Methodological issues in evaluating workplace interventions to reduce work-related musculoskeletal disorders through mechanical exposure reduction

by Donald C Cole, MD,1, 2 Richard P Wells, PhD,3 Mardy B Frazer, PhD,3 Mickey S Kerr, PhD,1, 4 W Patrick Neumann, LSc,5, 6 Andrew C Laing, MSc,3 the Ergonomic Intervention Evaluation Research Group7


Researchers of work-related musculoskeletal disorders are increasingly asked about the evidentiary base for mechanical exposure reductions. Mixed messages can arise from the different disciplinary cultures of evidence, and these mixed messages make different sets of findings incommensurate. Interventions also operate at different levels within workplaces and result in different intensities of mechanical exposure reduction. Heterogeneity in reporting intervention processes and in measuring relevant outcomes makes the synthesis of research reports difficult. As a means of synthesizing the current understanding of measures, this paper describes a set of intervention and observation nodes for which relevant workplace indicators prior to, during, and after mechanical exposure reduction can provide useful information. On the basis of this path of impacts from exposure reduction, an approach to the evaluation of multilevel ergonomic interventions is described that can assist fellow researchers in producing evidence relevant to the challenges faced by workplace parties and policy makers.

Key terms arm pain, back pain, ergonomics, human factors, occupational injury, repetitive strain injury, cumulative trauma disorders.

A broad range of physical, psychological, and work organizational factors have been epidemiologically established as plausible risk factors for the development of musculoskeletal disorders (1–3). More specifically, the risk of reporting pain increases when there are high peak compressions and shear forces and high cumulative tissue loading due to rapid rates or prolonged durations of worktasks and awkward, repetitive, or prolonged postures associated with worktasks (2, 4–7, 8–10). High force, frequent repetition, and awkward posture are components of the construct of high mechanical exposure, conceptualized as the magnitude, time variation pattern, and duration of forces on body tissues (1, 11).

A general aim of workplace intervention should be to reduce known mechanical exposures, often through “ergonomic” intervention; yet workplace parties have faced conflicting messages about the likely impact of such intervention from researchers (12). Existing review

1 Institute for Work & Health, Toronto, Ontario, Canada.
2 Department of Public Health Sciences, University of Toronto, Toronto, Ontario, Canada.
3 Department of Kinesiology, University of Waterloo, Waterloo, Ontario, Canada.
4 School of Nursing, Faculty of Health Sciences, University of Western Ontario, London, Ontario, Canada.
5 National Institute for Working Life - West, Göteborg, Sweden.
6 Department of Design Sciences, Lund University, Lund, Sweden.
7 The Ergonomic Intervention Evaluation Research Group [Members at the time of writing, in addition to the authors, included Theberge N, Norman RW, Reid M, Brawley LR, (Department of Kinesiology, University of Waterloo, Waterloo, Canada); Theberge N, Granzow K, Dixon S (Department of Sociology, University of Waterloo, Waterloo, Canada); Tompa E, Ferrier SE, Swift M, Ibrahim S (Institute for Work & Health, Toronto, Canada); Tompa E (Department of Economics, McMaster University, Hamilton, Canada); Kerton R (Department of Economics, University of Waterloo, Waterloo, Canada).]

Reprint requests to: Dr Donald C Cole, Institute for Work & Health, 481 University Avenue, Suite 800, Toronto, ON, Canada, M5G 2E9. [E-mail: dcole@iwh.on.ca]
papers on ergonomic intervention usually note that laboratory research has shown a reduction of forces on tissues or improved surface electromyographic responses as indicators of reductions in mechanical exposure; yet field research has been less conclusive (13, 14). The conclusiveness of workplace-based evaluations of ergonomic intervention has been reduced by the lack of attention to research design (15); inadequate accounting for the timing or impact of interventions (18). In their review of ergonomic interventions, Westgaard & Winkel (11) commented on each of these issues at some length.

Underlying such mixed messages on intervention effectiveness are different disciplinary cultures of evidence, intervention corresponding to different organizational levels from jobs to workplace policies, varying intensities of intervention, varying considerably by resources allocated by workplaces, and marked heterogeneity in documentation or indicators of both intervention processes and outcomes of interest by researchers and their workplace partners. In this narrative review and conceptual paper, we delineate cultures of evidence, describe issues associated with levels, set out a path linking nodes with associated indicators, and propose an approach to evaluating multilevel ergonomic intervention.

Cultures of evidence
Different disciplines and different social groups have different conceptions of evidence that they deem persuasive in guiding policy and practice (Upshur et al, unpublished). Douglas & Wildavsky (19) have described how different sets of shared values and supporting social institutions, or cultures, highlight certain risks and downplay other risks in a process of ongoing reconstruction of evidence and values. In this section, we highlight some of the key contrasts in evidentiary cultures relevant to the reduction of mechanical exposures at work.

Laboratory versus field or workplace. Scientists working predominantly in laboratories characterize a limited set of mechanical exposures using a vast array of intensive measures under highly controlled conditions. Their work has added considerably to our understanding of the biological mechanisms involved in damage to tissues due to various modes of loading (20, 21). Furthermore, laboratory testing of means to reduce mechanical exposures provides the closest approximation to the concept of efficacy in clinical research (ie, answers to the question “Can it work?” For example, testing lift assists can show that they reduce peak low-back loads when used according to specifications in lifting tasks for weights of interest (22). Yet direct application of such methods to workplaces is severely constrained due to the limits placed by the field environment on measurement options, the variety of mechanical exposures that may be operative, and the variability of conditions that fluctuate with production demands. Each of these independently precludes the kind of replication under exactly comparable conditions preferred by laboratory scientists.

Experimental versus observational. Linked to the preceding discussion, is the strong preference for more experimental designs in assessing the effectiveness of interventions, predominantly within the agricultural and clinical evidentiary traditions (23). Randomized control trials are carried out to provide the most conclusive type of evidence, followed by quasi-experimental designs (24). Although more common in safety effectiveness research (25, 26), randomized controlled trials with ergonomic intervention and individual randomization have contributed relevant evidence (27), and such trials are now underway at the workplace level (Riihimäki H, personal communication).

However, some researchers in the participatory research and action research traditions eschew the investigator control implicit in experimental designs as contrary to effective organizational intervention and the utilization of research results (28). For example, participatory ergonomic change processes produce a wide variety of changes not under the control of the investigators but, rather, under the control of the workplace parties making up the ergonomic change team (29). Furthermore, cointervention is the norm in workplaces. Workplaces suffer unexpected and often rapid market or business plan changes, workforce and manager turnover, and changes in production rates and processes that directly influence work assignments and exposure intensities in ways unforeseen at the time of both the intervention and evaluation planning (30). Griffiths (31) has delineated the limits of the natural science paradigm for organizational interventions and argued for observational designs with greater clarity in the conceptualization and examination of intervention processes. Reviewers coming from ergonomic and epidemiologic traditions have called for more adaptive quasi-experimental and observational designs (15, 16, 25, 32) to provide evidence of the reduction of hazards to inform broad, population-level public health interventions (33).

Quantitative versus qualitative. Although the dominant evaluation paradigm is quantitative, education and management researchers have used case study approaches for many years to capture the complex sets of relationships that define change in organizations (34, 35). Qualitative methods are associated with different views of
evidence and different ways of establishing rigor (36). Mergler (37) has called for greater use of qualitative approaches in occupational health research, while some organizational scientists argue that “standardized questionnaires, structured interviews, and statistical analyses cannot begin to grasp the complex fabric of organizational change [p 92]” (38). On a more moderate note, Robson et al (26) noted important roles for qualitative methods in effectiveness evaluations, particularly with respect to documentary implementation, the “how” of interventions, and the understanding of the “why” of program effectiveness (or lack thereof).

**Efficiency versus effectiveness.** In market economies, a “business case” based on cost-benefit or return-on-investment analyses often has greater persuasive value with management in workplaces and policymakers in governments than researchers’ evidence of effectiveness (39). Substantially reduced workers’ compensation costs in association with relatively small investments in workplace ergonomic programs have been a key message in reports by government agencies in the United States (3, 40). In such situations, “administrative” effect sizes (the cost savings resulting from an intervention deemed important by managers) may be far larger than biomechanical effect sizes (changes deemed important by ergonomists) or clinical ones (changes deemed important by clinicians). This circumstance may be especially true for ergonomic contributions at the design phase rather than those resulting in retrofits (41), although an estimation of avoided health risks is associated with uncertainties.

Difficulties can be associated with the estimation of efficiencies (40), often requiring close workplace–researcher partnerships to generate valid and essential data (42) similar to the requirements for valid effectiveness evaluations.

**Levels of action**

A complicating factor when workplace intervention to reduce mechanical exposures is considered is that it can be aimed at multiple levels, from workplace policies and organizational design (macro) through work group training (meso) to the level of individual tasks (micro). (See figure 1.) Earlier reviews have catalogued studies aimed at making changes across the macro–micro spectrum, from organizational structures and employee relations to job design and task requirements (3, 11, 13, 15, 25, 43, 44).

We have found it helpful to think of ways in which change at each of the levels in figure 1 (organization or company, plant or workplace, line or department, work group, job, worker, and task or tool) can alter the amplitude, time variation pattern, or duration of physical risk factors at the job or worker level. In our work, requirements for workplace ergonomic audits by corporate management in a multinational manufacturing corporation catalyzed development of a better way of releasing foam from a mold and thus reduced the amplitude of peak spinal loads associated with demolding. Similarly, a hospital workplace policy requiring at least two persons to lift a patient has the potential to reduce the amplitude of spinal loads substantially. Engineering intervention involving the redesign of the trimming and packing process on a production line reduced the frequency of twisting and pulling tasks on that job. The sharing of jobs through job rotation within a work group or job enlargement altered the time variation of physical exposures. In other work, training in workstation adjustment resulted in improvements in the proportion of workers carrying out adequate adjustments (45). In meat-processing plants, training to improve task performance, such as knife sharpening, and the modification of tools, can both be expected to reduce the amplitude of forces (46).

Evaluations need to be designed according to the primary level of an intervention. For example, the redesign of an entire production system of a plant should be evaluated with the use of a pre-post design with another similar plant, preferably within the same company, as reference. An evaluation of work-group-based ergonomic change processes would benefit from a stratification of the groups by key characteristics and then staggered implementation evaluation designs. Selecting workers with musculoskeletal problems for the random allocation of the timing of specific interventions (eg,
exercises, workstation adjustments) is appropriate for the evaluation of secondary prevention initiatives (47).

We suggest that, rather than make blanket statements about the evaluability of “ergonomic interventions”, the level of intervention be considered explicitly. Many interventions are applied as programs, to maximize change efforts (32, 40). Increasingly, multifaceted intervention is advocated (including engineering, behavioral, and administrative changes) at multiple levels within the organization, in keeping with macro- as well as micro-ergonomic approaches (48, 49). Some have argued that such an approach is the preferred way of implementing ergonomic programs (3) even though they pose challenges to the documentation and evaluation of changes at each level.

**Conditions and modifiers**

- Resources
- Commitment to change
- Management competence
- Adequate resources (personnel, monetary resources)
- Ergonomic competence
- Adherence or coverage
- Time delay
- Distribution of WMSD symptoms
- Work organization
- Workplace and work organization
- Reporting policies and procedures
- Psychological factors
- Psychosocial supports
- Therapeutic maintenance at work
- On-site physiotherapy, education
- Reactive workplace changes
- Psychosocial supports
- Therapeutic return to work
- Clinical treatment

**Nodes in path**

1. Management or workforce attitudes to practices of changing work conditions
2. Workplace changes
3. Mechanical exposures
4. Pain or discomfort
5. Reporting of pain or discomfort
6. Disability at work
7. Lost-time disability

**Indicators of conditions, interventions and outcomes**

In order to link interventions to reductions in mechanical exposure and then on to meaningful outcomes for workplace parties, we see the need for multiple indicators along a path connecting plausible steps or nodes. (See figure 2.) The model builds on earlier models (50, 51, 12) that link constructs in a cascade fashion from broader determinants at organizational levels (nodes 1–2) through targeted impacts at more micro-levels (nodes 3–4), mechanical exposure in particular, and back to broader organizational outcomes (nodes 5–7) (nodes identified in figure 2).

**Nodes 1, 2, and 3.** Of interest here are workplace attitudes and practices (node 1), which lead to changes in

---

**Figure 2.** Nodes in the path through which changes in mechanical exposures can be effected and have beneficial impacts on work-related musculoskeletal disorders (WMSD).
work organization, job design, equipment, layouts, and the like (node 2), which, in turn, may reduce measurable mechanical exposures (node 3). Considerable observational evidence exists for the important role of the first (52). Nevertheless, evaluations of intentional redesigns have had less success in demonstrating impacts of workplace changes (node 2) on mechanical exposures (node 3).

In Westgaard & Winkel’s (11) review of ergonomic interventions, of the 20 studies addressing mechanical exposure, only about one quarter evaluated the effect of the intervention on exposure. Among the studies using production system intervention or rationalization strategies, less than half evaluated the effect of the intervention on exposure. As an example, Aborg et al (53) found that work reorganization at the departmental level among Swedish office workers resulted in essentially no change in mechanical exposures as measured by surface electromyography and other means. Translating knowledge into effective actions is also a challenge that may be related to both resources and practical competencies. For example, Daltroy et al (54) showed that, although a “back school” improved knowledge, there was no observable improvement in work methods, the proxy for exposure.

Among the conditions modifying linkages between nodes are the extent of adherence by members of the workplace and the intensity of the intervention, often framed as threats to validity in quasi-experimental designs. For example, feedback from supervisors and workers on the use of recently introduced powered lifting assists in a stamping operation indicated that many people in the department were not familiar with the use of the lift assists and that breakdowns resulted in the lift assists not being available much of the time (14). Challenges in assessing intensity include issues of adherence, coverage (eg, how many workers have adequately designed ergonomic tools available), and metrics. For example, for a platform designed to reduce low back loading by angling and elevating product bins off the floor, what is the appropriate metric for load reduction: peak load (in Newtons), amount of time (in seconds or percent) less than a reference threshold (eg, 3400 N, NIOSH lifting guidelines) or cumulative load? Furthermore, should load reduction be assessed as if only one person does the task, weighted for the proportion of time at a workstation (given job rotation), or aggregated across the work group? Task-based measures may be more specific, but work-group-based measures may be more relevant to organizational outcomes (nodes 5 to 7).

Nodes 3 and 4. Linking reduced pain or discomfort (node 4) to reduced mechanical exposure (node 4) is crucial for delineating the effects of exposure reductions from other effects that interventions may have, as per Volin’s critique (12). A good example of making a plausible link is provided by Aarås (55), whose workstation changes reduced both measured physical exposures and shoulder symptoms among computer-assisted design operators. On the other hand, Demure et al (56) found a reduction in discomfort after ergonomic intervention but the reduction had little relationship to the extent of improvement. Marras and his colleagues (7) prospectively followed the impacts of workplace changes to address low-back pain. They observed that only some changes were effective in reducing exposure (using, for example, lifting aids) and hence subsequent injury reports.

Nodes 4, 5, and 6. Rates of incident reports, first aid, or first-time occupational visits are often the organizational outcomes most sensitive to the impact of ergonomic intervention (57). Unfortunately, in our experience, there is considerable variation in the threshold for reporting, the manner of data collection, the way incidents are classified, and the systems for collecting, aggregating and sharing incident reports. Linkage between levels of pain or discomfort (node 4) in a working population and reports to the workplace (nodes 5 and 6) may be loose (58, 59), as demonstrated by the “iceberg” of burden measures observable in a newspaper worker population, in which only about one-third of those with pain during the last year reported it to the workplace. Reporting practices may improve as a result of intervention, the result being paradoxical (60). Furthermore, at a given level of pain or discomfort (node 4), people may experience different problems with function (node 6), as measured by a health-related quality-of-life instrument such as the Disability of the Arm, Shoulder and Hand (61), and varying ability to carry out job tasks, as measured by an instrument such as the Work Limitations/Role Function Questionnaire (62). Webb and his colleagues (63) provide a useful filter model for understanding variation in injury reports across different levels in the workplace and in broader company administration. Because of such filters, joint measurement of pain or discomfort and function at the individual level and the recording of reporting rates at the departmental or workplace level need to become more standard practice in evaluation studies.

Nodes 6 and 7. The extent to which problems carrying out job tasks (node 6) are linked with absence from work (node 7) depends in large measure upon workplace disability management practices, including the provision of modified work (64). Variation in benefit levels, administrative procedures, and guidance given employees may result in such lost time and show up either in workers’ compensation or weekly indemnity or sickness absence rates. For this reason, some companies include
both types of indicators in their routinely reported workplace health and safety data (Reeves G, personal communication). Since wage replacement costs for absent employees are some of the major costs borne by workplaces, such data are crucial as organizational outcome, although some caution must be observed because of the relative rarity of lost-time events, the variable time required for intervention to have an effect on more serious outcomes, and the problems of carryover of assigned liabilities across years. For example, Norman & Wells (14) cite the example of an injury occurring in the year prior to the introduction of lifting assists in a manufacturing plant having a persistent and major impact on costs to the company.

An encouraging movement is to bring indicators across the nodes together within a scorecard of leading and lagging indicators (65). Leading indicators refer to the more upstream nature of programs and activities and lagging indicators belong to the more downstream nature of human health outcomes, with job conditions, such as mechanical exposure, in the middle. Unfortunately, few workplaces are currently tracking the full suite of appropriate measures relevant to nodes along the path, and therefore it is left to intervention researchers to invest resources in measuring missing indicators and to bring together such indicators into a coherent picture for workplace parties.

**Application**

So how might researchers draw on their understanding of cultures of evidence, levels of intervention, and indicators along causal paths to better evaluate workplace interventions to reduce mechanical exposures? Depending on the intervention of primary interest, the willingness of workplace parties to participate in such evaluations, and the resources available, one could imagine a diverse range of approaches employing multiple observational methods, across different levels, with multiple indicators. An area of active interest for us has been participatory ergonomic change processes of work and equipment redesign in collaboration with workplace parties.

In our research, ergonomic change teams, made up of workplace parties and researcher facilitators, go through a process of problem identification, problem characterization, solution building, and solution evaluation in an iterative fashion using a “blueprint model that draws from a variety of disciplines and parallels quality management approaches (66). Although primarily a meso-level strategy, ergonomic change teams aim at bringing about changes in the way ergonomics are incorporated into plant- or organization-wide decision making (macro level), as well as in concrete changes in work design (meso level) and equipment and tools (micro level). We have found qualitative methods of data collection (participatory observation, field notes, and interviews) and analysis (theme identification, thick description, linkage with theory) to be invaluable in better understanding the “how” of participatory ergonomic change (nodes 1 and 2) (67, 68).

To document the variety of concrete changes undertaken (eg, 27 changes on one line in one plant), we developed complementary methods to describe the equipment, task, or tool changes made (eg, an ergonomic change documentation form, node 2), perceptions of the changes by employees (ie, a 1-minute survey, node 3), and measurement of changes that occur in biomechanical exposure (node 3). The latter includes detailed observational checklists (eg, Manufacturing Operations Risk Factor) and software adaptations of laboratory-based measurement tools for more intensive measurement [eg, 4D Watbak, available at www.escs.uwaterloo.ca]. The former includes task breakdowns (aided where available by industrial engineering studies) to account for all significant loads and postures during a shift. Forces exerted on the body, as well as pinch and grip forces, are measured using appropriate force transducers. These forces, as well as postures obtained from video, are entered into biomechanical modeling and assessment software (69). This software estimates peak and cumulative loading on the low back and shoulder and can be extended to cover the distal arm (70). As well, we use surface electromyography of the shoulders and forearms to provide information on task demands, and this information appears to parallel results from epidemiologic investigations (71). Such tools permit a detailed characterization of the nature of a particular change and precise quantification of its impact on biomechanical exposures (Frazier et al, unpublished results).

For direct employee data, we primarily use an intervention-specific survey. Constructs in the survey include broad perceptions of workplace health and safety culture and communication (node 1), complementary to the qualitative data on group processes; ratings of perceived physical exertion (node 3) (eg, force required, repetitions, extent to which work is tiring, which is complementary to the ergonomic change specific descriptions); more general job characteristics [eg, influence, control, security, which both provide information on other relevant risk factors for work-related musculoskeletal disorders using well-tested instruments (8, 10)]; and pain or discomfort measures (node 4), including frequency, duration, intensity, and location, which, according to earlier work (72), demonstrate construct validity with respect to measures of function and disability. In particular, we have adapted the pain intensity measure developed by Von Korff and his colleagues (73) (pain over the last 7 days + average pain during the last 6 months + worst pain during the last 6 months) and have found
it has superior psychometric performance as a continuous measure of severity when compared with other possible outcome measures for work-related musculoskeletal disorders (Smith et al, unpublished results).

Finally, at the plant or departmental level, we have obtained existing human resources and operations management data to assess changes in both injury- and health-related outcomes [eg, first aid, absenteeism (nodes 5–7)] and production outcomes (eg, right-first-times) meaningful to workplace parties (74, 75). We carry out this activity over three time periods, before researcher involvement, during the ergonomic intervention process and after major researcher involvement with the plant when ergonomic change teams continue working but in a less intense manner.

Such a suite of indicators of both change processes and outcomes at different levels and for different links along the path can be grouped into case studies for each experience (76). Within the case studies, qualitative findings can be compared with quantitative findings and comparisons made across the different kinds of quantitative data with the intent of “triangulation” or cross-validation (77). Such enriched case studies can add to the extensive existing case study literature on ergonomic intervention.

To enhance inference on the impact of ergonomic intervention, several additional design and analytical steps are being taken. First, we are seeking to characterize the nature and extent of ergonomic changes and the factors influencing both the changes made and the impacts occurring in order to compare and contrast experiences in a multiple case-study design (82). Second, we have staggered the timing of ergonomic intervention in paired departments, plants, or worksites and permitted the later intervention sites of the pairs to act as time-based referents in keeping with quasi-experimental approaches (26). For each of these two actions, the full range of indicators is appropriate. Third, by concentrating on similar-sized worksites and using the same person-based measures in each worksite, our aim is to pool person-based data across experiences to achieve sufficient sample size for quantitative inference testing. Depending on the homogeneity in observed variable distributions, this action can occur either directly or using meta-analytic approaches. Through each of these approaches, we look for convergence or divergence in impacts and the reasons for each.

Our approach, which includes many of the features of both methodologically rigorous and societally relevant intervention evaluation, should provide evidence for the effectiveness of meso-level workplace ergonomic intervention that speaks to the majority of cultures of evidence that we described earlier. It does not exclude the possibility of considering more rigorous designs in future work, such as randomized trials across a range of medium-sized workplaces, provided that key elements of effective intervention in appropriate contexts can be identified in our current work and those of others, institutional support can be gleaned from organizations of workplaces that provide an appropriate sampling frame, workplace parties in participating worksites can provide access to production, quality and injury data and agree to randomized time-lagged intervention with ongoing participation and monitoring tasks, sufficient similarity of the intervention process and comparability of the specific changes can be assured, ongoing intervention can be documented and accounted for, and substantially greater amounts of research resources can be mobilized. Currently such a list is a tall order in most jurisdictions, but, as the pressure to produce better evidence increases, we can only hope that research responses can be similarly enriched in ways that are meaningful for workplace stakeholders, who make the fundamental decisions determining workplace biomechanical exposures (78).

**Acknowledgments**

We thank the workplace parties (companies, unions and employees) that we have worked with over the last 5 years and that have assisted us in understanding problems, testing responses, and building frameworks.

Financial support for our work has been provided by the National Centres of Excellence via the Health Evidence & Application Linkage Network (HEALNet) and the Workplace Safety and Insurance Board through the Ontario Research Advisory Committee’s Solutions for Workplace Change extramural grant program and directly to the Institute for Work & Health.

**References**

5. Punnett L, Bergqvist U. Visual display unit work and upper


38. Badham R, Couchman P, Little S. Getting smart: developing an action research approach to the integrated management of technical and organizational innovation. Hum Syst Manage

Scand J Work Environ Health 2003, vol 29, no 5 403


Received publication: 16 June 2003