Analysis should conveniently follow these criteria

### Concerning measures

**A1 Direction:** Rasch measures must vary monotonically on the time series. This pattern imply that the health status is positively improving and the nursing intervention is working as expected. Depending on the direction (or polarity) of the rating scale the user must identify if the variable must increase or decrease monotonically. A non-monotonic

pattern is inadequate and should require a different intervention or a new treatment for the patients.

**A2 Range:** A wide difference between initial and final measures (more than 1logit) is very convenient because it shows a significant improvement in the expected patient's status. A short range (less than 1logit) could be accepted but probably does not correlate with the economic impact of the treatment. Nevertheless in some patients, a stable condition may correspond to a null range and should be evaluated by the nurse or the physician.

### Concerning fit

**B1 Acceptable fit:** Item fit indicates that data correspond to the Rasch model and this tells the nurse that all persons have a pattern close to what is expected. The nurse may continue the intervention according to the planning. If item misfit occurs at the beginning but does not appear at the end of the intervention, then it confirms the pertinence of the treatment.

**B2 Misfit:** Item misfit at each step indicates that each patient does not respond as expected. The nurse has to modify the intervention for each patient, or suggest a new treatment in order to improve the health status of all the patients.

### Concerning medical care

**C1:** Some symptoms and signs associated with sleep disorders are related to important postoperative complications. With the Rasch analysis the health team may supervise each step to ensure the satisfactory evolution of all patients in a given trait.

**C2:** A calibrated instrument contributes to have objective, valid and reliable information about the measured traits. Objectivity is a cornerstone for a questionnaire, the outcomes and the proposal of criteria to increase the possibility of an opportune and effective attention.

## Acknowledgements

## Conflicts of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Vitello-Cicciu J, Fitzgerald C, Whalen D. On the Horizon: minimally invasive cardiac surgery. *J Cardiovasc Nurs*. 1998;12(3):1–16.

2. Deaton C, Namasivayam S. Nursing outcomes in coronary heart disease. *J Cardiovasc Nurs*. 2004;19(5):308–315.

3. Ariza OC. Nursing care model for patients in early postoperative of myocardial revascularization. *Ciencia y Enfermería*. 2016;22(1):35–46.

4. Rasch G. *Probabilistic models for some intelligence and attainment tests.* Copenhagen, Denmark: Danmarks Paedagogiske Institut; 1960.

5. Rasch G. *Probabilistic models for some intelligence and attainment tests (expanded edition).* Chicago, USA: University of Chicago Press; 1980.

6. Wright BD, Stone MH. *Best test design*. Chicago, USA: Mesa Press; 1979.

7. Bond TG, Fox CM. *Applying the Rasch model: Fundamental measurement in the human sciences*. New York, USA: Routdlege; 2015.

8. Tristan LA, Ariza C, Duran MM. Nursing treatment matches the rasch model. *Rasch measurement transactions*. 2008;21(4):1142–1143.

9. Andrich DA. A rating formulation for ordered response categories. *Psychometrika*. 1978;43(4):561–573.

10. Linacre JM. *Guidelines for Rating Scales and Andrich Thresholds*. Midwest Objective Measurement Seminar, Chicago, USA; 1997.

11. Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas*. 2002;3(1):85–106.

12. Panayides P, Gavrielides M, Galatopoulos C, et al. Using rasch measurement to create a quality of sleep scale for a non-clinical sample based on the pittsburgh sleep quality index (PSQI). *Europe's Journal of Psychology*. 2013;9(1):113–135.

13. Sargento P, Perea V, Ladera V, et al. Measurement properties of a screening questionnaire of obstructive sleep apnea risk: Little information, great prediction? *Sleep Sci*. 2014;7(2):89–95.

14. Hart PD, Kang M, Weatherby NL, et al. Evaluation of the short-form health survey (SF-36) using the rasch model. *American Journal of Public health Research*. 2015;3(4):136–147.

15. Gaukroger S. *Objectivity: A very short introduction*. Oxford, UK: Oxford University Press; 2002.

16. Tristán LA, Pedraza CNY. Objectivity in standardized tests. *Revista Iberoamericana de Evaluación Educativa*. 2017;10(1):11–31.

17. Stenner J. Specific objectivity-local and general. *Rasch Measurement Transactions*. 1994;8(3):374.

18. Golbin JM, Somers VK, Caples SM. Obstructive sleep apnea, cardiovascular disease, and pulmonary hypertension. *Proc Am Thorac Soc*. 2008;5(2):200–206.

19. Le Mone P, Burke KM. *Clinical Handbook for Medical-Surgical Nursing*. 4th ed. USA: Prentice Hall; 2008.

20. Herdman TH, Kamitsuru S. Nursing Diagnoses 2015-2017. *Definitions and Classification*. 10th ed. NANDA International, New York, USA: Wiley-Blackwell; 2014.

21. Barimboim E. Sleep apnea as a risk factor in surgery. *Revista Americana de Medicina Respiratoria*. 2014;3(14):272–279.

22. Jean-Louis G, Zizi F, Clark LT, et al. Obstructive sleep apnea and cardiovascular disease: role of the metabolic syndrome and its components. *J Clin Sleep Med*. 2008;4(3):261–272.

23. Lattimore JD, Celermajer DS, Wilcox I. Obstructive sleep apnea and cardiovascular disease. *J Am Coll Cardiol*. 2003;41(9):1430–1437.

24. Linacre JM. *Winsteps Users Manual*. Winsteps.com. Australia; 2017.

25. Royal KD, Ellis A, Ensslen A, et al. Rating scale optimization in survey research: an application of the rasch rating scale model. *Journal of Applied Quantitative Methods*. 2010;5(4):607–617.

26. Wilson M. Some Notes on the Term: Wright Map. *Rasch Measurement Transactions*. 2011;25(3):1331.