Reliability of a Health Questionnaire among Women with Brain Injury

Colantonio, A, Ph.D., OT, Saunderson Family Chair in Acquired Brain Injury Research, Toronto Rehabilitation Institute and Professor University of Toronto, Department of Occupational Science and Occupational Therapy
Harris, JE, PhD, OT, Post Doctorate Fellow, Toronto Rehabilitation Institute
Tarek, N, MSc. OT, Occupational Therapy Department, University of Toronto

Corresponding author:
Colantonio, A, Ph.D., OT,
University of Toronto
Department of Occupational Science and Occupational Therapy
500 University Ave, Room 950, Toronto, Ontario M5G 1V7
angela.colantonio@utoronto.ca
Phone: 416-978-1098
Abstract

Research on health outcomes post traumatic brain injury (TBI) has found women have unique issues in comparison to men. However, no tools have been developed to capture these unique health outcomes. We investigated rater agreement of a health questionnaire developed to identify health concerns among women with TBI. Thirteen women completed the questionnaire twice and provided feedback regarding content and ease of use. Kappa coefficients and percent agreement were generated to determine level of response agreement. Kappa coefficients ranged from 0.21 to 1.00 with the highest agreement for items requiring concrete responses (e.g. yes/no) and the lowest with items requiring more subjective answers. Participants rated the utility of the questionnaire high (73%-100%). This questionnaire proved to be a reliable and useful method to gather information regarding health in women with TBI. By identifying health issues early, diagnostic and treatment intervention may be delivery in a timelier and effective manner.
Introduction

There are approximately 18,000 hospitalizations associated with a traumatic brain injury (TBI) diagnosis annually in Canada (Canadian Institute for Health Information, 2006) and the Centers for Disease Control estimates that 2% of the U.S. population lives with the effects of a TBI (Thurman, Alverson, Dunn, Guerrero, & Sniezek, 1999). The survival rate for those who sustain a TBI has increased in recent years (Bushnik, Hanks, Kreutzer & Rosenthal, 2003; Willer, Allen, Liss, & Zicht, 1991). Subsequently, individuals frequently cope with life long impairments affecting their physical, cognitive, and social functioning (Dean, Colantonio, Ratcliff, & Chase, 2000; Martin, Viguier, Deloche, & Dellatolas, 2001). Although men are twice as likely to sustain a TBI, (Thurman et al., 1999) women are more likely to report worse somatic and psychiatric outcome (Farace & Alves, 2000). Further, Colantonio et al. (2009) found that compared to age matched controls, women with TBI reported significantly greater problems associated with the neuroendocrine system (e.g. amenorrhea). Despite this knowledge, the health-related consequences unique to women are rarely identified and studied.

The decision to provide diagnostic and therapeutic intervention is often dependent on the reliability and validity of the information collected from clinical and research findings; this extends to the outcome measures utilized to collect data. Often data is collected using questionnaires focused on ascertaining information in areas such as physical and mental health, social and community involvement, and quality of life status. Quality of data is a reflection of whether the measure is applied with the appropriate population (Baron, Hales, Hurrell, 1996; McHorney, 1999). Though reliability and validity of measures are often determined using a specific clinical population (e.g. people with cancer), seldom is the measure evaluated beyond
the original population. Yet, measures are often employed to collect data and thus inform health care decisions within a variety of clinical populations.

It is important to assess the properties of instruments prior to use with the TBI population due to the particular sequelae that arise from TBI such as memory, attention and awareness which can affect responses (Bazarian, Cernak, Noble-Haeusslein, Potolicchio, & Temkin, 2009; Dikmen et al., 2009). To assess the impact of injury, questions may need to be altered to evaluate present health and function in comparison to pre-injury health and function. Additionally, reliability of items are particularly important to establish when they can not be easily corroborated by an informant e.g. questions on symptoms/menstrual cycles. To date, there has been very little research on outcomes after TBI that are specific to women. Also, not all measures of health and function variables have been assessed in a TBI population. As such, it was important to test measurement properties prior to use with a larger sample.

The objectives of the study were 1) to measure the magnitude of rater agreement of a health questionnaire relevant to women who have sustained a TBI and 2) to determine the utility of this questionnaire with women who have sustained a TBI.

**Methods**

**Data collection**

Participants were recruited from community agencies that provide service for people with moderate to severe traumatic brain injury. Women were invited to participate in the study if they met the following criteria: between the ages of 20 and 55, sustained a moderate to severe brain injury, at least two years post injury, able to read and speak English, living in the community, and able to provide informed consent. Participants were asked to attend the research clinic for an hour appointment to complete the study questionnaire on two separate occasions one week apart.
Post-Print version

Assistance to participants who required help understanding questions and writing responses was offered. The questionnaire was administered by one study member (NT) on both test sessions. Ethical approval was given by the university board of ethics.

**Outcome Measures**

We chose to focus our health questionnaire on collecting information on women’s health issues, general health status, and participation in important and meaningful activities. The Women’s Health Questionnaire was developed by compiling relevant and validated scales and questionnaires and collecting demographic information. Additionally, we developed a utility questionnaire to measure participant’s response to the appropriateness of the questions and the format of the measure.

**Demographic Information**

Participants age, living situation, marital status, number of children, education, employment history, household income, duration of injury, severity of injury (length of unconsciousness after injury) was collected.

**Women’s Health**

Validated items were drawn from the June 2000 National Population Health Survey in Canada to measure the frequency and extent of breast and cervical cancer screening of participants (Statistics Canada, 2002). Questions include whether or not participants had a Papanicolaou Test (Pap smear) or a mammogram, if so, when the last time was, or if they examine their breasts regularly.

The Women’s Health Questionnaire (Hunter, 1992) was developed to measure the physical and emotional health of women in relation to pre-, peri- and postmenopausal transitions. For this study we used the menstrual difficulties sub-scale. Participants are asked to identify
Post-Print version

physical and emotional symptoms during the menstrual cycle (e.g. sad, anger, cramps, abdominal discomfort) using a using a four-point scale with answers ranging from “yes definitely” to “no not at all”. This questionnaire has been used to study women’s health in a number of studies (Liao, Hunter, & Weinman, 1995; Vickers, 1999; Shulman, Yankov, & Uhl, 2002; Van Leusden, Van Buuren, & Pryse-Davies, 1997) and has been translated into 20 languages (Hunter, 2003). Test-retest reliability was conducted resulting in correlations ranging from 0.78- 0.96 (Hunter, 2003; Silva Filho, Baracat, Conterno, Haidar, & Ferraz, 2005). Concurrent validity was established using correlations ranging from 0.70 to 0.86 (Liao et al., 1995; Hunter, 2000) and using Cronbach alpha internal reliability of sub-scales ranged from 0.59 to 0.84 (Hunter, 2003; Siva Filho et al., 2005). To ascertain the impact of the brain injury on the menstrual cycle, participants were asked whether or not changes had occurred since their injury (e.g. if they had experienced a disruption in their menstrual cycle following their brain injury).

Regarding conception and pregnancy, questions were modified from The Nurses’ Health Study (Michels et al., 1996). A subset of 1,569 women from this study completed a questionnaire that asked about conception, pregnancy complications, and delivery. A study by Tomeo and colleagues (1999) assessed the reliability and validity of this questionnaire and found reliability coefficients ranging from 0.73-0.95 and validity ranging from 0.53-0.94. Women were asked about the regularity and duration of their menstrual periods before and after their injury, specifically, if their menses were absent (amenorrhea), if cycles were irregular, if there were any changes in intensity of periods, and to indicate any other menses-related changes. They were asked if they ever became or tried to become pregnant either before or after their injury. For those who responded positively, further questions covered such issues as difficulty
conceiving, complications during pregnancy, miscarriages, abortions, and postpartum difficulties.

The Greene Climacteric Scale (Greene, 1998) was used to evaluate menopause symptoms. This scale consists of 21 items. Each item is rated by the individual according to its severity using a four point scale: not-at-all (0); a little (1); quite a bit (2); extremely (3). Items address psychological symptoms (e.g. anxiety and depression), somatic aspects, vasomotor symptoms, and sexual dysfunction. This section was only answered by those individuals who identified themselves as in menopause or post menopause. Several studies have utilized this scale to measure climacteric symptoms (Ziaei, Moghasemi, & Faghihzadeh, 2009; Barentsen, van de Weijer, van Gend, & Foekema, 2001; McPhail & Smith 2000). Good test-retest reliability was found with coefficients ranging from 0.83 to 0.87 (Greene, 1998; Green, 1976). Additionally, satisfactory content and construct validity has been demonstrated through factor analysis studies (Greene, 1998; Indira & Murthy, 1980; Kaufert & Syrotuik, 981).

Thyroid disease was examined using an adapted version of The Thyroid Health Survey (Canaris, Manowitz, Mayor, & Ridgway, 2000; Canaris, Steiner, & Ridgway, 1997). All participants were asked about whether or not they were currently experiencing 1) any of the listed symptoms of thyroid dysfunction, 2) if this was a symptom they had prior to their brain injury, and 3) whether or not any of the listed symptoms had emerged in the past year. Symptoms included: constipation, hoarse voice, dry skin, puffy eyes, and weaker muscles.

General Health

Two questions were used to measure pre and post injury general health. Responses were measured using a 5-point scale ranging from “excellent” to “poor”.

Participation in Activities
The Role Checklist, designed and empirically tested by Oakley and colleagues (Oakley, Kielhofner, Barris, & Reichler, 1986), was developed to obtain information on an individual’s perception of their participation in various life roles. Part one of the checklist assesses, along a temporal continuum (past, present, and future), the major roles (e.g. worker, volunteer, caregiver, home maintainer, friend) that is considered part of daily life. Part two identifies the degree to which each role is valued: very valuable, somewhat valuable, and not at all valuable. Studies using the Role Checklist have been published involving diverse populations (Hakansson, Eklund, Lidfeldt, Nerbrand, Samsioe, & Nilsson, 2005; Crowe, VanLeit, Berghmans, & Mann, 1997; Hachey, Boyer, & Mercier, 2001). Two studies evaluated the methodology and reliability of translating the Role Checklist into Spanish (Colón & Haertlein, 2002) and French (Hachey, Jumoorty, & Mercier, 1995) with satisfactory results. This checklist has shown to have both good content validity and satisfactory test-retest reliability, with the majority of kappa statistics showing moderate (0.41-0.60) to substantial (0.61-0.80) agreement (Oakley et al., 1986; Barris, Dickie, & Baron, 1988).

**Utility Questionnaire**

A feedback questionnaire was developed to determine if the questions in the Women’s Health Questionnaire were unclear, misleading or hard to understand, and whether the questionnaire was too long. Participants were asked to comment on whether any important areas regarding women’s health were omitted and were provided with space for additional comments or suggestions. Further, we asked participants if they experienced mental and or physical fatigue while completing the questionnaire.

*Data Analysis*
Demographic data and results from the utility questionnaire was analyzed using descriptive statistics. Level of agreement (intra-rater agreement) between a participant’s response during test session one and two was calculated using Kappa statistics (kappa coefficient) as all variables of interest were dichotomous and categorical data (Portney & Watkins, 2000). Values of kappa 0.61 and above were substantial, values between 0.41-0.60 moderate, and values less than 0.40 represented fair to slight agreement (Landis & Koch, 1977). Percent agreement was calculated for all variables. Data analysis was conducted using SAS version 9.1.

Results

Participants/Demographics

Thirteen women with a mean age of 41.4 (range from 25-53) and who sustained a moderate to severe brain participated in this study. The amount of time since injury ranged from two to 30 years with a mean of 13 years. Self-report loss of consciousness post injury ranged from 20 minutes to four months (kappa coefficient 0.45, 95% CI: 0.01 to 0.90). All women lived in the community in a non-institutional setting and eight women were living with a partner.

Women’s Health

Kappa coefficients for cervical cancer screening demonstrate substantial agreement of participant response (1.00, 95% CI: 1.00 to 1.00) as did breast cancer screening self-examination (0.67, 95% CI: 0.26 to 1.0) and mammogram participation (1.0, 95% CI: 1.0 to 1.0). Percent agreement for these items ranged from 85% to 100%.

The kappa coefficients for the menstrual cycle items ranged from 0.42 to 1.00 (Table 1). Percent agreement for these items ranged from 77% to 100%. The items measuring conception
Post-Print version

and pregnancy demonstrated moderate to substantial agreement with kappa coefficients ranging from 0.45 to 1.00. Percent agreement for these items ranged from 75% to 100%.

Four participants indicated being menopausal and were administered the Green Climacteric Scale. Items measuring menopausal symptoms produced a diverse range of Kappa coefficients (0.21 to 1.0) indicating fair to substantial agreement. Percentage agreement ranged from 75% to 100%.

Results from the thyroid survey indicated moderate to substantial agreement with kappa coefficients ranging from 0.40 to 1.0. Percent agreement ranged from 78% to 100% (Table 1).

*General health*

Results for these items showed moderate agreement for current health (0.53, 95% CI: 0.26 to 0.80) and substantial agreement for health prior to brain injury (0.77, 95% CI: 0.51 to 1.0). Percent agreement was 92% for current health and 77% for health prior to brain injury.

*Participation in Activities*

Table 2 shows the kappa and percent agreement for the Occupational Role Checklist. Kappa statistics for the role variables ranged from fair to substantial agreement, 0.33 to 1.00. Percent agreement for these items ranged from 38% to 100% with the majority exceeding 75%. The items with the lowest Kappa coefficients (0.16-0.33), Participant in Organizations and Hobbyist, may be reflective of the descriptive nature of these roles and therefore, the definition was not always consistent. In turn, the lower Kappa coefficients (0.43-0.58) for items Family member, Friend, and Home maintainer may have been influenced by the perceived and or actual loss of skills and abilities required to fulfill these roles.

*Utility questionnaire*
Table 3 presents the results obtained from the feedback questionnaire. Overall, the results demonstrate that the questionnaire was easy to use and follow, produced little mental or physical fatigue, and covered many areas seen as important to women who have sustained a TBI.

**Discussion**

The evaluation of the reliability and utility of a measure should be addressed in the context of the measurement tools purpose. This questionnaire was designed to be used as a screening tool for identifying women’s health issues post TBI. This is the first paper to examine rater agreement on a measure of women’s health in a population of women with TBI. Further testing of reliability and validity (e.g. menstrual disturbances could be recorded using diaries) are required but this is a preliminary step to provide valuable information for the assessment and treatment of women with TBI.

A preliminary step to collecting data is to determine who will provide the most reliable and valid information. We have shown that women with severe to moderate TBI can produce reliable results on self-report measures; this is similar to other studies using individuals with moderate to severe TBI (Hawthorne, Gruen, & Kaye, 2009; Wertheimer, Hanks, & Hasenau, 2008; van Baalen, Oding, van Woensel, van kessel, Roebroeck, & Stam, 2006). However, studies have shown physical and mental health outcome differences between individuals with mild, moderate, and severe TBI (Katz, Polyak, Coughlan, Nichols, & Roche, 2009; Anderson, Brown, Newitt, & Hoile, 2009; Willemse-van Son, Ribbers, Verhagen, & Stam, 2007) which indicates that sub-group analysis is recommend to delineate possible variation in response due to severity. In fact, two studies on women with TBI indicated that injury severity was a predictor of increased problems in menstruation and sexual function (Colantonio et al., 2009; Hibbard, Flanagan, Haddad, & Labinsky, 2000).
Higher levels of agreement were found for items more concrete in nature. In contrast, agreement was lower on more subjective items. For example, kappa coefficients for questions on whether the participant had a child prior to or after the TBI compare to whether they had difficulty conceiving a child were much higher (1.00 compare to 0.45). Additionally, magnitude of agreement for items concerned with symptoms of menses (subjective) was lower compared to whether participants were menstruating (0.42 compared to 1.00). This pattern is evident in the results of the Thyroid questionnaire which is based on symptoms. This discrepancy may be due to participants’ perception of symptom difficulty or importance which is not captured by the assessment questions. In turn, several of the symptom specific questions require memory and we acknowledge that individuals with TBI are more likely to have difficulty recalling events due to disturbances in long and short term memory (Bazarian, Cernak, Noble-Haeusslein, Potolicchio, & Temkin, 2009). It would be beneficial to capture corroborating data by using multiple methods of recall such as diaries, personal assistive devices, or family members. Further, items on the Role Checklist that had the lowest Kappa coefficient could be open to definition interpretation e.g. hobbyist or participant in an organization or could be a result of self perceived loss of the skills and abilities required for many day to day activities post TBI and not a true reflection of role change. These issues could be resolved by having strict item definition with examples, and instructions on distinguishing a true change in role (e.g. not able to be a friend) compared to a decrease in abilities (e.g. not able to be as social post TBI but able to be a friend in a different capacity).

Limitations

We recognize that this is a small sample and there were only four participants who were eligible to complete the Green Climacteric Scale. However, given that the investigation of the
effects of TBI on women’s health is in its infancy many studies are small and exploratory in nature. Our study was self-report with no validation from other sources. It is also important to note that due to the subjectivity of the items, one might expect variability within the testing period. Given the sensitive nature of some of the questions, individuals may have under-reported symptoms or issues. However, we attempted to control for this issue by having a female as the interviewer.

Conclusion

Although this study has limitations, there are important clinical implications that can be drawn from the data. As has been found in other studies (Valovich McLeod, Bay, Heil, & McVeigh, 2008; van Baalen, Odding, van Woensel, & Roebroeck, 2006; Ocampo, Dawson, & Colantonio, 1997) the highest level of agreement is found on concrete items and therefore, clinicians may wish to verify more subjective, symptoms based information using corroborating sources. Women with moderate to severe TBI reported no problems completing the questionnaire and felt that it reflected many of the important health concerns faced by women post TBI. However, future investigation involving women with mild TBI should be completed as these women may have different issues.

This study has begun the process of examining the psychometric properties of a women’s health questionnaire developed to specifically for women who have sustained a TBI. Though there was a range of Kappa coefficients across items, the majority indicated moderate to substantial agreement, providing support for the reliability of this questionnaire.
Acknowledgments

Funding for the study was obtained from the Canadian Institutes for Health Research, The Toronto Rehabilitation Institute, and a grant from the Ministry of Health and Long Term Care to the Toronto Rehabilitation Institute. Support for this study was given to JEH in a CIHR Fellowship Award and Strategic Training Fellowship in Health Care, Technology and Place (FRN: STP 53911).

Conflicts of Interest/Disclosures

All authors declare that the answer to the questions on competing interest are ‘No’ and therefore have nothing to declare. The funding source for this study had no involvement in the study design, analysis, interpretation of data, or writing of the report.
Reference List


