

Older Adult Mistreatment Risk Screening: Contribution to the Validation of a Screening Tool in a Domestic Setting

Jeannette M. Lindenbach, Sylvie Larocque, Anne-Marise Lavoie, and Marie-Luce Garceau
Laurentian University

RÉSUMÉ

Le dépistage du risque de maltraitance de la personne âgée vivant à domicile est difficile en raison de sa nature cachée. Afin de surmonter ce défi, un instrument valide évaluant systématiquement les facteurs de risque est requis. Un tel instrument, le "expanded-Indicators of Abuse" [e-IOA], fut originellement validé en hébreu dans un milieu hospitalier. Cette étude a contribué à la validation du "e-IOA" dans un nouveau contexte, celui d'un milieu communautaire anglophone en Ontario, Canada. L'étude s'est déroulée en deux phases: la validation du contenu et l'adaptation de l'instrument par des experts partout en Ontario, suivi de l'évaluation de la fidélité inter-juges par des infirmières. Cette adaptation, nommée le "Mistreatment of Older Adult Risk Factors", est un outil de dépistage clair, compréhensif et représentatif du risque de maltraitance à domicile. Cet instrument permettra aux professionnels œuvrant auprès des aînés d'être mieux outillés pour évaluer ce risque.

ABSTRACT

The hidden nature of older adult mistreatment renders its detection in the domestic setting particularly challenging. A validated screening instrument that can provide a systematic assessment of risk factors can facilitate this detection. One such instrument, the "expanded Indicators of Abuse" tool, has been previously validated in the Hebrew language in a hospital setting. The present study has contributed to the validation of the "e-IOA" in an English-speaking community setting in Ontario, Canada. It consisted of two phases: (a) a content validity review and adaptation of the instrument by experts throughout Ontario, and (b) an inter-rater reliability assessment by home visiting nurses. The adaptation, the "Mistreatment of Older Adult Risk Factors" tool, offers a comprehensive tool for screening in the home setting. This instrument is significant to professional practice as practitioners working with older adults will be better equipped to assess for risk of mistreatment.

Manuscript received: / manuscrit reçu : 26/09/11

Manuscript accepted: / manuscrit accepté : 30/11/11

Mots clés : aîné, maltraitance, risque, dépistage, domicile, aidant naturel, vieillissement

Keywords: older adult, mistreatment, risk, screening, domestic, caregiver, aging

Correspondence concerning this article should be addressed to

Jeannette Lindenbach, M.Sc.N.
École des sciences infirmières
Laurentian University
Ramsey Lake Road
Sudbury, Ontario, Canada, P3E 2C6
(jm_lindenbach@laurentian.ca)

Risk Screening Tool for Older Adult Mistreatment in the Domestic Setting

The mistreatment of older adults, which includes abuse and neglect, is a deeply troubling phenomenon which is predicted to escalate given the current health care and social climates (Baker & Heitkemper, 2005; Lachs & Pillemer, 2004). When those most vulnerable in society are mistreated or at risk of mistreatment, early detection is crucial. Unfortunately, many obstacles

may be present when the victim is an older adult living at home and being mistreated by a caregiver. Such complexity, especially within the sacred context of the family, renders identification inherently challenging for health care professionals (Lachs & Pillemer, 2004). Screening tools used in practice must be valid and reliable, thereby permitting early identification of mistreatment risk and possibly preventing further escalation (Shugarman, Fries, Wolf, & Morris, 2003). This

study contributes to the validation of a screening tool to detect risk of mistreatment of older adults; the tool is intended for use by English-speaking professionals visiting older adults in their homes.

Background

In Canada, although 78 per cent of Canadian older adults who require assistance live at home (Cranswick & Dosman, 2008), formal home care services have continued to suffer significant reductions, leaving family members with increased – sometimes unrealistic – responsibilities (National Advisory Council on Aging [NACA], 2006). Many studies have demonstrated that it is indeed family, including adult children or spouses, who are most frequently the perpetrators of mistreatment of the older adult (Choi & Mayer, 2000; National Research Council [NRC], 2003). For example, an American incidence study found that in almost 90 per cent of cases, the perpetrator was a family member (National Center on Elder Abuse [NCEA], 1998). More recently, a study in Turkey confirmed that the prevalence of mistreatment increased significantly when older adults lived with their spouse or child, who are the most frequent perpetrators (Kissal & Beser, 2011).

As isolated older adults may eventually come into contact with the health care system, health professionals who visit the home are in an ideal position to detect mistreatment risk (Allan, 2002; NCEA, 1998). Although experts believe that mistreatment of older adults is frequent enough to be faced routinely in health professionals' daily practice (Lachs & Pillemer, 2004), studies demonstrate that many fail to detect it (Fulmer, Guadagno, Dyer, & Connolly, 2004; Ortmann, Fechner, Bajanowski, & Brinkmann, 2001; World Health Organization/International Network for the Prevention of Elder Abuse [WHO/INPEA], 2002). To aid in detection, it is imperative to increase knowledge of its predisposing risk factors (Anetzberger, 2001; Baker, 2007). Although risk factors are not causal factors, they are associated with an increased probability of victimization, and the greater their presence in a family milieu, the higher the likelihood that mistreatment will occur (Wolf, 1997). The screening instrument therefore becomes a very important tool for the health professional as it provides for systematic and objective documentation of the phenomenon (Anetzberger, 2001).

Although many screening instruments have been developed, most have seen inadequate assessment of their psychometric properties and have significant limitations (Fulmer et al., 2004; Wolf, 2000). Important policy documents have reiterated the need for valid and reliable screening tools for mistreatment of older adults (NCEA, 1998; NRC, 2003; WHO/INPEA, 2002).

Several challenges have impeded progress in the development of valid and reliable instruments for screening of mistreatment. First, the divergence of definitions is believed to have rendered the development of psychometrically valid, reliable measures of mistreatment virtually impossible: the construct of older adult mistreatment must be very clear if it is to be measured (Kozma & Stones, 1995). A related problem has been the creation of instruments based on flawed or incomplete theories (Fulmer et al., 2004; Reis & Nahmiash, 1998; Shugarman et al., 2003). For example, if based on the theory of vulnerability, the instrument would only assess for risk factors of vulnerability in the older adult and therefore ignore the characteristics of both the caregiver and the sociocultural environment. This lack of comprehensiveness is problematic (NRC, 2003) insofar as theorists have now concluded that no one theory can explain mistreatment of older adults; consequently, assessment must include all risk factors of mistreatment (Ansello, 1996).

Some tools are ineffective because of their format. For example, self-report questionnaires that require accurate responses from older adults who suffer from cognitive or emotional difficulties are ineffective (Fulmer et al., 2004). This also applies to instruments that seek to obtain responses directly from a caregiver who may be incapable or unwilling to provide accurate responses (NRC, 2003). In addition, some measures have been adapted from other fields and, therefore, could neglect important factors contributing to the mistreatment of older adults (Fulmer et al., 2004; NRC, 2003). The complexity of mistreatment of older adults demands the use of a valid and reliable tool developed specifically to assess this concept.

Other more comprehensive tools have been developed but suffer from vaguely operationalized indicators (Nagpaul, 2001; NRC, 2003). Such ambiguity has led to poor specificity and sensitivity which are unacceptable when dealing with such a serious phenomenon. Finally, some screening instruments have combined two constructs: risk of mistreatment and actual mistreatment of older adults. The identification of those who are mistreated versus those who are at risk of mistreatment requires different methods of assessment, either case identification or screening (Kozma & Stones, 1995). As such instruments may lead to confusion because they lack conceptual clarity, this clear distinction is essential for effective screening (Anetzberger, 2001).

The Indicators of Abuse Screen [IOA], designed by Reis and Nahmiash (1998), is considered an important milestone in the assessment of older adult mistreatment (Wolf, 2000). Developed specifically for use by home visiting social service practitioners, the inclusion of caregiver and environment assessment in the IOA

does reflect the complexity of the older adult mistreatment (NRC, 2003; Wolf, 2000). A study of 341 cases supported the validity of the set of IOA indicators, which discriminated 84.4 per cent of "likely abuse" cases and 99.2 per cent of "likely nonabuse" cases (Reis & Nahmiash, 1998). Also, Chronbach alpha testing revealed an excellent internal consistency of 0.92 and 0.91 on two separate samples (Fulmer et al., 2004). However, the 27 items to be assessed require interpretation by the practitioner, and therefore its ability to objectively measure mistreatment has been questioned (Fulmer et al., 2004). Although the specificity of the instrument was 100 per cent, it achieved a sensitivity of 78.4 per cent (Reis & Nahmiash, 1998).

Some experts, commenting on the status of risk assessment of older adult mistreatment, have stated that little progress has been made in this respect because tools remain qualitative assessments based on clinical judgment (Wolf, 2000). As quantitative measurement is preferred for consistency when screening for mistreatment risk (Kozma & Stones, 1995), researchers have declared that improvements in measurement methods are urgently required for progress to occur in the study of older adult mistreatment (Fulmer et al., 2004). These instruments must have acceptable reliability and validity, be appropriate for the varying clinical contexts where mistreatment occurs, and either address screening or case identification (NRC, 2003). Furthermore, accurate and efficient measurement methods are essential given the important consequences of screening for older adults and their caregivers, the potential devastating effects on older adults and their families in the event of false positive or negative findings, and limited resources within the health and social service sectors (Lachs & Pillemer, 2004; NRC, 2003).

The 'Expanded Indicators of Abuse' Tool

In 2006, Cohen, Halevi-Levin, Gagin, and Friedman developed the *Expanded Indicators of Abuse* [e-IOA] tool. The comprehensive e-IOA includes all dimensions of the concept of mistreatment of the older adult. The e-IOA is based on Kosberg and Nahmiash's (1996) conceptual framework. Developed to address the complexity of older adult mistreatment, the framework includes four overlapping areas of risk identified in research on mistreatment of older adults. Like its predecessor, the Indicators of Abuse [IOA] (Reis & Nahmiash, 1998), the e-IOA is a semi-structured instrument that comprises 21 risk indicators. To rectify the subjectivity associated with the risk indicators in their original format in the IOA, these were operationalized in the e-IOA with sub-indicators based on the content of relevant academic literature in psychiatry and geriatric social work (Cohen, Halevi-Levin, Gagin, & Friedman, 2006). The operationalization standardizes

the tool and reduces its subjectivity as differences in professional interpretation are avoided (Nagpaul, 2001; NRC, 2003). Furthermore, the developers of the e-IOA proposed an instrument permitting early identification of high-risk older adults prior to the appearance of actual indicators, thereby addressing the problem of under-identification (Cohen et al., 2006). In doing so, prevention, identification of risk, and early intervention by health professionals were made possible (Cohen et al., 2006). Note that despite reference to indicators of abuse in the title of the tool, the authors considered abuse to also encompass neglect and measured risk factors, not indicators (Cohen et al., 2006).

Cohen et al. (2006) evaluated and confirmed content, criterion, and construct validity as well as inter-rater reliability. Content validity of the e-IOA was first assessed by an interdisciplinary team of health professionals in the fields of psychiatry, social work, and geriatrics. After a pilot test, sub-indicators found to be problematic by the interviewers were restructured. Criterion validity was assessed with a small group of older adults known to social workers as having been mistreated or not. The e-IOA correctly identified 92.7 per cent of those at high risk for mistreatment and 97.9 per cent of the non-mistreated cases. Discriminant validity was assessed by verifying if the e-IOA could differentiate older adults probably mistreated from those probably not mistreated on the basis of a list of evident signs of abuse. High agreement of 93 per cent was obtained between interviewers (Cohen et al., 2006).

The e-IOA was again evaluated in 2007 along with two other measures: direct questioning of older adults and an assessment conducted using an instrument of evident signs of mistreatment. These three measures were then compared. Because the e-IOA did capture a wider circle of older adults at risk (32.6%) than those who were actually mistreated (21.4%), the authors concluded that it could safely be used in other settings (Cohen, Halevi-Levin, Gagin, & Friedman, 2007). Caution was expressed, however, about relying on a cutoff score since, of those identified for evident mistreatment, only 74.4 per cent were also at high risk, on the basis of a cutoff of 1.7.

Research Purpose

The aforementioned studies were conducted in Hebrew and in a hospital context. Consequently, a new study was considered necessary to evaluate the validity and reliability of this instrument when used in an English-speaking community context (Fortin, 2006). The authors confirmed having translated the English instrument by a process of back translation as recommended in the literature (Behling & Law, 2000). However, the English version had not been validated (M. Cohen, personal

communication, August 1, 2008). Therefore, the purpose of this study was to adapt the expanded Indicators of Abuse screening tool and to contribute to the validation of this English adaptation in an Ontario community context. The specific research questions were: (1) What is the content validity of the e-IOA according to experts in mistreatment of the older adult? and (2) What is the inter-rater reliability of the adaptation of the e-IOA?

Research Method

This study assessed two critical psychometric properties: (a) content validity, which is the ability of the items of an instrument to adequately measure the concept chosen for study (Grant & Davis, 1997), and (b) inter-rater reliability, which is the degree of agreement or concordance between raters (Myers & Winters, 2002). As recommended by Waltz, Strickland, and Lenz (1991), we requested permission from Dr. Cohen, who developed the e-IOA, to validate the instrument. Also, permission was granted to revise the instrument as per experts' recommendations in order to reflect the specific community context – that is, Ontario communities (Burns & Grove, 2005). Furthermore, prior to commencing the study, we obtained ethical approval from the Laurentian University REB and that of the participating agencies.

Phase 1 – Determination of Content Validity

A preferred means of evaluating content validity is a review by a panel of experts (Waltz et al., 1991). In this study, a modified Delphi method defined as a “reactive Delphi” was used as it guided participants to “react” to the e-IOA in its present format and make recommendations to adapt the tool for use within the proposed context of this study (McKenna, 1994). This method consists of a series of sequential rounds of seeking knowledge and expertise from a panel of experts with the aim of sharing and building consensus (Mead & Moseley, 2001). Three rounds of Delphi were conducted over a seven-month period.

Purposive sampling guided the selection of the content experts. Expertise could stem from empirical knowledge originating from scientific study or experiential knowledge resulting from the work of clinicians in the field (Anetzberger, 2005). Inclusion criteria for this study therefore consisted of knowledge regarding mistreatment of the older adult and, more specifically, the clinical experience of applying this knowledge on a practical level within the community setting (Baker, Lovell, & Harris, 2006; Petry, Maes, & Vlaskamp, 2007). Experts throughout Ontario in both rural and urban community settings were invited to serve as panellists, providing a heterogeneous sample for all geographical

areas of the study (Hardy et al., 2004; Mead & Moseley, 2001). When utilizing the Delphi process, the quality of the judges is of greater importance than their quantity (Grant & Davis, 1997; Zink & Fisher, 2007). A total of seven experts who met the inclusion criteria agreed to participate.

Experts received a standardized detailed information package that included the conceptual framework to ensure that the instrument continued to rest upon its theoretical framework (Davis & Grant, 1993). As well, a new Content Review Questionnaire was created for each new round based on the data previously gathered and guided a structured assessment of each sub-indicator (Springer, Abell, & Hudson, 2002).

In the first round, panelists judged the elements of a content review: representativeness, clarity, and comprehensiveness (Grant & Davis, 1997). Representativeness of the content coverage was assessed to determine if items reflected, sampled, and measured the construct of risk of older adult mistreatment in the home setting (Berk, 1990). Responses ranged from (1) “*item is not representative of risk of older adult mistreatment in the home setting and should be removed*” to (4) “*item is very representative and should be kept*”. Item clarity indicated if items were clear, well-written, and distinct (DeVellis, 1991). Responses were coded as (0) “*item is not clear and/or contains inappropriate wording and/or is not easily discriminated from other items and requires revisions to be clear*” or (1) “*item is clear, without bias and easily discriminated from other items*”.

Panellists' recommendations and descriptive feedback were summarized and used to improve either clarity or representativeness. Alternatively worded sub-indicators were then constructed for ranking in Round 2 (Grant, Kinney, & Guzzetta, 1990; Hasson, Keeney, & McKenna, 2000). The comprehensiveness of the instrument was assessed, and additional sub-indicators were recommended by the judges to ensure that the included items were sufficient to represent the total content domain (Lynn, 1986). These were returned to subsequent rounds for consideration by all (Parks, Cintas, Chaffin, & Gerber, 2007). Finally, experts were asked to rate the appropriateness of the response scales as the original e-IOA contained two different rating scales.

As transparency between rounds of a Delphi review is recommended (Bowles, 1999), feedback consisted of simple statistical summaries (Keeney, Hasson, & McKenna, 2006). Using SPSS, we calculated the median and semi-quartile range for representativeness (Bello & Singh, 2004; Statistics Canada, 2009). As clarity was evaluated with a dichotomized 0 or 1 scale, the mode was the most appropriate measure of central tendency.

Round two began with a ranking process (Evans, 2005). Judges were provided with the original sub-indicator from the e-IOA along with the alternatively worded sub-indicators constructed from Round 1 feedback and were asked to rank items from 1 to 3 with 1 being “the most comprehensive, representative and clear choice”. The average rank was calculated indicating the preferred wording for that sub-indicator (Siegel & Castellan, 1988). To measure the degree of association among these rankings, we applied the Kendall’s coefficient of concordance W , a measure particularly useful when calculating agreement among rankings (Siegel & Castellan, 1988). This value ranges from 0 (no agreement) to 1 (complete agreement). Second, for the recommended additions from Round 1, we also calculated the percentage of agreement. We retained only those items with acceptable agreement (60%) for reassessment.

In the third and final round, panelists were asked to judge if the preferred sub-indicators from Round 2 as well as the recommended deletions and additions from the previous rounds should be accepted. Percentage agreement was calculated as a measure of consensus reached after this final round. If the judges were not in agreement to modify, add, or delete the sub-indicator, we instructed them to reject the recommendations with supporting comments. There is a lack of guidance in the literature in regards to the determination of importance in Delphi studies (Hardy et al., 2004). Rounds should continue as required until consensus is reached or until returns begin to diminish (Hasson et al., 2000). For this study, three rounds were conducted with the occurrence of both events: acceptable consensus and diminished returns.

Phase 2 – Assessment of Inter-Rater Reliability

The assessment of inter-rater reliability was achieved through the application of the adapted tool to various paper case studies. We selected this method versus testing in a natural setting because of the sensitive nature of older adult mistreatment research (i.e., deeply personal, may involve deviance or social control, and may be threatening for older adults and their caregivers; Dresser, 2003). As well, to have proceeded with actual vulnerable older adults in the community would have carried numerous ethical challenges at this early stage since the findings could have had potentially devastating legal, financial, and social consequences for the older adult and caregiver (Lachs & Pillemer, 2004; Waltz et al., 1991). Others have successfully used a similar paper design when dealing with sensitive research with vulnerable participants (Endacott, Clifford, & Tripp, 1999; Selwood, Cooper, & Livingston, 2007).

For this study, we chose selected case studies on the basis of their realistic illustration of risk characteristics of the victim, the perpetrator, and the family context. These studies were obtained from a resource guide prepared by the Ontario Network for the Prevention of Elder Abuse (ONPEA), the agency mandated to provide expert consultation on older adult mistreatment at the community level for professionals, volunteers, and seniors (ONPEA, 2006). Therefore, we felt that these case studies reflected accurate and current knowledge of mistreatment of older adults in the community. The case studies were meant to be explicit and clear but not detailed enough to render rating totally predictable or to lead the panel (Endacott et al., 1999). Various forms of mistreatment of older adults and risk factors were described, and we obtained permission for the use of each study from the ONPEA.

Phase 2 drew on a convenience sample of home visiting nurses employed by a nursing agency in a north-eastern Ontario city. The participation of nurses visiting older adults in their homes (versus hospital nurses) was essential as this adaptation and validation was geared to the domestic setting (Springer et al., 2002). Selection criteria consisted of registration with the College of Nurses of Ontario as a registered nurse (RN) or a registered practical nurse (RPN) as evidenced by active employment within the nursing agency. Although these two categories have different levels of foundational knowledge, we included both because they may function autonomously in the home setting (College of Nurses of Ontario, 2009) and therefore could encounter instances of mistreatment of older adults. The ability to understand English was also required.

Initial recruitment was carried out using a formal mailed invitation to each nurse followed by a general voice mail message. As the first message was only successful in attracting two nurses, we sent reminders again informing nurses of this opportunity. Washington and Moss (1988) have recommended a minimum of 10 subjects to adequately assess inter-rater reliability. Despite the numerous strategies we attempted, only six registered nurses attended and completed the packages. Although a larger sample of nurses would have been preferable, the size of the sample is considered to be of less importance than the number of items being measured (Lefrançois, 1992). Whereas the instrument consisted of 95 items to be assessed with 10 scenarios by six nurses, we tested each sub-indicator 60 times, resulting in 5,700 observations. This large number of sub-indicators, as well as the nonparametric statistical technique chosen and the ability to test significance of the latter with small samples, satisfied the consulted statistician as to the small sample size.

Due to the limitations imposed by traveling, the training of home visiting nurses was offered in various geographical locations and at different times. Consistency was maintained for the duration of the two-hour sessions with the individual nurses as indicated in the literature (Springer et al., 2002; Washington & Moss, 1988). The session began with a Microsoft PowerPoint presentation which consisted of the dynamics of older adult mistreatment, the benefits of screening tools, the adapted instrument, and its conceptual underpinnings (Washington & Moss, 1988). Packages were then distributed containing instructions: 10 case studies with 10 copies of the adapted screening tool. A case demonstration was completed by the researcher, and nurses were then directed to complete their package independently without discussion with other participants, which minimized threats to internal validity (Burns & Grove, 2005). We modified the order of the paper scenarios for each participating nurse in order to prevent maturation with the last scenarios as nurses became more experienced but more fatigued in completing the instrument (Burns & Grove, 2005).

Descriptive statistics were used to describe the sample of participating nurses.

We measured agreement among these six nurses using the Kendall's coefficient of concordance W , a nonparametric measure strongly recommended in studies of inter-judge reliability and appropriate for ordinal measures (Siegel & Castellan, 1988). For each case, the nurses rated the 95 sub-indicators according to a five-item Likert-scale ranking. This permitted the calculation of the Kendall's W for each of these sub-indicators. Afterwards, the values for that specific sub-indicator for all 10 cases were added and their mean was calculated representing the performance of that specific sub-indicator in the 10 cases by the six judges. In testing the significance of W , critical values have been tabled for an N of 5 (rating options) and a k of 6 (number of nursing raters). For this concordance to be significant at the $\alpha = .01$ level, the observed mean Kendall's W was required to be .489 or larger (Siegel & Castellan, 1988, p. 365).

Results

Phase 1 – Content Validity Assessment and Adaptation of the E-IOA

The panellists had from 2 to 30 years of professional experience. From the perspective of clinical knowledge and expertise, 43 years was the cumulative number of years of experience with older adult mistreatment. The minimum level of educational achievement was a university baccalaureate degree with three experts being masters' prepared. Both nursing and gerontology

backgrounds were reported. Examples of activities demonstrating clinical expertise included (a) directly responding to older adult mistreatment calls in the community; (b) case management of mistreatment of older adults within families; (c) mistreatment education; (d) local, provincial, and national collaboration to develop strategies and response programs for mistreatment; (e) service on various advocacy groups for the older adult; and (f) expert consultation. Moreover, research, authorship, and policy work were listed among this expert panel's many attributes.

Data collection consisted of three subsequent rounds of Delphi review. All seven experts completed the first round. However, only five experts completed the study. Reasons offered for withdrawal consisted of (a) workload issues not related to the study and (b) personal matters.

Overall, 20 per cent of the items (18/90) were accepted in Round 1 as they met the following three criteria: (a) a median of 4.0 for representativeness, (b) a mode of 1.0 for clarity, and (c) without need for improvement as evidenced by a lack of recommendations for modifications. These are indicated by an asterisk in the Table 1, Round 1. Because these were accepted in this round, they did not reappear in Round 2 in order to prevent the fatigue which may result from a prolonged study (Hasson et al., 2000). In regards to comprehensiveness, experts recommended 13 additional sub-indicators which were forwarded to the next round for further consideration.

In Round 2, the experts ranked the 72 remaining items along with the recommended modifications and the average rank was determined. To measure agreement among the raters, we calculated a Kendall's coefficient of concordance W . Only 10 items achieved elevated agreement (.760 to 1) and were therefore accepted in this round (see items marked by a double asterisk in Table 1, Round 2). The remaining 62 items did not achieve high consensus among the experts as evidenced by Kendall's W being under .760, and so we retained them for reassessment. Of the recommended additions, we retained only those with acceptable percentage of agreement (60%) for reassessment (see Table 1).

In the third and final round, the percentages of agreement were as follows: 49 sub-indicators were agreed upon by 100 per cent of the judges, 20 sub-indicators were agreed upon by 80 per cent of the judges, and three sub-indicators received agreement by 60 per cent of the judges. Although these last three sub-indicators did not achieve high agreement, two factors motivated our retaining them in the adapted instrument. First, in all three instances, the sub-indicators were few in number and therefore to remove it would leave very little to measure the indicator. Second, if the item was

Table 1: Results: Adaptation of "expanded Indicators of Abuse" (e-IOA) Tool throughout Phase 1 – Modified Delphi and Phase 2 – Interrater Reliability of "Mistreatment of Older Adult Risk Factors" (MOARF) Tool

Original e-IOA indicators and sub-indicators	Phase 1 – Modified Delphi				Phase 2				
	Round 1		Round 2		Round 3		Inter-rater reliability of MOARF sub-indicators		
	Representativeness	Clarity	Modification of sub-indicator and recommended additions in MOARF	Average rank	Kendall W	Percentage of agreement	Mean Kendall W		
	Mdn	SQR	Mode	SQR					
1. Behavior Problems (Caregiver)									
Has outbursts	3.0	0.50	0.0	0.00	Has verbal or physical outbursts towards older adult	1.60	0.160	100	.6667
Committed to his/her responsibilities	4.0	0.62	1.0	0.12	Not committed to his/her caregiving responsibilities	1.60	0.120	80	.4653***
Engages in conflicts with family, friends, or neighbours	4.0	0.50	1.0	0.00	Engages in verbal or physical conflict with family, friends, or neighbours**	1.40	0.760	Accepted Round 2	.8003
Has poor personal functioning	3.0	1.00	0.0	0.50	Lack of competency with personal life skills	1.60	0.160	100	.6160
Has poor family functioning	3.0	0.50	0.0	0.50	Behaviour contributes to poor family functioning	1.40	0.280	80	.4792***
Blames external forces for his/her situation	4.0	0.50	1.0	0.00	Blames external forces for his/her situation	1.40	0.040	100	.8264
Angry and bitter towards his/her environment	4.0	0.50	1.0	0.00	Angry and hostile towards his/her environment	1.40	0.040	100	.6383
2. Emotional/Cognitive Difficulties (Caregiver)									
Expresses helplessness*	4.0	0.50	1.0	0.00	Expresses helplessness	Accepted Round 1			.8403
Expresses dependence	4.0	0.00	1.0	0.00	Expresses dependence	1.20	0.360	80	.9028
Irritable or nervous	4.0	0.50	1.0	0.00	Easily provoked/ angered or nervous	1.20	0.360	100	.7586
Thinks that people harass him/her, plot against him/her, or are unfair	4.0	0.00	1.0	0.00	Thinks that people harass him/her, plot against him/her, or are unfair	1.40	0.040	100	1.000
Exaggerated, inappropriate, or fluctuating emotional reactions	4.0	0.00	1.0	0.00	Exaggerated, inappropriate, or fluctuating emotional response**	1.00	1.000	Accepted Round 2	.7847
Paranoid ideations	4.0	0.50	1.0	0.00	ELIMINATE "Paranoid" as similar to "Distorted perception (exaggerated or not reality-based) of self, other people, and events"***	1.00	1.000	DELETED	
Shallow emotional expression	3.0	1.00	1.0	0.50	Shows little emotional connection towards older adult	1.60	0.160	80	.5625
Distorted perception (exaggerated or not reality-based) of self, other people, and events*	4.0	0.00	1.0	0.00	Distorted perception (exaggerated or not reality-based) of self, other people, and events	Accepted Round 1			.8334
Depressed (moody, lack of motivation and of interest, low self-esteem, low energy, sleep disturbance, or poor appetite or overeating)	4.0	0.00	1.0	0.00	Depressed (sad, lack of motivation and of interest, low self-esteem, low energy, sleep disturbance, or poor appetite or overeating)	1.40	0.040	100	1.000

Continued

Table 1 . Continued

Original e-IOA indicators and sub-indicators	Phase 1 – Modified Delphi						Phase 2	
	Round 1			Round 2			Round 3	
	Representativeness	Clarity	Modification of sub-indicator and recommended additions in MOARF	Average rank	Kendall W	Percentage of agreement	Inter-rater reliability of MOARF sub-indicators	Mean Kendall W
Mdn	SQR	Mode	SQR					
Anxious (fears, vigilance, automatic hyperactivity, motor tension, restlessness)*	4.0	0.00	1.0	0.00	0.00	Anxious (fears, vigilance, automatic hyperactivity, motor tension, restlessness)	Accepted Round 1	1.000
Is confused or disoriented	4.0	0.50	1.0	0.50	0.50	Is confused or disoriented to time, person, or place**	1.00	1.000
Has memory problems	4.0	0.50	1.0	0.50	0.50	Has memory problems which interfere with care giving tasks	1.20	1.000
Is mentally retarded (mark degree of retardation)	2.0	1.00	0.0	0.50	0.50	Low cognitive functioning	1.40	1.000
Alcohol abuse	4.0	0.00	1.0	0.00	0.00	Alcohol abuse	1.40	.8681
Drug abuse	3.0	1.00	1.0	0.50	0.50	Abuses prescription, street or over the counter drugs	1.60	.8542
						ADD Gambling addiction	100	.9653
						ADD Communication barrier (language barrier, aphasia, stroke, developmental disability)	80	1.000
3. Financial Dependence (Caregiver)								
Has satisfactory means for living	3.0	1.00	1.0	0.12	0.12	Has unsatisfactory means for living	1.40	.7569
Receives financial support from older adult	4.0	0.00	1.0	0.00	0.00	Receives financial support from the older adult	1.20	.6459
						ADD Holds older adult's power of attorney for property	2.20	.9653
						ADD Loss of employment	2.60	.8472
4. Unrealistic Expectations (Caregiver)								
Has unrealistic perception of the older adult's physical or mental situation*	4.0	0.00	1.0	0.00	0.00	Has unrealistic perception of the older adult's physical or mental situation	Accepted Round 1	.6042
Disregards older adult's functional disabilities*	4.0	0.00	1.0	0.00	0.00	Disregards older adult's functional disabilities	Accepted Round 1	.5973
5. Caregiving Reluctancy (Caregiver)								
Provides the required physical treatment (dressing, washing)	4.0	0.00	1.0	0.50	0.50	Does not provide the required assistance with ADLs or IADLs	1.40	.5764
Understands the essence of the treatment	4.0	0.00	1.0	0.50	0.50	Does not understand the essence of the treatment	1.40	.6041
Is confident in providing the needed treatment	4.0	0.00	1.0	0.00	0.00	Is not confident in providing the needed treatment	1.25	.7084

Treats the older adult with care and goodwill	4.0	0.00	1.0	0.00	0.00	1.40	0.040	100	.6484
Does not treat the older adult with care and goodwill								100	.9653
ADD Disrespectful of older adult's cultural preferences								80	.7431
ADD Lack of/poor external supports (community, siblings)								100	.8056
ADD Knowledge deficit regarding normal aging and/or disease process									
6. Marital/Family Conflicts (Caregiver)									
Is involved in conflicts and disagreements within the family *	4.0	0.00	1.0	0.00	0.00	Accepted Round 1			.6597
Expresses guilt or anger, and bitterness towards the family*	4.0	0.00	1.0	0.00	0.00	Accepted Round 1			.7484
Unspoken family secrets exist	3.0	1.00	1.0	0.50	0.040	1.80	0.040	100	.8612
Family copes with special problems or trauma (disabled child, recent loss)	4.0	0.50	1.0	0.00	0.360	1.20	0.360	100	.9097
7. Poor Current Relationships (Caregiver)									
Has close friends	4.0	0.00	1.0	0.00	0.040	1.40	0.040	80	.9653
The existing relationships are conflictual	4.0	0.00	1.0	0.00	0.840	1.20	0.840	Accepted Round 2	.5834
Breaks his/her relationships with others	4.0	1.00	1.0	0.50	0.040	1.40	0.040	100	.9306
8. Is a Blamer (Caregiver)									
Attributes the cause of his/her difficulties to others	4.0	0.00	1.0	0.00	0.360	1.20	0.360	100	DELETED
Blames others for his/her situation*	4.0	0.00	1.0	0.00	Accepted Round 1				.8403
9. Emotional Dependency (Caregiver)									
Has difficulty making decisions*	4.0	0.00	1.0	0.00	Accepted Round 1				.8334
Needs full-scale daily support	4.0	1.00	1.0	0.50	0.360	1.20	0.360	60	1.000
Needs attention and approval from his/her surroundings	3.0	1.50	1.0	0.50	0.040	1.40	0.040	80	.9653
10. Lacks Understanding (Caregiver)									
Understands the medical condition	4.0	0.00	1.0	0.00	0.040	1.40	0.040	100	.6042
Understands the cognitive condition	4.0	0.00	1.0	0.00	0.040	1.40	0.040	100	.7431
11. Behavior Problems (Older adult)									
Has outbursts	3.0	0.00	0.0	0.50	0.280	1.60	0.280	80	.8889
Committed to his/her responsibilities	3.0	1.00	0.0	0.50	0.360	1.40	0.360	100	DELETED

Continued

Table 1. Continued

	Phase 1 – Modified Delphi				Round 2			Round 3		Phase 2
	Round 1				Modification of sub-indicator and recommended additions in MOARF	Average rank	Kendall W	Percentage of agreement	Mean Kendall W	
	Representativeness	Clarity	Mdn	SQR						Inter-rater reliability of MOARF sub-indicators
Engages in conflicts with family, friends, or neighbours	4.0	1.0	0.50	0.00	Engages in conflicts with family, friends or neighbours	1.25	0.250	100	.9028	
Has poor personal functioning	3.0	1.0	0.50	0.50	Has poor coping mechanisms	1.20	0.360	100	.5348	
Has poor family functioning	3.5	0.0	0.62	0.50	Has poor interpersonal relationships within family	1.80	0.040	100	.4654***	
Blames external forces for his/her situation	4.0	1.0	0.50	0.00	Blames others for current situation	1.40	0.040	100	.7918	
Angry and bitter towards his/her environment*	4.0	1.0	0.00	0.00	Angry and bitter towards his/her environment	Accepted Round 1			.7709	
12. Social Isolation (Older adult)										
Attends activities outside the house	4.0	1.0	0.00	0.00	Does not attend activities outside the house (medical appointments)	1.40	0.040	100	.6320	
Meets friends and associates	4.0	1.0	0.50	0.50	Does not meet people outside the home (church, social gatherings)	1.80	0.040	100	.5972	
People come to meet him/her	4.0	1.0	0.50	0.50	Family or friends do not visit him/her in home	1.60	0.120	80	.7640	
					ADD Caregiver restricts access (family, friends, home health services)			100	.8506	
13. Emotional/Cognitive Difficulties (Older adult)										
Expresses helplessness*	4.0	1.0	0.00	0.00	Expresses helplessness	Accepted Round 1			.5139	
Expresses dependence*	4.0	1.0	0.00	0.00	Expresses dependence	Accepted Round 1			.5519	
Irritable or nervous	4.0	1.0	0.50	0.00	Lack of eye contact, reluctant to answer	1.40	0.000	100	.8889	
Thinks that people harass him/her, plot against him/her, or are unfair	4.0	1.0	0.50	0.00	Thinks that people harass him/her, plot against him/her, or are unfair	1.40	0.040	100	.9306	
Exaggerated inappropriate, or fluctuating emotional reactions	4.0	1.0	0.00	0.00	Exaggerated, inappropriate, or fluctuating emotional reactions	1.40	0.040	100	.8889	
Paranoid ideations	4.0	1.0	1.00	0.00	ELIMINATE "Paranoid ideations" as similar to "thinks that people harass him/her, plot against him/her, or are unfair**"	1.00	1.000	Accepted Round 2	DELETED	
Shallow emotional expression	3.0	1.0	1.00	0.50	Shows little emotional connection	1.40	0.280	80	.9053	
Distorted perception (exaggerated or not reality-based) of self, other people, and events*	4.0	1.0	0.00	0.00	Distorted perception (exaggerated or not reality-based) of self, other people, and events	Accepted Round 1			.8403	

Depressed (moody, lack of motivation and of interest, low self-esteem, low energy, sleep disturbance, or poor appetite or overeating)	4.0	0.00	1.0	0.00	1.40	0.040	100	.5757
Anxious (fears, vigilance, automatic hyperactivity, motor tension, restlessness)*	4.0	0.37	1.0	0.00	Accepted Round 1			.4583***
Is confused or disoriented	4.0	0.50	1.0	0.50	1.00	1.000	Accepted Round 2	.8473
Has memory problems	4.0	0.50	1.0	0.50	1.20	0.480	100	.8403
Is mentally retarded (mark degree of retardation)	2.0	1.50	0.0	0.50	1.20	0.520	100	.8195
Alcohol abuse	4.0	0.00	1.0	0.00	1.40	0.040	100	.9306
Drug abuse	4.0	0.50	1.0	0.50	1.80	0.040	100	1.000
					ADD Communication barrier (language, sensory, aphasia, stroke, developmental disability)		100	.8334
14. Financial Dependence (Older adult)								
Has satisfactory means for living	4.0	0.50	1.0	0.00	1.40	0.040	60	.5070
Receives financial support from the caregiver*	4.0	0.00	1.0	0.00	Accepted Round 1			.8959
15. Unrealistic Expectations (Older adult)								
Has unrealistic perception of his/her own physical or mental situation	4.0	0.00	1.0	0.50	1.20	0.360	100	.8265
Disregards own functional disabilities	4.0	0.50	1.0	0.00	1.40	0.040	100	.8889
					ADD Unrealistic expectations of caregiver (unrealistic multiple requests)		80	.9097
16. Marital/Family Conflicts (Older adult)								
Is involved in conflicts and disagreements within the family*	4.0	0.00	1.0	0.00	Accepted Round 1			.6250
Expresses guilt or anger, and bitterness towards the family*	4.0	0.00	1.0	0.00	Accepted Round 1			.6736
Unspoken family secrets exist	3.0	1.50	0.0	0.50	1.80	0.040	100	.8263
Family copes with special problems or trauma (disabled child, recent loss)	4.0	0.50	1.0	0.00	1.40	0.040	100	.6041
17. Poor Current Relationships (Older adult)								
Has close friends	4.0	0.00	1.0	0.00	1.40	0.040	80	.7360

Continued

Table 1. Continued

Original e-IOA indicators and sub-indicators	Phase 1 – Modified Delphi				Phase 2			
	Round 1		Round 2		Round 3		Phase 2	
	Representativeness	Clarity	Modification of sub-indicator and recommended additions in MOARF	Average rank	Kendall W	Percentage of agreement	Inter-rater reliability of MOARF sub-indicators	
	Mdn	SQR	Mode	SQR			Mean Kendall W	
The existing relationships are conflictual	4.0	0.00	1.0	0.00	The existing relationships are conflictual**	1.40	0.760	.6461
Breaks his/her relationships with others	3.5	1.00	1.0	0.50	Does not maintain relationships with others	1.40	0.040	.9028
18. Is a Blamer (Older adult)								
Attributes the cause of his/her difficulties to others	4.0	0.00	1.0	0.00	ELIMINATE " Attributes the cause of his/her difficulties to others" as too closely related to "Blames others for his/her situation"***	1.00	1.000	DELETED
Blames others for his/her situation	4.0	0.00	1.0	0.00	Blames others for his/her situation**	1.00	1.000	.7640
19. Emotional Dependency (Older adult)								
Has difficulty making decisions	4.0	0.00	1.0	0.00	Has difficulty making decisions	1.40	0.040	.5556
Needs full-scale daily support	4.0	0.50	1.0	0.50	Total emotional dependence	1.80	0.040	.4722***
Needs attention and approval from his/her surroundings*	4.0	0.50	1.0	0.00	Needs attention and approval from his/her surroundings	Accepted Round 1		.8612
20. Lacks Social Support (Older adult)								
Has friends/ associates or family members	4.0	0.00	1.0	0.00	Does not have friends/ associates and or family members	1.40	0.040	.7848
Friends/ associates extending help	4.0	0.00	1.0	0.00	Does not have friends/associates providing support	1.40	0.040	.5486
Friends/ associates show interest in his/her condition	4.0	0.00	1.0	0.00	Friends/associates do not show interest in his/her condition	1.40	0.040	.7848
21. Lacks Understanding (Older adult)								
Understands the medical condition	4.0	0.00	1.0	0.00	Does not understand his/her own medical condition	1.40	0.040	.7431
Understands the cognitive condition	4.0	0.00	1.0	0.00	Does not understand his/her own cognitive condition	1.40	0.040	.6875

* = represents three criteria met for acceptance in Round 1: (a) a median of 4.0 for representativeness, (b) a mode of 1.0 for clarity, and (c) without need for improvement as evidenced by a lack of recommendations for modifications

** = items that achieved elevated agreement (.760 to 1) and accepted in Round 2

*** = items which did not achieve sufficient interrater agreement in Phase 2

Mdn = Median

MOARF = Mistreatment of Older Adult Risk Factors

SQR = semi-quartile range

problematic, it was felt that this would be triggered by the home visiting nurses trialing the instrument (see Table 1, Round 3).

Over the course of the three rounds of Delphi, 60 per cent of the original e-IOA items were revised. Five of the original items were deleted and 10 new items were added, resulting in a 95-item instrument. The experts agreed to keep one of the rating scales from the original tool which measured frequency (very often, often, not so often, very seldom, not at all, not possible to receive information). As well, a new title was deemed essential in order to correctly reflect the measurement of risk versus actual mistreatment and to be comprehensive in measuring both abuse and neglect as represented in the term mistreatment. The adaptation was entitled the "Risk Factors of Elder Mistreatment" screening tool. To be respectful of culture, we replaced "elder" by "older adult". Therefore, the tool is now entitled "Mistreatment of Older Adult Risk Factors".

Phase 2 – Inter Rater Reliability Assessment

The adapted instrument was then tested for inter rater reliability by home visiting nurses. The years of general professional experience of the participating nurses ranged from 1 month to 22 years in the home visiting community setting. Half were college graduates while the other half was university prepared. Four of the nurses had never received any training on the topic of mistreatment of the older adult while two had attended one in-service session.

Ninety of the 95 sub-indicators achieved satisfactory agreement among the nurses (see Table 1, Phase 2). The five sub-indicators which did not achieve the critical value of $W = .489$ for $\alpha = .01$ significance levels are marked by a triple asterisk.

Discussion

This study has contributed to the validation of a screening instrument to detect the risk of mistreatment of older adults in the domestic setting. The Delphi method was successful in harnessing the experiential knowledge of experts, representing various geographical regions of Ontario, in the field of older adult mistreatment. Some changes to the instrument, including modifications, deletions, and additions, will be discussed based on the three structural elements of clarity, representativeness, and comprehensiveness.

The assessment of clarity of the original e-IOA sub-indicators revealed some important flaws such as inappropriate wording. For example, 71.4 per cent of the judges rated the sub-indicator "Is mentally retarded (mark degree of retardation)" as being unclear and considered it to be offensive. The literature indicates

that scale items should never be offending to the patient or professional (Berk, 1990; Springer et al., 2002). The content of the item was changed to "low cognitive functioning". Improvements to clarity were also required for those sub-indicators that applied to both the caregiver and the older adult. For example, the sub-indicators "Marital/Family Conflicts – Family copes with special problems or trauma (disabled child, recent loss)" were identical for both the caregiver and the older adult. As this instrument is used within a familial context, the clarification of these sub-indicators better reflects current societal stressors facing the modern family (Levine, 2003). Such stressors are illustrated in the new wording chosen by the judges: "Caregiver faced with special problems or trauma (disabled child, divorce, recent loss)" and for the older adult, "Older adult faced with special problems or trauma (recent loss, death of spouse, sale of home)". Such specificity contributes to accurate measurement of the indicator. Clarity of the sub-indicator also meant that it was easily discriminated from other sub-indicators (Grant et al., 1990). Lack of clarity resulted in the deletion of four sub-indicators.

The assessment of representativeness resulted in the deletion of the sub-indicator "Older Adult Behavior Problem – Committed to his/her obligations". The following judges' comments illustrate the lack of representativeness of this sub-indicator: "Does not fit. What obligations? Excessively loyal to family? Is this related to finances?" It is evident that such an item could create confusion among evaluators because it was not representative of "Older adult behavior problems" contributing to mistreatment risk. Another example is that of "Unspoken family secrets exist" which was changed to "Past history of family violence, conflict, or trauma". Such wording correctly represents the risk of older adult mistreatment within the family context as studies have indeed identified a poor pre-morbid relationship as a significant contributor to mistreatment of the older adult (Cooney, Howard, & Lawlor, 2006).

Lack of comprehensiveness has been a serious flaw of past instruments. To improve this structural element, we sought recommendations for new sub-indicators in the first two rounds and fed them back into subsequent rounds for consideration by all judges. High levels of agreement among judges justified the addition of 10 new sub-indicators. First, "Communication barrier (language barrier, aphasia, stroke, developmental disability)" was added for the caregiver. Its inclusion is supported in the literature because language comprehension and memory deficits have been studied in caregivers who mistreat their older adult care receivers (Miller et al., 2006). The experts then recommended adding a similar sub-indicator for the older adult inasmuch as speech and language impediments may impair

ability to communicate needs or concerns to caregivers, family, and service providers or to disclose actual mistreatment (Cooney et al., 2006; Peguero & Lauck, 2008).

“Gambling addiction” was then added. Although we do not yet fully understand the impact of gambling addiction within the caregiving relationship, studies are beginning to study its emotional and financial consequences. The research literature does speak of dependence or “other special problems” (Hwalek, Goodrich, & Quinn, 1996, p. 129). The judges then added “Holds older adult’s power of attorney for property” and “Loss of employment” to address financial dependency risk. This is supported in the literature as researchers Hwalek et al. (1996) spoke of “parasitic or opportunistic behaviors” in regards to the older adult’s financial assets (p. 131). Also, research has confirmed financial mistreatment to be the most prevalent type of mistreatment of older adults in Canada (Podnieks, Pillemer, Nicholson, Shillington, & Frizzell, 1990), and it is frequently perpetrated by the victims’ adult children (NCEA, 1998). Next, the experts recommended adding the sub-indicator “Disrespectful of older adult’s cultural preferences” which reflects the recommendations that instruments should be sensitive to culture (Fulmer et al., 2004). Although current study findings are inconclusive, factors studied in the literature include values of loyalty, interdependence, and the intergenerational gap resulting from contradictory traditional and modern cultural values (Litwin & Zoabi, 2004).

Although the original e-IOA addressed lack of social support for the older adult, it did not do so for the caregiver. The new sub-indicator “Lack of/poor external supports (community, siblings)” reflects the “reciprocal relationship” between social support for the older adult, the caregiver and mistreatment which explains that health support networks for the caregiver mitigate the risk to the dyad (Peguero & Lauck, 2008, p. 65). In the present health care system in which family caregiving is a societal expectation and fiscal restraint a priority, caregiver needs and supports are seldom addressed (NACA, 2006). A perception of inadequate support has been demonstrated to be significantly associated with higher levels of burnout and actual mistreatment of older adults (Almberg & Granfström, 1997). Also added was “Knowledge deficit regarding normal aging and/or disease process” which the literature supports as a risk factor (Hwalek et al., 1996) and which reflects current reality as care continues to be delegated to families despite their lack of knowledge and perhaps complete unsuitability to provide care (Wiles, 2003).

Next, the experts recommended the addition of the sub-indicator “Caregiver restricts access (family, friends, home health services)”. Research has demonstrated

that isolation by the caregiver is one of the most significant contributing factors of mistreatment risk of older adults (NCEA, 1998; WHO/INPEA, 2002). The last addition made – “Unrealistic expectations of caregiver (unrealistic multiple requests)” – reflects the fact that many aspects of caregiving can be catalysts for older adult mistreatment, such as the intimacy required for personal care provision; the required commitment in time, effort, and resources; and the lack of respite (Anetzberger, 2000).

All additions recommended by the judges are supported by theory and research in this field. Furthermore, these sub-indicators place this instrument within the caregiver and community contexts and are supported by the conceptual framework of Kosberg and Nahmiash (1996).

We provided the expert judges with a definition of mistreatment as including both abuse and neglect and examined risk factors versus indicators. As a result, they chose a new title for the adaptation which reflects both the concepts of risk and of mistreatment. The adaptation is now entitled the “Mistreatment of Older Adult Risk Factors” (MOARF) screening tool.

During the second phase of this study, the assessment of inter rater reliability of the adapted instrument, the Kendall’s coefficient of concordance W provided a measure of inter-rater reliability. Based on these statistical results, 90 of the 95 sub-indicators achieved acceptable agreement at a .01 significance level. The other five sub-indicators did not achieve acceptable agreement. Discussion of these variations will now be provided with recommendations for problematic sub-indicators.

Despite the positive modifications to the sub-indicators resulting from the Delphi process, some sub-indicators remained subjective. Differences in ratings can be attributed to three primary sources of difficulty: (a) ambiguity related to lack of clarity, (b) inaccurate measurement due to negative wording, and (c) inconsistency related to the rating scale.

In several instances, different understandings of expressions were noted. For example, the acronym IADLs (which is used to refer to independent activities of daily living) and the term “means for living” were perceived differently by the nurses. Some correctly understood “IADLs” to include grocery shopping, bill payment, and meal preparation whereas others did not. Even though all participants were nurses, the researchers should not have assumed that all expressions would be understood by all (Polit & Beck, 2004). Other expressions remained very subjective contributing to poor measurement (Grant & Davis, 1997). For example, the sub-indicator “*Has poor coping mechanisms*”

rated inconsistently because only some participants viewed addictions such as gambling and alcohol abuse to be poor coping mechanisms. To remedy such ambiguity, rephrasing these items would be necessary as well as adding descriptors in parentheses to improve clarity. Uncertainty also resulted when more than one behavior was targeted by a sub-indicator. For example, "*Expresses guilt or anger, and bitterness towards the family*" described incongruent emotions. Whereas guilt was described with some of the older adults, they were not angry and bitter towards their family. Thus, addressing more than one concept in an item should be avoided (Burns & Grove, 2005).

The negative wording of sub-indicators such as, "*Does not attend activities outside the house*" combined with negative ratings of "*Not at all*" also contributed to inaccurate measurement (Burns & Grove, 2005; Polit & Beck, 2004). To resolve this difficulty, item wording could be modified to affirmative statements such as "*Attends few/no activities outside the home*".

Finally, in several situations, the nurses agreed on the presence of a sub-indicator but assessments varied in regards to its frequency on the Likert scale. For example, in two cases where the caregiver demonstrated verbal and physical outbursts towards the older adult, the nurses disagreed in regards to frequency. However, when reflecting upon this sub-indicator, the occurrence of such outbursts, regardless of frequency, was considered a risk factor. As the literature indicates that the response format should be matched to the measurement's purpose (Myers & Winters, 2002), the Likert-type scale truly is not helpful in this sense. Therefore, we recommend a dichotomised Yes/No scale which would serve to flag the presence of a risk factor.

The limitations for both phases of the study merit discussion. First, only five of the seven judges completed the study. Due to fatigue and to the effects of attrition that may result from a prolonged Delphi study (Hasson et al., 2000), the study would have benefited by starting with a larger panel. As well, a larger sample would have permitted the testing of significance of the second round of Delphi as the k (number of judges) of 5 was insufficient (Siegel & Castellan, 1988). In regards to the Delphi method itself, the time required for each round of the study varied from 8 to 10 weeks and therefore, data collection spanned over seven months. There is little guidance in the literature regarding an acceptable level of consensus or recommended number of required rounds, and consequently, the researcher must decide when to stop collecting data based on achieved consensus and the "law of diminishing returns" (Keeney et al., 2006, p. 207). This study was therefore ended after three rounds with a percentage of agreement of 60 to 100 per cent for the final sub-indicators. Continuation

of the study may have resulted in stronger consensus or refinement of the sub-indicators.

In regards to the inter-rater reliability trials, despite various recruitment strategies, only six nurses from one agency participated. It would have been beneficial to enlist the participation of more than one nursing agency to increase the number of participants. The present small convenience sample restricts generalizability of the study (Fu, McDaniel, & Rhodes, 2007). The time allotted for the participant sessions was insufficient since the majority of the nurses had not received any prior in servicing on mistreatment of older adults, and considerable training with a new instrument is recommended to ensure inter-rater reliability (Myers & Winters, 2002). As the work required to complete the 10 screening instruments was lengthy, possible haste by the participants may have led to careless assessments (Springer et al., 2002).

With respect to the paper scenarios, the content did not sufficiently describe the caregivers. Therefore, when insufficient information prevented actual assessment of a sub-indicator, the nurses rated "*Not at all*". Furthermore, some sub-indicators would have performed better with face to face contact such as "*Lack of eye contact, reluctant to answer*". Despite this limitation, the chosen design was convenient and assured that the instrument was validated in a controlled environment (Selwood et al., 2007). Lastly, the content of both risk factors and actual mistreatment in the paper scenarios may have contributed to lower inter-rater agreement. Scenarios strictly describing risk factors may have better served to assess inter-rater reliability of this risk screening instrument. For the paper scenario design to be effective, specific scenarios could be built based on seminal works in the literature on mistreatment risk of older adults (Endacott et al., 1999).

To further refine its psychometric properties, the MOARF requires further revision and revalidation before it can be used. Such work should initially consist of reconstructing the problematic sub-indicators based on the findings of this study and implementing a dichotomized Yes/No rating scale. Afterwards, a pilot study of the revised MOARF with a larger sample of home visiting nurses would be required. To correctly use a paper case design for the pilot, new scenarios should be developed focusing strictly on risk factors and containing sufficient information to properly assess each sub-indicator. Once satisfactory inter-rater reliability was established, actual testing of the MOARF with a large sample of older adults receiving home health services would be possible. This subsequent study would permit the length of the instrument to be revised as it may not be suitable for busy health care professionals in its present format of 95 sub-indicators.

Conclusion

The original e-IOA was developed to be used in an acute hospital setting by social workers to assess the risk of mistreatment with Hebrew-speaking older adults and their caregivers. The modifications recommended by the experts in this study have contributed to the tool's content validity in a new context, namely a community context in Ontario, Canada. Sub-indicators are clearer and more representative of the risk of older adult mistreatment in the community home setting. The instrument's comprehensiveness has been augmented because the tool's conceptual underpinnings are now more closely reflected in the additional sub-indicators being measured. Despite the study limitations, elevated Kendall's *W* values for most sub-indicators in the inter-rater trials by the home visiting nurses suggest that this is a positive step towards the development of a valid and reliable screening tool for risk of older adult mistreatment in the domestic setting.

Acknowledgements

The author wishes to acknowledge the assistance of her master's advisory committee in the preparation of this manuscript.

References

- Allan, M. (2002). Elder Abuse: A challenge for home care nurses. *Home Healthcare Nurse*, 20(5), 323–330.
- Almberg, B., & Grafström, M. (1997). Caring for a demented elderly person—burden and burnout among caregiving relatives. *Journal of Advanced Nursing*, 25, 109–116.
- Anetzberger, G.J. (2000). Caregiving: Primary cause of elder abuse? *Generations*, 24(2), 46–51.
- Anetzberger, G.J. (2001). Elder abuse identification and referral: The importance of screening tools and referral protocols. *Journal of Elder Abuse and Neglect*, 13(2), 3–22.
- Anetzberger, G.J. (2005). The reality of elder abuse. In G.J. Anetzberger (Ed.), *The clinical management of elder abuse* (pp. 1–25). Kentucky, USA: Routledge.
- Ansello, E.F. (1996). Causes and theories. In L.A. Baumhover & S.C. Beall (Eds.), *Abuse, neglect and exploitation of older persons: Strategies for assessment and intervention* (pp. 9–29). Baltimore, MD: Health Professions Press.
- Baker, J., Lovell, K., & Harris, N. (2006). How expert are the experts? An exploration of the concept of 'expert' within Delphi panel techniques. *Nurse Researcher*, 14(1), 59–70.
- Baker, M.W. (2007). Elder mistreatment: Risk, vulnerability, and early mortality. *American Psychiatric Nurses Association*, 12(6), 313–321.
- Baker, M.W., & Heitkemper, M.M. (2005). The roles of nurses on interprofessional teams to combat elder mistreatment. *Nursing Outlook*, 53, 253–259.
- Behling, O., & Law, K.S. (2000). *Translating questionnaires and other research instruments: Problems and solutions*. Thousand Oaks, CA: Sage Publications.
- Bello, A., & Singh, M.D. (2004). Descriptive data analysis. In G. LoBiodondo-Wood & J. Haber (Eds.), *Nursing research in Canada: Methods, critical appraisal, and utilization*. Toronto, Ontario, Canada: Elsevier Mosby.
- Berk, R.A. (1990). Importance of expert judgment in content-related validity evidence. *Western Journal of Nursing Research*, 12(5), 659–671.
- Bowles, N. (1999). The Delphi technique. *Nursing Standard*, 13(45), 32–36.
- Burns, N., & Grove, S.K. (2005). *The practice of nursing research: Conduct, critique and utilization* (5th ed.). St. Louis, MO: Elsevier Saunders.
- Choi, N.G., & Mayer, J. (2000). Elder abuse, neglect, and exploitation: Risk factors and prevention strategies. *Journal of Gerontological Social Work*, 33(2), 5–25.
- Cohen, M., Halevi-Levin, S., Gagin, R., & Friedman, G. (2006). Development of a screening tool for identifying elderly people at risk of abuse by their caregivers. *Journal of Aging and Health*, 18, 660–685.
- Cohen, M., Halevi-Levin, S., Gagin, R., & Friedman, G. (2007). Elder abuse: Disparities between older people's disclosure of abuse, evident signs of abuse, and high risk of abuse. *Journal of the American Geriatrics Society*, 55, 1224–1230.
- College of Nurses of Ontario. (2009). *Utilization of RNs and RPNs*. Toronto, Ontario, Canada.
- Cooney, C., Howard, R., & Lawlor, B. (2006). Abuse of vulnerable people with dementia by their carers: Can we identify those most at risk? *International Journal of Geriatric Psychiatry*, 21, 564–571.
- Cranswick, K., & Dosman, D. (2008). *Elder Care: What we know today*. Component of Statistics Canada Catalogue no. 11-008-X, Canadian Social Trends. Ottawa, Ontario, Canada: Statistics Canada.
- Davis, L., & Grant, J.S. (1993). Focus on psychometrics: Guidelines for using psychometric consultants in nursing studies. *Research in Nursing and Health*, 16, 151–155.
- DeVellis, R.F. (1991). *Scale development: Theory and applications*. Newbury Park, CA: Sage.
- Dresser, R. (2003). Ethical and policy issues in research on elder abuse and neglect. In R.J. Bonnie & R.B. Wallace (Eds.), *Elder mistreatment: Abuse, neglect, and exploitation in an aging America*. Panel to review risk and prevalence of elder abuse and neglect. Committee on national statistics and Committee on law and justice, Division of behavioral and social sciences and education. Washington, DC: The National Academies Press. Retrieved from <http://www.books.nap.edu/books/0309084342/html/index.html>.

- Endacott, R., Clifford, C.M., & Tripp, J.H. (1999). Can the needs of the critically ill child be identified using scenarios? Experiences of a modified Delphi study. *Journal of Advanced Nursing*, 30(3), 665–676.
- Evans, B.C. (2005). Content validation of instruments: Are the perspectives of Anglo reviewers different from those of Hispanic/Latino and American Indian reviewers? *Journal of Nursing Education*, 44(5), 216–224.
- Fortin, F. (2006). *Fondements et étapes du processus de recherche*. Québec, Québec, Canada: Chenelière éducation.
- Fu, M.R., McDaniel, R.W., & Rhodes, V.A. (2007). Measuring symptom occurrence and symptom distress: Development of the symptom experience index. *Journal of Advanced Nursing*, 59(6), 623–634.
- Fulmer, T., Guadagno, L., Dyer, C.B., & Connolly, M.T. (2004). Progress in elder abuse screening and assessment instruments. *Journal of the American Geriatrics Society*, 52, 297–304.
- Grant, J.S., & Davis, L.L. (1997). Focus on quantitative methods: Selection and use of content experts for instrument development. *Research in Nursing and Health*, 20, 269–274.
- Grant, J., Kinney, M., & Guzzetta, C. (1990). A methodology for validation nursing diagnoses. *Advances in Nursing Science*, 12(3), 65–74.
- Hardy, D.J., O'Brien, A.P., Gaskin, C.J., O'Brien, A.J., Morrison-Ngatai, E., Skews, G., et al. (2004). Practical application of the Delphi technique in a bicultural mental health nursing study in New Zealand. *Journal of Advanced Nursing*, 46(1), 95–109.
- Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, 32(4), 1008–1015.
- Hwalek, M.A., Goodrich, C.S., & Quinn, K. (1996). The role of risk factors in health care and adult protective services. In L.A. Baumhover & S.C. Beall (Eds.), *Abuse, neglect and exploitation of older persons: Strategies for assessment and intervention* (pp. 123–136). Baltimore, MD: Health Professions Press.
- Keeney, S., Hasson, F., & McKenna, H. (2006). Consulting the oracle: Ten lessons from using the Delphi technique in nursing research. *Journal of Advanced Nursing*, 53(2), 205–212.
- Kissal, A., & Beser, A. (2011). Elder abuse and neglect in a population offering care by a primary health care center in Izmir, Turkey. *Social Work in Health Care*, 50, 158–175.
- Kosberg, J.I., & Nahmiash, D. (1996). Characteristics of victims and perpetrators and milieus of abuse and neglect. In L.A. Baumhover & S.C. Beall (Eds.), *Abuse, neglect and exploitation of older persons: Strategies for assessment and intervention* (pp. 31–49). Baltimore, MD: Health Professions Press.
- Kozma, A., & Stones, M.J. (1995). Issues in the measurement of elder abuse. In M. MacLean (Ed.), *Abuse and neglect of older Canadians: Strategies for change* (pp. 117–128). Toronto, Ontario, Canada: Thompson Educational Publishing.
- Lachs, M.S., & Pillemer, K. (2004). Elder abuse. *Lancet*, 364, 1263–1272.
- Lefrançois, R. (1992). *Stratégies de recherche en sciences sociales: Application à la gérontologie*. Montréal, Québec, Canada: Les Presses de l'Université de Montréal.
- Levine, J.M. (2003). Elder neglect and abuse: A primer for primary care physicians. *Geriatrics*, 58(10), 37–44.
- Litwin, H., & Zoabi, S. (2004). A multivariate examination of explanations for the occurrence of elder abuse. *Social Work Research*, 28(3), 133–142.
- Lynn, M.R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382–385.
- McKenna, H.P. (1994). The Delphi technique: A worthwhile approach for nursing? *Journal of Advanced Nursing*, 19, 1221–1225.
- Mead, D., & Moseley, L. (2001). The use of the Delphi as a research approach. *Nurse Researcher*, 8(4), 4–23.
- Miller, L.S., Lewis, M.S., Williamson, G.M., Lance, C.E., Dooley, W.K., Schultz, R., et al. (2006). Caregiver cognitive status and potentially harmful caregiver behavior. *Aging and Mental Health*, 10(2), 125–133.
- Myers, K., & Winters, N.C. (2002). Ten-year review of rating scales. I: Overview of scale functioning, psychometric properties, and selection. *Journal of the American Academy of Child and Adolescent Psychiatry*, 41(2), 114–122.
- Nagpaul, K. (2001). Application of elder abuse screening tools and referral protocol: Techniques and clinical considerations. *Journal of Elder Abuse and Neglect*, 13(2), 59–78.
- National Advisory Council on Aging (NACA). (2006). *Seniors in Canada 2006 report card*. Ottawa, Ontario, Canada: Public Health Agency of Canada.
- National Center on Elder Abuse (NCEA). (1998). The national elder abuse incidence study. Final report. Retrieved from http://www.aoa.gov/AoARoot/AoA_Programs/Elder_Rights/Elder_Abuse/docs/ABuseReport_Full.pdf.
- National Research Council (NRC). (2003). Elder mistreatment. Abuse, neglect, and exploitation in an aging America. In R.J. Bonnie & R.B. Wallace (Eds.), *Panel to review risk and prevalence of elder abuse and neglect*. Committee on national statistics and Committee on law and justice, Division of behavioral and social sciences and education. Washington, DC: The National Academies Press. Retrieved from <http://www.books.nap.edu/books/0309084342/html/index.html>.
- Ontario Network for the Prevention of Elder Mistreatment (ONPEA). (2006). *Core curriculum and resource guide*. Retrieved from <http://www.onpea.org/>.

- Ortmann, C., Fechner, G., Bajanowski, T., & Brinkmann, B. (2001). Fatal neglect of the elderly. *International Journal of Legal Medicine*, 114, 191–193.
- Parks, R., Cintas, H.L., Chaffin, M.C., & Gerber, L. (2007). Brief assessment of motor function: Content validity and reliability of the fine motor scale. *Pediatric Physical Therapy*, 19, 315–325.
- Peguero, A.A., & Lauck, A.M. (2008). Older adults and their vulnerabilities to the exposure of violence. *Sociology Compass*, 2(1), 62–73.
- Petry, K., Maes, B., & Vlaskamp, C. (2007). Operationalizing quality of life for people with profound multiple disabilities: A Delphi study. *Journal of Intellectual Disability Research*, 51(5), 334–349.
- Podnieks, E., Pillemer, K., Nicholson, J.P., Shillington, T., & Frizzell, A.F. (1990). *National survey on abuse of the elderly in Canada*. Toronto, Ontario, Canada: Ryerson Polytechnical Institute.
- Polit, D.F., & Beck, C.T. (2004). *Nursing research: Principles and methods* (7th ed.). Philadelphia, PA: Lippincott Williams & Wilkins.
- Reis, M., & Nahmiash, D. (1998). Validation of the indicators of abuse (IOA) screen. *The Gerontologist*, 38(4), 471–480.
- Selwood, A., Cooper, C., & Livingston, G. (2007). What is elder abuse—who decides? *International Journal of Geriatric Psychiatry*, 22(10), 1009–1012.
- Shugarman, L.R., Fries, B.E., Wolf, R.S., & Morris, J.N. (2003). Identifying older people at risk of abuse during routine screening practices. *Journal of the American Geriatrics Society*, 51, 24–31.
- Siegel, S., & Castellan, N.J. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). Toronto, Ontario, Canada: McGraw-Hill Company.
- Springer, D.W., Abell, N., & Hudson, W.W. (2002). Creating and validating rapid assessment instruments for practice and research: Part 1. *Research on Social Work Practice*, 12, 408–439.
- Statistics Canada. (2009). *Range and quartiles*. Retrieved from <http://www.statcan.gc.ca/edu/power-pouvoir/ch12/5214890-eng.htm>.
- Waltz, C.F., Strickland, O., & Lenz, E. (1991). *Measurement in nursing research* (2nd ed.). Philadelphia, PA: F. A. Davis.
- Washington, C.C., & Moss, M. (1988). Pragmatic aspects of establishing inter rater reliability in research. *Nursing Research*, 37(3), 190–191.
- Wiles, J. (2003). Informal caregivers' experiences of formal support in a changing context. *Health and Social Care in the Community*, 11(3), 189–207.
- Wolf, R. (1997). Elder abuse and neglect: An update. *Reviews of Clinical Gerontology*, 7, 177–182.
- Wolf, R. (2000). Risk assessment instruments. *National Center on Elder Abuse Newsletter*, September 2000. Orange, CA: US Administration on Aging.
- World Health Organization/International Network for the Prevention of Elder Abuse (WHO/INPEA). (2002). *Missing voices: Views of older persons on elder abuse*. Geneva, Switzerland: World Health Organization.
- Zink, T., & Fisher, B.S. (2007). Family violence quality assessment tool for primary care offices. *Quality Management in Health Care*, 16(3), 265–279.