Journal of Safety Research xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Developing leading indicators from OHS management audit data: Determining the measurement properties of audit data from the field

Lynda S. Robson, * Selahadin Ibrahim, Sheilah Hogg-Johnson, Ivan Steenstra, [1] Dwayne Van Eerd, Benjamin C. Amick III [2]

Institute for Work & Health, 481 University Ave., Suite 800, Toronto, Ontario M5G 2E9, Canada

ARTICLE INFO

Article history:
Received 11 January 2016
Received in revised form 5 October 2016
Accepted 20 February 2017
Available online xxxx

Keywords:
Performance measurement
Audit
Leading indicator
Management system

ABSTRACT

Introduction: OHS management audits are one means of obtaining data that may serve as leading indicators. The measurement properties of such data are therefore important. This study used data from Workwell audit program in Ontario, a Canadian province. The audit instrument consisted of 122 items related to 17 OHS management elements. The study sought answers regarding (a) the ability of audit-based scores to predict workers' compensation claims outcomes, (b) structural characteristics of the data in relation to the organization of the audit instrument, and (c) internal consistency of items within audit elements. *Method:* The sample consisted of audit and claims data from 1240 unique firms that had completed one or two OHS management audits during 2007–2010. Predictors derived from the audit results were used in multivariable negative binomial regression modeling of workers' compensation claims outcomes. Confirmatory factor analyses were used to examine the instrument's structural characteristics. Kuder–Richardson coefficients of internal consistency were calculated for each audit element. *Results:* The ability of audit scores to predict subsequent claims data could not be established. Factor analysis supported the audit instrument's elementbased structure. KR-20 values were high (≥0.83). *Conclusions:* The Workwell audit data display structural validity and high internal consistency, but not, to date, construct validity, since the audit scores are generally not predictive of subsequent firm claim experience. Audit scores should not be treated as leading indicators of workplace OHS performance without supporting empirical data. *Practical applications:* Analyses of the measurement properties of audit data can inform decisionmakers about the operation of an audit program, possible future directions in audit instrument development, and the appropriate use of audit data. In particular, decision-makers should be cautious in their use of audit scores as leading indicators, in the absence of supporting empirical data.

1. Introduction

1.1. OHS performance measurement and leading indicators

Measuring an organization's occupational health and safety (OHS) performance is a key element of managing OHS well (ANSI/AIHA/ASSE, 2012; OHSAS Project Group, 2007; Redinger & Levine, 1998). Comprehensive performance measurement includes determining whether an organization is meeting its OHS objectives and monitoring whether the various OHS programs and risk controls are operating as intended. Various means may be used to measure performance (e.g., administrative data, employee survey, analytical instruments).

OHS professionals typically distinguish between two types of performance measures: leading and trailing indicators. Trailing indicators (also known as lagging, reactive or negative indicators) measure OHS outcomes of interest (e.g., injury and illness rates). Leading indicators (also known as proactive or positive indicators) measure workplace activities, conditions, and events that are relevant to or may determine OHS outcomes. A few examples of leading indicators are safety climate measures, frequency of workplace inspections, and noise exposure levels. Other examples can be found in Glendon and Booth (1995), a special issue in Safety Science (Hopkins & Hale, 2009) and in National Occupational Health & Safety Commission (1999), as well as in guidance for implementing OHS management systems (ANSI/AIHA/ASSE, 2012; OHSAS Project Group, 2008).

Some researchers and guidance documents offer conceptual frameworks with which to consider OHS performance measurement. For

* Corresponding author.
E-mail addresses: lrobson@iwh.on.ca (L.S. Robson), sibrahim@iwh.on.ca (S. Ibrahim), shoggjohnson@iwh.on.ca (S. Hogg-Johnson), ivan.steenstra@ryerson.ca (I. Steenstra), dvaneerd@iwh.on.ca (D. Van Eerd), bamick@iwh.on.ca, bamickii@fiu.edu (B.C. Amick).
[1] Present address: Ted Rogers School of Management, 350 Victoria Street, Toronto, Ontario M5B 2K3, Canada.
[2] Present address: Department of Health Policy and Management, Robert Stempel College of Public Health & Social Work, Florida International University, 11200 SW 8 St, AHC 5-453, Miami, FL 33199, United States.

http://dx.doi.org/10.1016/j.jsr.2017.02.008
0022-4375/

example, van Steen (1996) recommends measuring four domains: failures, plant and equipment, systems and procedures, and people. The Health & Safety Executive (2001) recommends measuring hazard burden, three levels of the OHS management system, OHS culture, and OHS outcomes. Like the HSE, Kjellén (2000) uses an input–process–output model to identify several performance domains, though these differ from those identified by the HSE. Hinze, Thurman, and Wehle (2013) suggest considering "passive" and active" leading indicators, with the latter being more subject to change in the short term and therefore having more potential for leading to intervening proactive action. Wurzelbacher and Jin (2011) considered OHS performance measurement more broadly than most by explicitly including leading and trailing indicators of not only primary prevention, but also secondary and tertiary prevention of illness, injury and disability. Finally, even more broadly, Warrack (1998) considered OHS performance measurement not only at the workplace-level, but also at the government jurisdiction-level.

1.2. OHS management audit data as a measure of organizational performance

Some have suggested that the data derived from OHS management audits could serve as a measure of OHS performance (Glendon & Booth, 1995; NOHSC, 1999; Redinger & Levine, 1998). An OHS management audit is used to evaluate the state of a workplace's management structure and processes related to OHS. It determines whether an organization is conforming with a particular standard, such as its own policies and procedures, applicable legislation and regulations, or another standard external to the organization. Common external OHS management standards are those developed by national or international standard organizations (e.g., ANSI/AIHA/ASSE Z10-2012). Besides conformance to a standard, a management audit may also seek to answer whether the workplace's system for managing OHS is effective (relative to internal or external OHS objectives).

The processes carried out by auditors include: (a) gathering audit evidence through systematic data collection; typically by reviewing documentation, conducting interviews and observing worksites; (b) evaluating the audit evidence against audit criteria; and (c) summarizing and reporting the results (ISO, 2011; Robson, Macdonald, Gray, Van Eerd, & Bigelow, 2012). Often step 2 yields categorical data, with categories such as yes/no or levels of conformance. Then in step 3 the categorical data are summarized as a count or a percentage score. In addition to the final quantitative summary, there is usually a qualitative summary that identifies areas for improvement in OHS management.

Audit data are often used to make evaluative judgments about organizational performance in OHS: for example, determining the relative performance of organizational sub-units (as in Gunningham & Sinclair, 2009); determining whether there has been a change in performance over time (as in Bunn, Pikelny, Slavin, & Paralkar, 2001; LaMontagne et al., 2004; Nielsen, Rasmussen, Glasscock, & Spangenberg, 2008; Pearse, 2002); and assessing whether a target level of performance is being met, as in some recognition or reward schemes (Blewett & O'Keeffe, 2011; Eisner & Leger, 1988; Robson et al., 2012). In these situations, the measurement properties of audit data become germane. In particular, one would want to know that audit results are *reliable* (i.e., they have minimal measurement error) and *valid* (i.e., they truly measure OHS performance); and if one were comparing audit results over time, one would also want to know that audit data were *responsive* to change (i.e., have the ability to detect change in performance when real change has taken place) (Mokkink et al., 2010).[3] More specific aspects of reliability and validity can also be considered, such as internal consistency, interrater reliability, test–retest reliability, content validity, construct validity, and criterion validity.

Despite the need for knowledge about the measurement properties of management audit data used in evaluative decision-making, a thorough search of the research literature found little in this area (Robson & Bigelow, 2010). Moreover, much of the extant research had been conducted with data collected for research purposes, rather than on data from an active audit program. The review identified a few examples of well-executed published research related to content validity (Dyjack et al., 2008; Redinger & Levine, 1998) and inter-rater reliability (Dyjack et al., 2006; Kuusisto, 2000). Only one study was found that examined the relationship between audit results and injury outcomes (Eisner & Leger, 1988). Its null findings were not surprising, given its small sample size and the crude audit-based measure (1- to 5-stars, based on the award achieved). Since the time of the review, a study of the Singapore construction sector established that OHS management audit data can be used to predict concurrent injuries (Goh & Chua, 2013). As well, Robson, Macdonald, Van Eerd, Gray, and Bigelow (2010) conducted a content validity analysis of five audit instruments used in the province of Ontario, Canada.

1.3. Research questions

Our aim in the present study was to address the research gap in the area of the measurement properties of audit data, by studying data from an operating auditing program, in the context of the audit data being potential leading indicators of OHS performance. After reviewing the OHS management audit programs operating in the broader public sector in the province of Ontario, the Workwell program, operated by the Workplace Safety and Insurance Board of Ontario (WSIB), was identified as the most promising because of its relatively high volume of activity and sophisticated data capture methods. Given the lack of knowledge in the area of measurement properties of audit data, the research questions were exploratory in nature. Arguably, a critically important measurement test of an OHS leading indicator is a determination of the direction and strength of the relationship with a corresponding OHS outcome. This would be a test of the indicator's *construct validity*; i.e., the degree to which indicator scores are consistent with theoretically based hypotheses about relationships between those scores and other variables. This is the basis of this study's *primary* research question:

> RQ1: How predictive are the metrics derived from the audit data of subsequent firm injury claim experience?

Our underlying premise was that, if audit data could serve as good leading indicators, there would be a statistically significant and negative relationship between audit scores and lost-time injury rates, because higher audit scores should reflect better organizational practices in primary prevention and return-to-work following injury, which should then lead to lower lost-time injury rates. We also explored the relationship of audit data with no-lost-time injury rates in the study, but there has been no prior expectation about this relationship. There was reason to believe it could be negative, due to better primary prevention; but also reason to believe that it could be positive, due to more effective return-to-work programs (a section of the WW audit) or better reporting of injuries.

The remaining research questions were *secondary* to RQ1. One question was concerned with *structural validity*, which is the degree to which scores are an adequate reflection of the dimensionality of the construct to be measured:

> RQ2: Are the structural characteristics of the audit data consistent with the organization of the audit tool?

[3] Though the taxonomy of measurement properties presented by Mokkink et al. (2010) was developed by clinical health measurement experts to measure individuals rather than organizations, we use it here and in the framing of the Research Questions below, because those researchers' applications, like audits, are similarly evaluative in nature and typically involve the use of questionnaire-based assessment tools.

Please cite this article as: Robson, L.S., et al., Developing leading indicators from OHS management audit data: Determining the measurement properties of audit data from the field, *Journal of Safety Research* (2017), http://dx.doi.org/10.1016/j.jsr.2017.02.008