Predictive Modeling of risk factors in slaughterhouses using Low-cost inertial sensors

Thesis Defense

Adolfo Villalobos

Magister en Ciencias de la Ingenieria
Departamento de Ingenieria Industrial
Pontificia Universidad Catolica de Chile

15 November 2020

Outline

Section 1

Work-Related Muscukuloskeletal Disorders in Slaughterhouses

Motivation

State of the Slaughterhouse Industry (BMUB 2015/16)

1. Labour is a high pertcentage of the costs.
2. Fatigue and bad practices lead to injuries, absenteesim and costs.

Motivation

State of the IoT industry

1. IoT market has grown dramatically in the last five years.
2. IoT market share by sector shows that wearable technology still represent a small proportion of the total (3%).

Motivation

Oportunity

1. Sensing technologies have been used in human activity recognition tasks, particularly in sports.
2. This shows an opportunity of improvement in occupational domains. The benefits are: Risk and lesion reduction, lower fatigue, increase productivity and increase efficiency.

Literature Review

Assessment of WRMSDs

1. Work-Related Muscukuloskeletal Disorders (WRMSDs) are the leading cause of work disability, sickness and absence from work [Bevan, 2015].
2. RULA and OCRA Check List surveys are the standard assessment methodology [McAtamney and Nigel Corlett, 1993, Occhipinti, 1998].
3. Microsoft Kinect 3D cameras or intricate arrangements of IMUs, placed over the worker's body, have been used with good results [Vignais et al., 2013, Buisseret et al., 2018, Chen et al., 2018]
4. Others have used Deep Learning black-box models in work environments to detect bad postures [Barkallah et al., 2017, Abobakr et al., 2017, Hu et al., 2018].

Literature Review

Assessment of WRMSDs

A review of predictive modeling techniques in occupational domains showed that practitioners still haven't fulfilled the promised of using IoT technologies to assess risk factors [Lim and D'Souza, 2020].

Limitations

In summary, currently approaches are powerful but hard to be implemented outside the lab due to high model complexity, high sensor cost, highly dependent on controlled conditions or resisted by workers.

Section 2

Thesis Hypothesis and Objectives

Objectives

Strategic Guidelines

1. System is innocuous to the worker and the product.
2. Generate an indicator that can be used to benchmark and evaluate workers.
3. Design a methodology that can be extrapolated to other cutting-tasks.
4. Avoid complex models if not needed.

Objectives

Objective 1

Model the presence of the risk factors, during a meat-cutting task, as a time series classification problem.

Objectives

Objective 2

Determine if the information obtained from low-cost sensors, placed in the wrists of the worker, and combined with expert supervision, are sufficient to accurately assess the presence of risk factors. For this, we focus on two risk factors

1. The prediction of the RULA score, related to ergonomic risk and improper technique [Viikari Juntura, 1983].
2. The presence of a knife with a compromised blade, since bluntness has been found to increase the likelihood of WRMSDs by increasing the necessary exertions [Marsot et al., 2007, Karltun et al., 2016, Savescu et al., 2018].

Objectives

Objective 3

Determine if the developed predictive modeling tools for the assessment of WRMSDs can be used to quantify the economic benefits of preventive decision making.

Objectives

Objective 4

Analyze if its possible to accurately predict whenever the worker begins and ends a cut.

Hypothesis

Research Questions

1. Is it possible to gather information from low-cost accelerometers, placed wrists of slaughterhouse workers and use it as input for machine learning algorithms that accurately predict the presence of risk factors in cutting activities.
2. Can we use the predictions of risk factors as a replacement for human ergonomic supervision and prevention.
3. Is it possible to assess risk factors with limited human supervision, and only relying on auxiliary predictions.
4. There exists a positive relationship between a reduction in ergonomic risk and improving productivity.

Section 3

Methodology, Prototype & Experiments

Methodology

Initial approach

1. Measure workers.
2. Process data and try to detect patterns, i.e, observe box-plots, clustering.
3. Consult with experts and review ergonomic literature to see if insights made sense.
4. Iterate.

Methodology

Key components

1. We model each risk factor of interest as a supervised learning problem.
2. Hence, for each problem we will need to provide supervision. This is done by performing measurements on workers.
3. Transform sensor data into features that can be used in supervised learning. (e.g. feature matrix in regression problems).
4. Ensemble a supervised learning pipeline: Pre-processing, encoding, standardization, feature engineering, training, validation (via cross validation) and test. On each iteration a set of parameters is used. We choose the best one.
5. Obtain ergonomic insight for decision making.

Methodology

Strategic Guidelines

1. System is innocuous to the worker and the product.
2. Generate an indicator that can be used to benchmark and evaluate workers.
3. Design a methodology that can be extrapolated to other cutting-tasks.
4. Avoid complex models if not needed.

Prototype

Components

The prototype had 3 components.

1. The Graphical User Interface (GUI).
2. The wristband-based devices.
3. The Batch Layer.

Experiments

Sample population

2 instructors and 18 Non-senior workers with varying degrees of experience.

Experiments

Work tasks

Three work tasks.

1. Instructor 1: Complete processing of the meat product
2. Instructor 2: Femur and coxal deboning
3. Non-senior workers: Femur deboning.

Experiments

Assessment of WRMSDs

We propose a workflow for the assessment of WRMSDs based on the application of consecutive machine learning problems, using their predictions as input for a decision making rule.

Section 4

Predictive Modeling & Decision Making

Supervised Learning Fundamentals

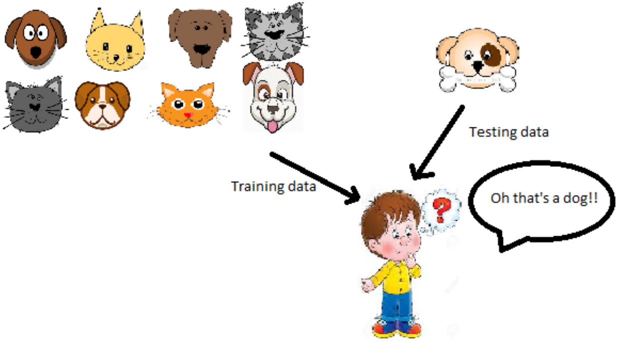


Figure: Supervised Machine Learning. URL:
https://medium.com/campusx/machine-learning-from-a-5-year-old-perspective-93278fbf73ce

Supervised Learning I

Mathematical Framework

Supervised Learning II

1. Consider a collection of n examples $z_1, \ldots, z_n$ represented by a tuple $z_i := (x_i, y_i)$ composed by a vector of features $x_i \in \mathbb{R}^n$ and a label y_i.
2. The objective is to estimate a function $f : X \to Y$ that maps the observed features into their respective label.
3. The function is approximated from a set of candidates indexed by a parameter:

$$F = \{f_\theta : \theta \in \Theta\}$$

4. The quality of the approximation is measured by defining a loss function $L(\theta)$ that measures the distance l between the true value y_i and the predicted values $\hat{y}_i = f_\theta(x_i)$.

$$L(\theta) := \sum_{i=1}^{n} l(y_i, \hat{y}_i)$$

5. The optimization problem then becomes finding

$$\hat{\theta} = \arg\max_{\theta in \Omega} \{L(\theta) + \Omega(\theta)\}$$

Time Series Classification

TSC

1. Cutting-tasks are modeled as multivariate time series. A time series is a discrete sequence of data points $\{x_t\}_{t \in \mathbb{N}}$, where each x_i lives in some vector space (e.g. $\mathbb{R}^m$)
2. Time Series Classification (TSC) is a hard problem, highly dependent on data types and domain context.

TSC

Three types of models: Distance-based, Model-based and Feature-based.

TSC Example

Visual example of Time series classification. Use tikz.

Feature Engineering

Features

We considered:

1. Statistical Features.
2. Ergonomic Threshold Features.

The threshold features are randomly sampled from a grid, and selected by evaluating performance on the test set.

Feature Engineering

Hyperextension Feature

The threshold feature T_1 is defined as the number of times that the kinematic variable X surpasses a threshold value φ, indicating the presence of hyperextension.

Maximum Exertion Feature

The threshold feature T_2 is defined as the percentage of the time in which the kinematic variable X surpasses a proportion $\gamma \in [0, 1]$ of the Maximum observed value, indicating the proportion of time in which the worker is performing dangerous tasks.

Performance

Evaluation Metrics

1. The classifier is evaluated in the test set using evaluation metrics.
2. Common metrics are Accuracy, Precision and Recall.
3. In the binary classification case, there are TP, TN, FN and FP. Metrics are calculated acording different aspects of the distirbution.
4. For the multivariate case, the extension is simple.

Performance

For each machine learning problem assessing each risk factor we consider different metrics. Accuracy focuses on overall performance, precision focuses in reducing false positives and recall focuses on detecting all the negatives.

Performance

Table: Proposed evaluation metrics for WRMSDs ML problems

ML Problem	Objective	Dataset	Training Evaluation	Test Evaluation
CTC-6, CTC-11	Multi-class	D_1	$F_{1(m)}$	Accuracy
CTC-2	Binary	D_2	F_1	Accuracy
KEC-6*, KEC-11	Binary	D_1	F_1	Recall
KEC-2	Binary	D_2	F_1	Recall
PRP	Multi-class	D_3	$F_{1(m)}$	Accuracy

*: Two versions of KEC task are studied on dataset D_1 (6 and 11 cutting-task types); and one on D_2 (2 cutting-task types). However, the problem is still binary (blunt and sharp k

Section 5

Results

Implementation

Stack

The implementation was done using the following frameworks

1. Python 3.7 (NumPy, Pandas and Scikit-Learn).
2. MySQL 8.0.
3. Jupyter Notebooks for experimentation.
4. AWS EC2 for computing.

Implementation

ML Parameters

The implementation was done using the following parameters

1. DT, SVM, RF and ET models.
2. Statistical and Ergonomic Features.
3. Jupyter Notebooks for experimentation.
4. AWS EC2 for computing.

Classification Results

Task	Model	NF	Evaluation Metrics				F	RH						LH					
			A	P	R	F_1		a	v	p	Pitch	Roll	Yaw	a	v	p	Pitch	Roll	Yaw
CTC-11	*DT*	42	0.44	0.45	0.42	0.41	T_1	-	-	1.9	-	-	179.3	0.28	-	-	-	-	150.3
							T_2	-	-	-	-	-	0.41	-	-	-	-	0.41	0.41
							S	a,b,c,d,e	a,b,c	a,b,c,d,e	b,c,d	a,c,d	d,e	a,c,d	a,d,e	a,c	b,d	-	b,d,e
	ET	32	0.51	0.55	0.53	0.51	T_1	0.65	69.8	-	-	171.2	-	0.34	-	7.2	37.5	-	-
							T_2	-	-	-	0.47	-	-	-	-	-	-	-	-
							S	b,c,d,e	a,b,c,e	a,c,e	a,b,c,e	b,c,d	-	-	a,b,e	a,e	e	-	-
CTC-2	*DT*	6	0.76	0.89	0.74	0.81	T_1	1.35	106.2	-	-	-	-	-	-	-	-	-	-
							T_2	-	-	-	0.64	-	-	-	-	-	-	-	-
							S	e	-	-	-	a,d	-	-	-	-	-	-	-
	RF	57	0.76	0.86	0.78	0.82	T_1	1.07	98.2	-	-	92.6	-	-	287.6	2.5	20.1	52	-
							T_2	-	-	-	0.44	0.44	-	-	-	-	-	0.44	0.44
							S	a,b,c,d,e	b,c,e	a,b,c,d,e	a,c,e	a,b,c,d,e	a,b,d,e	e	a,c,e	b,c,e	c,d,e	a,b,c,d,e	a,b,c,d,e
CTC-6	*ET*	47	0.71	0.76	0.72	0.71	T_1	0.35	363.8	0.4	11.2	173.1	-	0.03	124.7	10	56.6	-	-
							T_2	-	-	-	-	-	0.87	-	-	-	0.87	0.87	-
							S	b,c,d	a,b,c,e	a,b,c,e	a,b,c,d,e	b,c,d	d	b,c,d	a,b,c,e	a,b,c,e	-	b,c	d
	SVM	12	0.67	0.71	0.67	0.67	T_1	-	61.1	4.1	-	-	-	0.67	-	-	-	-	-
							T_2	-	-	-	-	-	-	-	-	-	-	-	-
							S	-	a,c,e	a,c	c	b,c	-	-	-	-	-	-	-
KEC-11	*ET*	51	0.81	0.8	1	0.89	T_1	-	346.7	7.4	-	-	19.5	-	238	13.6	-	159.5	-
							T_2	-	-	-	0.48	0.48	0.48	-	-	-	-	0.48	0.48
							S	a,b,c,e	a,c,e	a,b,d	a,c	a,c	b,c,d,e	b,c,d	a,b,d,e	a,b,c,d,e	b,d	a,b,c,e	b,c,e
	SVM	22	0.81	0.81	0.98	0.89	T_1	-	-	-	-	-	108.8	-	-	-	-	-	-
							T_2	-	-	-	-	-	-	-	-	-	-	-	0.72
							S	b	a,d,e	a	a,e	c	a,b,c,d,e	b	a,e	a,e	-	-	b,d
KEC-2	*DT*	2	1	1	1	1	T_1	-	-	-	-	-	-	-	-	-	-	-	95.4
							T_2	-	-	-	-	-	-	-	-	-	-	-	-
							S	-	-	-	-	-	-	-	-	-	-	-	d
	ET	7	1	1	1	1	T_1	-	-	-	-	-	-	-	-	-	-	-	25.1
							T_2	-	-	-	-	-	-	-	-	-	-	-	-
							S	-	-	-	-	-	c	-	-	-	c	-	a,c,d,e
KEC-6	*ET*[a]	32	0.86	0.85	1	0.92	T_1	-	375	-	52.5	-	102.3	-	-	-	-	-	-
							T_2	-	-	-	-	-	-	-	-	-	0.49	-	-
							S	a,e	d	-	a	-	b,c,e	a,b,e	b,e	a,b,c,d,e	a,b,d,e	b,c,d	a,e
	SVM	12	0.74	0.82	0.85	0.84	T_1	-	-	-	-	-	-	-	-	-	-	-	-
							T_2	-	-	-	-	-	-	-	-	-	-	-	-
							S	-	-	-	a,e	-	a,b,c,d,e	-	e	a,d,e	b	-	-

NF: Number of features selected by the algorithm; F: Feature type (T_1, T_2 or Statistical)
T_1: Threshold parameter selected for the hyper-extension feature
T_2: Threshold parameter selected for the maximum exertion feature
S: Statistical Feature selected: a-mean, b-std., c-maximum, d-minimum, e-median

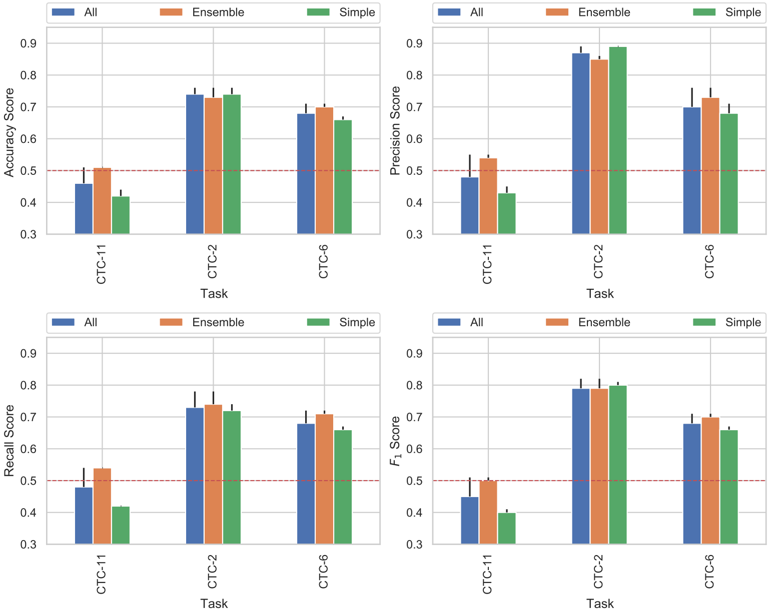
Classification Results

Table 6: Summary of selected features and parameters of the best performing algorithms on the PRP tasks

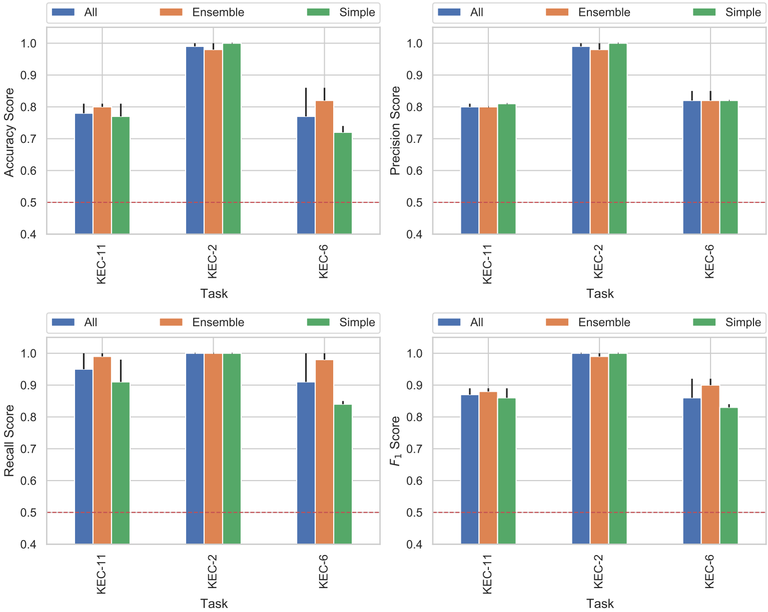
Task	Model	NF	Evaluation Metrics A	P	R	F_1	F	RH Pitch	RH Roll	RH Yaw	LH Pitch	LH Roll	LH Yaw
PRP	RF^{exp}	32	0.98	0.99	0.97	0.98	T_1	25.8	-	13.6	-	43.2	-
							T_2	-	-	0.93	-	-	-
							S	a,b,c, d,e	a,b,c, d,e	a,b,c, d,e	a,c	a,b,c, d,e	a,b,c, d,e
	$SVM^{exp,t}$	42	0.96	0.97	0.95	0.96	T_1	87.4	78.8	23.2	-	33.2	11.1
							T_2	0.76	0.76	0.76	-	0.76	0.76
							S	a,b,c, d,e	a,b,c, d,e	a,b,c, d,e	a,b,c, d,e	a,b,c, d,e	a,b,c, d,e

NF: Number of features selected by the algorithm; F: Feature type (T_1, T_2 or Statistical)
T_1: Threshold parameter selected for the hyper-extension feature
T_2: Threshold parameter selected for the maximum exertion feature
S: Statistical Feature selected: a-mean, b-std., c-maximum, d-minimum, e-median
"-": The feature was not selected by the associated algorithm
^{exp}: In the PRP task, the algorithm selected the experience of the worker as relevant
^{t}: In the PRP task, the algorithm selected the duration of the task as relevant

Classification Results



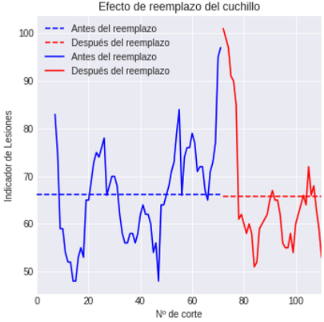
Classification Results



Effect of Knife Replacement

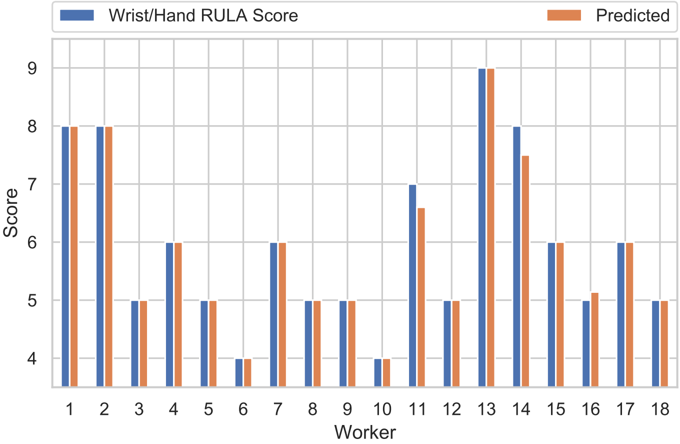
(a) Operario n15

(b) Operario n8



(c) Operario n2

Assessment of WRMSDs Risk



Assessment of WRMSDs Risk

(a) Confusion matrix of the WRMSDs risk assessment of the average cutting-task (n=18)

		Predicted Low	Medium	Very High
	Low	100%(2)	0	0
True	Medium	0	100 % (11)	0
	Very High	0	20% (1)	80% (4)

(b) Confusion matrix for the WRMSDs risk assessment of all cutting-task

		Predicted Low	Medium	Very High
	Low	100%(15)	0	0
True	Medium	0	100 % (102)	0
	Very High	0	5% (3)	95% (57)

Section 6

Conclusions

Conclusions

Findings

We showed that using Low-Cost IMUS and a data-driven approach:

1. Is possible to detect the type cut that the worker is performing.
2. We developed an accurate classifier to predict the presence of risk factors due to hyperextensions.
3. We can detect when the knife has lost his sharpness.

Conclusions

[

Limitations] There are limitations

1. Accuracy in cutting-task type is not high enough.
2. Deterioration of knife is discrete, not continuous.
3. We need supervision of the beginning and end of each cutting task.

Conclusions

[

Future Work] Future work:

1. Algorithms to determine when worker is performing a task or not.
2. Use IMU for retraining purposes.
3. Develop a DSS framework.
4. Determine the productivity of the collaborator
5. Build a production-ready version of the prototype.

Personal Conclusions

Challenges

1. Chance to work on a real industry problem.
2. Problem definition and strategy to solve it.
3. Interact with end-users and decision makers.
4. Opportunity to learn about techniques outside my domain of expertise.
5. Work with real-world dataset.

References I

Abobakr, A., Nahavandi, D., Iskander, J., Hossny, M., Nahavandi, S., and Smets, M. (2017).
RGB-D human posture analysis for ergonomic studies using deep convolutional neural network.
In *2017 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2017*, volume 2017-Janua, pages 2885–2890. Institute of Electrical and Electronics Engineers Inc.

Barkallah, E., Freulard, J., Otis, M. J., Ngomo, S., Ayena, J. C., and Desrosiers, C. (2017).
Wearable devices for classification of inadequate posture at work using neural networks.
Sensors (Switzerland), 17(9).

Bevan, S. (2015).
Economic impact of musculoskeletal disorders (MSDs) on work in Europe.

References II

- Buisseret, F., Dierick, F., Hamzaoui, O., and Jojczyk, L. (2018).
  Ergonomic risk assessment of developing musculoskeletal disorders in workers with the microsoft kinect: Track tms.
  IRBM, 39(6):436–439.

- Chen, D., Cai, Y., Cui, J., Chen, J., Jiang, H., and Huang, M. C. (2018).
  Risk factors identification and visualization for work-related musculoskeletal disorders with wearable and connected gait analytics system and kinect skeleton models.
  Smart Health, 7-8:60–77.

- Hu, B., Kim, C., Ning, X., and Xu, X. (2018).
  Using a deep learning network to recognise low back pain in static standing.
  Ergonomics, 61(10):1374–1381.

- Karltun, J., Vogel, K., Bergstrand, M., and Eklund, J. (2016).
  Maintaining knife sharpness in industrial meat cutting: A matter of knife or meat cutter ability.
  Applied Ergonomics, 56:92–100.

References III

Lim, S. and D'Souza, C. (2020).
A narrative review on contemporary and emerging uses of inertial sensing in occupational ergonomics.
International Journal of Industrial Ergonomics, 76(November 2019):102937.

Marsot, J., Claudon, L., and Jacqmin, M. (2007).
Assessment of knife sharpness by means of a cutting force measuring system.
Applied Ergonomics, 38(1):83–89.

McAtamney, L. and Nigel Corlett, E. (1993).
RULA: a survey method for the investigation of work-related upper limb disorders.
Applied Ergonomics, 24(2):91–99.

Occhipinti, E. (1998).
OCRA: A concise index for the assessment of exposure to repetitive movements of the upper limbs.
Ergonomics, 41(9):1290–1311.

References IV

- Savescu, A., Cuny-Guerrier, A., Wild, P., Reno, G., Aublet-Cuvelier, A., and Claudon, L. (2018).
  Objective assessment of knife sharpness over a working day cutting meat.
  Applied Ergonomics, 68(April 2017):109–116.

- Vignais, N., Miezal, M., Bleser, G., Mura, K., Gorecky, D., and Marin, F. (2013).
  Innovative system for real-time ergonomic feedback in industrial manufacturing.
  Applied Ergonomics, 44(4):566–574.

- Viikari Juntura, E. (1983).
  Neck and upper limb disorders among slaughterhouse workers. An epidemiologic and clinical study.
  Scandinavian Journal of Work, Environment and Health, 9(3):283–290.