Title: Development and Validation of Providers' and Patients' Measurement Instruments to Evaluate Adverse Events after Spinal Manipulation Therapy

Author names:
KA Pohlman\textsuperscript{a}, M O’Beirne\textsuperscript{b}, H Thiel\textsuperscript{c}, JD Cassidy\textsuperscript{d}, S Mior\textsuperscript{e}, EL Hurwitz\textsuperscript{f}, M Westaway\textsuperscript{g}, S Ishaque\textsuperscript{h}, JY Yager\textsuperscript{h}, S Vohra\textsuperscript{i}

\textsuperscript{a}- University of Alberta, CARE Program, Department of Pediatrics, Faculty of Medicine and Dentistry; School of Public Health; and Women’s and Children’s Health Research Institute, 8B19 Edmonton General Hospital, 11111 Jasper Avenue, Edmonton, AB T5K 0L4, Canada

\textsuperscript{b}- University of Calgary, Department of Family Medicine, G012 HSC, 3330 Hospital Drive NW, Calgary, AB T2N4N1 Canada

\textsuperscript{c}- Anglo-European College of Chiropractic, 13-15 Parkwood Road, Bournemouth BH5 2DF, England

\textsuperscript{d}- Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto and Division of Health Care and Outcomes Research, University Health Network, Suite 200, Box 115, 700 Bay Street, Toronto, Ontario M5G1Z6

\textsuperscript{e}- Canadian Memorial Chiropractic College, 6100 Leslie Street, Toronto, Ontario M2H 3JI, Canada

\textsuperscript{f}- University of Hawaii, Office of Public Health Studies, 1960 East-West Road, Biomed D104H, Honolulu, Hawaii 96822, USA

\textsuperscript{g}- LifeMark Health, 30 Spingborough Blvd SW, Calgary AB T3H 0N9 Canada; McMaster University, Faculty of Physiotherapy, 1280 Main St. West, Calgary AB L8S 4L8 Canada

\textsuperscript{h}- University of Alberta, Department of Pediatrics, Faculty of Medicine and Dentistry, 11405 87\textsuperscript{th} Avenue NW, Edmonton, AB T6G 1C9 Canada

\textsuperscript{i}- University of Alberta, CARE Program, Department of Pediatrics, Faculty of Medicine and Dentistry; School of Public Health; and Women’s and Children’s Health Research Institute, 8B19 Edmonton General Hospital, 11111 Jasper Avenue, Edmonton, AB T5K 0L4, Canada

*Corresponding author.
Phone: 01 (780) 342-8592
E-mail: svohra@ualberta.ca
8B19-11111 Jasper Ave
Edmonton Continuing Care Centre
Edmonton, Alberta
Canada
T5K 0L4
Abstract

Aim of the study/Introduction. Although spinal manipulation therapy (SMT) is used throughout the world by chiropractors, osteopaths, physiotherapists and other manual therapists, yet there are no systematic data collection mechanisms in place to monitor and adjudicate SMT-related adverse events (AE) that occur after SMT. We established a reporting and learning system ("SafetyNet") to fill this void and to address several aims, one of which is a prospective population-based active surveillance study to (a) document AE after SMT, (b) identify potential risk factors, and (c) develop potential strategies to mitigate risk. The purpose of this paper is to describe the development and validation of provider and patient measurement instruments to allow for assessment and identification of potential SMT AE in provider offices.

Materials and Methods/Methodology. Instrument development and validation occurred in a step-wise fashion: 1) definition of terms (e.g. adverse event, seriousness, etc.); 2) identification and development of key domains, items, and sub-items; and 3) assessment of relevant measurement properties.

Results. Two provider short form instruments, a provider long form instrument, and a pre and post treatment patient comment form were developed, refined, and pilot tested with 12 providers and 300 patients.

Discussion/Conclusions. The development and validation of instruments to evaluate SMT AEs may benefit the SMT research community as well as clinicians and their patients by providing the opportunity for rigorous prospective assessment of potential SMT-related AEs and their risk factors, thus enhancing patient safety and the promotion of a safety culture. Placing the instruments in providers' offices for use on consecutive patients is next on the SafetyNet research agenda.

Keywords

Spinal Manipulation Therapy, Chiropractic, Physiotherapist, Validation, Instrument, Adverse Event

Highlights

- We developed and validated an instrument to evaluate SMT AEs.
- Operational definitions for all relevant terms were first established.
- Identification of key domains, items, and sub-items was the second step.
- Relevant psychometric properties were assessed.
- Benefits of this instrument include the collection of rigorous prospective SMT AE assessment.
**Introduction**

The patient safety movement began in earnest with the 1991 report, *To Err Is Human: Building a Safety Health System* which found that U.S. hospital medical errors killed between 44,000 and 98,000 patients each year [1]. This report called for a shift in health care culture, moving away from a “blame and shame” culture towards a systems-based approach, promoting the identification and mitigation of adverse events. However, cultural shift is multifactorial and highly complex [1]. Barriers include litigation, professional protection, peer criticism, and potential respective governing body disciplinary actions. Understanding the multidimensionality and dynamic nature of culture particularly in community-based primary care is required if transformation to a safety culture is to occur [2]. Spinal manipulation therapy (SMT) is a regulated treatment, practised in community-based settings by several health care professions in community-based settings, including such as chiropractors, osteopaths, naturopaths, physiotherapists, and physicians. Potential safety concerns, including the potential for an adverse event/event (AE) related to the delivery of SMT have exists within all of these professions. Although the need to improve the identification of SMT AEs has been identified documented [3, 4]. Despite this, no formal safety reporting and learning mechanisms exist in North America to monitor, assess and reduce SMT-related AEs.

Reporting and learning systems have emerged as a key strategy to identify and mitigate risks associated with health care delivery [5, 6]. They are typically anonymous and confidential methods of monitoring the occurrence of clinical or administrative incidents, and used to develop improvement strategies to address the cause of the incidents. Good reporting and learning systems move beyond pure reporting element and lead into an environment of continuous learning (Kirk 2007). Most often these systems are found in association with hospital-based quality assurance and patient safety initiatives; community-based reporting and learning systems remain quite scarce. This gap is relevant, as the majority of health care delivery occurs in the community, not in hospitals [7]. As the first step in developing a reporting and learning system, AE identification, reporting, and assessment are vital to patient safety, as it promotes the identification of modifiable risk factors thus reducing harms. Reporting and learning systems are particularly useful in promoting culture change for participating health care providers, yet they have rarely been developed for complementary/alternative medicine can reduce harms system.

Adverse events AEs associated with SMT have been studied in different research designs, including clinical trials [8-10]. AEs that have been reported by providers after adult SMT. Clinical trials are not the optimal design to collect rare AEs [10] and most observational studies lack standardized instruments and operational definitions for relevant terms [11]. Reported AEs following SMT in adult patients are most often self-limiting and usually consist of symptoms such as radiating musculoskeletal pain, nausea, dizziness, or tiredness [11-13]. Other, there have been other more serious, but rare potential AEs have been reported, such as cauda equina syndrome [13, 14] and stroke. However, there is currently no concrete evidence that SMT is a recent case control study suggests the cause. In a large randomised controlled trial that actively sought patients' feedback after cervical SMT, 30% of the patients reported an AE [7]. This trial demonstrated that data collection from providers alone may not provide a complete picture association between

**Abbreviations**

Adverse events (AE), spinal manipulation therapy (SMT)
manipulation and patient data are required for a complete picture regarding SMT AE. Stroke is confounded by indication", raising doubt about a causal relationship [15].

To help overcome the absence of high quality data about SMT AE in North America, we developed SafetyNet, a reporting and learning system. It is comprised of a number of research projects that aim to support the development of a patient safety culture for regulated SMT providers. SafetyNet reflects the efforts of a large multidisciplinary research team and with expertise in physiotherapy, chiropractic, and various medical specialties. SafetyNet has several coordinated aims, one of which is to lead objectives, including conducting a prospective population-based active surveillance study to document moderate and serious AEs after SMT, to identify potential risk factors, and to develop potential strategies to mitigate risk. The team is based in Alberta, Canada, with steering committee members from across Canada, as well as from the United States and Europe. Thus far, given that As chiropractors and physiotherapists provide the majority of SMT care in Alberta, our team has focused on developing instruments for use in their practices. We describe one of the first projects undertaken by members of this team to develop and validate provider and patient measurement instruments to allow for assessment of potential SMT AE in provider offices.

Abbreviations
Adverse events (AE), spinal manipulation therapy (SMT)
Research Approach

The research approach we took was to develop standardised instruments with clear definitions of relevant terms. This development and validation of these instruments occurred in a step-wise fashion: 1) definition of terms (e.g. adverse event, seriousness, etc.); 2) identification and development of key domains, items, and sub-items; and 3) assessment of relevant measurement properties. The instruments needed to be brief enough to facilitate their implementation, yet detailed enough to be informative. A multi-disciplinary team of content and/or SMT experts and providers (n= 16) were involved, as their experience was needed in each step. The completion of a step was not considered to have been achieved until consensus was reached. This took a period of about 18 months.

Methods and Findings

Step 1: Definition of Terms

Unclear definitions are one of the major methodological flaws when reporting on manual therapy adverse event data [4, 11]. Our team’s first step was to define AE and determine other variables that needed to have operational definitions to allow for meaningful study. As shown in Table 1, we identified existing definitions of AE from relevant organisations. The team adapted the definition of AE from the International Conference of Harmonisation (ICH) [16, 17]: Any unfavourable sign, symptom, or disease temporally associated with the treatment, whether or not caused by the treatment.

Our team decided the following variables were necessary for meaningful AE assessment: (i) seriousness; (ii) causality (i.e. relatedness); (iii) preventability; and (iv) patient disposition. Similar to the AE process, definitions for these variables were sought from relevant organisations and the published literature. Table 2 provides all the definitions that were considered for seriousness. For our study’s purposes, we adapted the definition proposed by the National Cancer Institute [24]:

Mild: asymptomatic or mild symptoms, self-care only (e.g. ice/heat, over-the-counter analgesic);

Moderate: limiting age-appropriate activities of daily living (e.g. work, school) OR sought care from a medical doctor;

Severe: medically significant but not immediately life-threatening; temporarily limits self-care (e.g. bathing, dressing, eating); OR urgent or emergency room assessment sought; and

Serious: results in death OR a life-threatening adverse event OR an AE resulting in inpatient hospitalisation or prolongation of existing hospitalisation for more than 24 hours; a persistent or significant incapacity or substantial disruption of the ability to conduct normal life functions; a congenital anomaly/birth defect.
For causality, we modified the definition proposed by the WHO, de-emphasizing health products and making the language more inclusive of practice-based health care interventions [27] (see Table 3):

- **Certain**: a clinical event occurring in a plausible time relationship to treatment, and which cannot be explained by concurrent disease or other drugs or therapies;

- **Probable/Likely**: a clinical event with a reasonable time sequence to treatment, unlikely to be attributed to concurrent disease or other drugs or therapies;

- **Possible**: a clinical event with a reasonable time sequence to treatment, but which could also be explained by concurrent disease or other drugs or therapies; and

- **Unlikely**: a clinical event with a temporal relationship to treatment which makes a causal relationship improbable, and in which drugs, other therapies or underlying disease provide plausible explanations.

For patient disposition, we adopted the definition proposed by the National Institute of Arthritis and Musculoskeletal and Skin Diseases [30]:

1: Resolved, No Sequelae
2: AE still present- no treatment
3: AE still present-being treated
4: Residual effects present-no treatment
5: Residual effects present- treated
6: Death
7: Unknown

We also adopted a definition of preventability from Baker et al. [34]:

1: Virtually no evidence of preventability
2: Slight to modest evidence of preventability
3: Preventability not quite likely (less than 50/50, but “close call”)
4: Preventability more than likely (more than 50/50, but “close call”)
5: Strong evidence of preventability
6: Virtually certain evidence of preventability

Step 2: Identification and Development of Key Domains, Items, and Sub-Items

To be able to assess the relationship between exposure and outcome, separate patient and provider instruments were developed with the following domains: (i) details of the intervention, including anatomic location and dose; (ii) details of any AE reported, including time to occurrence, seriousness, patient disposition; and (iii) potential confounders, including patient’s underlying health concerns and other therapies used.
For feasibility reasons, the measurement instruments also needed to: a) be easy to complete by the users; b) collect essential information without being too burdensome; c) avoid promoting hypervigilence or stress about potential AE; and d) collect information for a reasonable duration, balancing. Finally, we balanced our desire to collect all potential related AE while recognizing the diminishing return from AEs that occurred more than a week after treatment.

We used an iterative process for developing and refining items and sub-items until consensus was reached on both the questions and response options. Four forms were developed (see Appendix A-C):

a) **Two Provider Short Forms**: Since terminology differs amongst SMT professions, the treatment section was designed to be profession-specific; thus both a physiotherapy and chiropractic versions were developed. We designed these forms to be completed on all consecutive patients seen during the study period, hence the majority of information is collected through check-boxes. This design allows the forms to only take a few seconds to complete. (Appendix A)

b) **Provider Long Form**: This instrument is designed to be completed for all moderate, serious, or severe patient reported AEs. (Appendix B) These forms contain text boxes to allow for narrative descriptions, allowing for better understanding of the events leading to the AE [16].

c) **Two Patient Comment Form**: The first version of this instrument was a two-sided form that collects information about the SMT visit from the patient’s perspective. Patient feedback was evaluated by our study team, and the instrument was modified into two separate pre- and post-treatment instruments. The pre-treatment instrument addresses items such as medical history and current symptoms. At the recommendation of SMT provider groups, the form starts by gathering information about the overall patient satisfaction and then, treatment sought and overall experience, positive or negative. Only patients, who report a negative experience, are asked additional questions regarding a potential AE and its nature, severity, and duration as well as follow-up care required and current disposition. Both a paper and web-based versions were created for the post-treatment instrument; they are identical except for 6 extra questions on the web-based version, allowing for more space for patient responses. (Appendix C)

Step 3: **Assessment of Relevant Measurement Properties**

Good measurement properties legitimize a health status questionnaire / instrument [17, 27, 35]. The quality criteria for a health instrument’s measurement properties are outlined in Figure 1. Only two measurement properties were completely relevant for the validation of these instruments: content validity and hypotheses testing. A portion of reliability was evaluated.
The other measurement properties are not relevant or too early in development to assess. Internal consistency and structural validity are not relevant as no total score from these instruments is sought. These instruments have only been developed and validated in English in two Canadian provinces; it is therefore premature to consider cross-cultural validity. Since there is no gold standard for assessing SMT AE, criterion validity cannot be evaluated. Responsiveness and measurement error are not relevant because this study is not looking for change over time and measurement error will be assessed in future studies.

Content validity assesses the instrument to ensure that the concepts of interest are embodied [35, 36]. For this instrument, the development included the following aspects:

*Measurement aim of the questionnaire:* The aim or specific definitions were clearly defined at the start of the study, which was followed up to ensure that each question would allow the terms to be adequately assessed.

*Target population:* Both SMT providers and their patients reviewed and provided feedback during the pre-testing period of the instrument development.

*Concepts:* The overall concept was to measure AEs associated with SMT and this was revisited by the multi-disciplinary team throughout the development of the instruments.

*Item selection and item reduction:* Questions were identified through literature reviews, expert consensus, pilot testing with field practitioners, and discussion with regulatory bodies. Each revision included a thorough review of all forms to ensure all relevant items were included, while removing redundancies.

*Interpretability of the items:* Pre-testing was used to examine the readability and question comprehension by both the providers and the patients. We also developed 2 provider short forms so that profession-specific terminology could be accommodated (provider feedback suggested this was important to prevent misinterpretation).

Hypotheses testing (part of construct validity) assesses the instrument's ability to measure the specific question that it was designed to do so [35]. For this instrument, our questions, [i.e. hypotheses,] and definitions were determined first (*Step 1*), followed by the development of the instruments to address our study questions (*Step 2*). Throughout the development of these instruments there was a consistent ongoing and iterative feedback to ensure that the questions asked were aimed at answering our specific study aim.

Reliability is the extent for which respondents who have not changed are the same when repeated measures are taken under several conditions [27, 35]. There are 3 main components: test-retest, inter-rater, and intra-rater. Of these components the first two are not relevant, in that we expect a change over time and different respondents (both providers and patients) should be expected to have different perceptions. Intra-rater reliability was evaluated on a limited basis during patient and
provider pretesting, where the instruments were found to collect the same information that was described during the interviews.

**Pretesting**

The penultimate version of the provider instruments was pretested by providers (n=12) and patients (approximately n=300) in Alberta and British Columbia, Canada. The Health Research Ethics Board at the University of Alberta approved the pretesting of the instruments.

All providers found that the short form instrument was quick and easy to use and could be implemented within existing practice procedures. General feedback on the long form instrument indicated that the questions were relevant when reporting a moderate, serious, or severe AE.

The penultimate version of the patient instrument was discussed with a small convenience sample of patients (n=15) following their visit with a SMT provider. One-on-one interviews were conducted until data saturation was achieved. The interviews were not recorded. A few patients found the form too long and some would not be willing to take the extra time to complete it. A common statement heard was ‘I would complete the form if my provider asked me to. If it was important to him / her, then I would make it important for me to do.’ Minor clarifications were requested. All patients stated that the list of potential AEs did not concern them or make them feel any less comfortable with the care that they had just received. Non-English speaking patients were unable to complete the patient comment form. The team therefore decided that for Non-English speaking patients, only the provider form instruments were to be completed.

**Discussion**

The patient safety movement began in earnest with the 1991 report, *To Err Is Human: Building a Safety Health System* which found that U.S. hospital medical errors killed between 44,000 and 98,000 patients each year [22]. This report called for a shift in health care culture, moving away from a “blame and shame” culture towards a systems-based approach, promoting the identification and mitigation of adverse events. Cultural shift is multifactorial and highly complex. Barriers include litigation, professional protection, peer criticism, and potential respective governing body disciplinary actions.

Limitations of current established This project started with definition of terms to be used consistently throughout measurement and assessment and then developed and validated the measurement instruments to assess AEs after SMT. A limitation of current AE reporting systems include the lack of ownership by professionals [37]. For example, in Australia the system was developed for acute care settings and therefore only used by those providers, as opposed to primary care providers in community settings. To try and engage the SMT community, a multi-disciplinary team of experts in epidemiology, SMT and patient safety research, providers and professional associations/regulators collaborated on the development of our study definitions and instruments. Instrument refinement occurred in an iterative process involving extensive
conversation and debate; the process was complete when consensus was reached. Our goal was for each participating profession to feel that the instruments “belonged” to them.

The importance of patients’ perspectives and experience to the patient safety movement was recognized as one of the six aims to the 2001 Institute of Medicine report, *Crossing the Quality Chasm* [38]. While most passive reporting systems are designed for provider reporting only, we have designed a system that provided both patients and clinicians the opportunity to report potential SMT AE. Patient perspective is especially important as health care providers are notoriously have demonstrated poor at reporting of suspected AEs [39]. Additionally, patient reports should come directly to a third party, since patients may be reluctant to report AEs to their providers in fear of being labeled ‘difficult’ [40]. On the basis of patient feedback, we had divided the patient instrument into 2 parts, which allow will reduce recall bias. Another important virtue is the use of standardised terminology and definitions on both the provider and patient instruments [11, 41, 42]. Similar to Carlesso et al.’s approach, this study used their team of experts and patients to develop the study’s definitions for AE and other related terms.

Surveillance for AE may be passive or active. Passive surveillance systems have been developed for SMT providers, such as the CPIRLS system currently open to all European chiropractors to anonymously report incidents [43, 44]. Like other passive surveillance systems (e.g. pharmacovigilance), it is challenged by considerable under-reporting [20, 45, 46]. Active surveillance systems have shown themselves to improve both the quality and quantity of AE reports, such that they can be evaluated in a meaningful fashion [47].

Both active and passive surveillance systems rest on a foundation of the identification of incidents, or “cases”. Considerable debate has occurred regarding whether or not case reports can be used to infer causation [48, 49], including the role of case reports in patient safety. While case reports are the base of the evidence hierarchy when evaluating effectiveness [50], some have proposed an inverted pyramid when evaluating harms, in light of the tremendous amount of information provided by well-reported cases [51]. The majority of harms identified in healthcare first emerged as case reports, which have served to generate hypotheses subsequently evaluated through other study designs [52]. Confounding by indication, or protopathic bias, is a major concern whenever AEs may be due teassociated with the patient’s underlying health condition, rather than due to the intervention. For example, one large case-crossover study recently suggested that vertebrobasilar stroke following SMT was the result ofreflected patients with cervical dissection-related head and neck pain seeking care from chiropractors, and that the SMT was coincidental and not in the causal pathway of the subsequent strokes [10].

In our study, we prospectively collect SMT exposure data on all patients, whether or not AE occur. We also request outcome data whether or not an AE occurs, allowing us to compare cases (those who experience AE) to controls (those who do not experience AE). Finally, we have developed an in-depth process to assess moderate, serious, and severe AEs by a multi-disciplinary team using validated approaches for harms assessment. While the instruments described in this paper do not evaluate administrative or other non-clinical incidents, these are included in other parts of the SafetyNet research program.
Our approach combines expert judgment and standardized tools, the gold standards in patient safety [53]. Our research will contribute to knowledge on patient safety and SMT. It will help to gauge the frequency and seriousness of the most common AEs. Most importantly, it will stimulate a dialogue on patient safety amongst practitioners of SMT. This in turn will help to develop more advanced study methodologies to assess causal relationships and preventive measures to ensure patient safety. Our goal is to collect high quality data that will make a meaningful contribution to our current understanding of SMT AE.

Conclusions

The development and validation of instruments to evaluate SMT AEs may benefit SMT research by providing the opportunity for rigorous prospective assessment of potential SMT-related AEs and their risk factors. We have developed profession-specific forms and engaged members of each profession who can act as champions, promoting patient safety culture for community-based SMT providers. Future efforts with these instruments include putting them into providers’ offices for use on consecutive patients in an effort to assess AE after SMT.
Acknowledgements

Katherine Pohlman receives an educational fellowship from NCMIC. Sunita Vohra receives salary support as an AIHS Health Scholar.

The work was done as part of a team grant; the authors would like to thank their team members, Greg Kawchuk and Michael Hill, as well as the EPICORE Centre for their insightful comments and suggestion to improve the work.

Funding Support

This work was supported by funding from the Canadian Institutes of Health Research, Alberta Innovates – Health Solutions, and the Women and Children's Health Research Institute, University of Alberta.
Figure 1. Quality criteria for a legitimized health instrument’s measurement properties.
References


19. Good clinical practice; ICH-GCP (E6), glossary art. 1.2. [Internet]. Available from: http://ichgcp.net/1-glossary.


44. The chiropractic patient incident reporting and learning system [Internet]. Available from: www.cpirls.org.


