DUTCH: A new tool for practitioners for risk assessment of push and pull activities

Douwes, Marjolein; Könemann, Reinier; Hoozemans, Marco; Kuijer, Paul; Vermeulen, Hetty

published in
Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018) - Volume III
2019

DOI (link to publisher)
10.1007/978-3-319-96083-8_79

document version
Publisher's PDF, also known as Version of record

document license
Article 25fa Dutch Copyright Act

DUTCH: A New Tool for Practitioners for Risk Assessment of Push and Pull Activities

Marjolein Douwes[1(✉)], Reinier Könemann[1], Marco Hoozemans[2], Paul Kuijer[3], and Hetty Vermeulen[4]

[1] TNO, PO Box 3005, 2301 DA Leiden, The Netherlands
marjolein.douwes@tno.nl
[2] Department of Human Movement Sciences, Faculty of Behavioural and Movement Sciences, Vrije Universiteit, Amsterdam Movement Sciences, Amsterdam, The Netherlands
[3] Netherlands Center for Occupational Diseases, Coronel Institute of Occupational Health, Academic Medical Center (AMC), University of Amsterdam, Amsterdam Public Health Research Institute, Amsterdam, The Netherlands
[4] vhp human performance, The Hague, The Netherlands

Abstract. Pushing and pulling at work is an undervalued theme within occupational health policy. This is unjustified, because these activities are very common and potentially increase the risk of shoulder symptoms. Gaining insight into the possible health risks of specific push or pull activities at the workplace is a first step towards prevention of musculoskeletal symptoms. Existing instruments proved to be insufficiently suitable to give that insight in a simple way. This was the motivation for developing the Push and Pull Check (DUTCH). This method makes clear whether the push or pull activity is acceptable or not, which risk factors exist, and which measures can reduce the risk. This article describes the operation of the DUTCH, as well as the development of the tool.

Keywords: Physical workload · Pushing · Pulling · Risk assessment Prevention · Shoulder · Low back

1 Introduction

Manual force exertion and pushing and/or pulling at work increase the risk of developing low back and shoulder pain [1, 2]. In particular, performing push or pull activities at work poses a high risk of shoulder symptoms. This was concluded from a systematic review of seven studies, covering a total of 8,279 employees, in which the risk of shoulder symptoms among workers with high exposure to pushing and pulling was between two and five times higher than in workers that did not perform pushing and pulling activities [2].

It is important that companies are aware of the health risks of pushing and pulling at work. Health and safety regulations do not offer specific health limits for pushing and pulling. It is also not possible to draw up these health limits based on available epidemiological literature [1]. The Dutch Health Council's recommendations stated

S. Bagnara et al. (Eds.): IEA 2018, AISC 820, pp. 604–614, 2019.
https://doi.org/10.1007/978-3-319-96083-8_79

that the ‘Mital tables’ [3] provide the best available information to assess and evaluate the physical work demands of pushing and pulling tasks. These tables provide data from psychophysical research on the - self reported – maximum manual force exertion under different conditions. However, the practical application of these tables requires force measurements at the workplace, which is often not feasible in practice. Existing practical instruments that do not require force measurements, such as the Key Indicator Method (KIM) [4, 5], the PushPullCalculator (PPC) [6, 7] and the pushing and pulling Operations Risk Assessment Tool [8] do not meet all criteria that apply to a good practical tool. For example, the KIM and the HSL tool are not sufficiently evidence-based, while the DTC does not take into account environmental factors of pushing and pulling tasks, such as characteristics of the floor surface and of the wheels of a trolley. These factors potentially have a large effect on the rolling resistance and thus on the hand force needed when pushing or pulling a container [9].

Thus, a new practical tool was developed, with the aim to give companies quick and simple insight into the presence of risk factors of pushing and pulling at work.

2 Pre-study on Inter-rater Reliability and Face Validity of Existing Tools

2.1 Objective

An important criterion for a risk assessment method is that it produces reliable and valid results. No information was available on the reliability and validity of existing tools for the evaluation of pushing and pulling at work. To get an impression of these qualities, we conducted a brief study on the inter-rater reliability and the ‘face-validity’ of the KIM and the DTC. In this paper, face-validity means: the level of agreement between the results of the instruments with the judgments of a group of experts in the field of physical workload on the evaluation of the potential health-risk associated with performing specific pushing and pulling activities. The final version of the HSL assessment tool and data presented by Weston et al. [10] were not yet available for this study.

2.2 Methods

For this study we selected ten push-pull tasks that had been studied before by one of the experts involved. These ten tasks involved various pushing and pulling activities with horizontally-oriented hand forces, with a variable level of force exertion, that had been measured according to a valid protocol. Characteristics of the tasks are listed in the first column of Table 1. The experts provided descriptions of these tasks in a standardized format, which enabled the KIM and DTC to be applied. Some of the tasks lacked information on the frequency of pushing or pulling, because they were experimental situations. In those cases, realistic estimates were used. Eight experts independently evaluated the push and pull tasks with the KIM and the DTC. They translated the task descriptions to the required input data of the two instruments. In addition, they answered some questions on the exact data and methods they used for the evaluation, the difficulties they encountered, the advantages and disadvantages of the two instruments and the extent to which the outcome was in line with their expert judgement on the physical demands of the tasks (red, orange or green). Green means ‘at least 90% of the population can maintain this task for eight hours’, orange means ‘between 25% and 90% can maintain this task’ and red means ‘25% of the population or less can maintain this task for eight hours’.

2.3 Data Analysis

To determine the face validity of the KIM and DTC, their results were compared with that of the experts’ judgments, which were used as the reference standard in this study. To reach a single final judgement from the three parts of the DTC (assessment of the hand force, low back load and shoulder load), the lowest limit value (strictest assessment) was used. For the inter-rater reliability, the percentage of absolute agreement between red-orange-green assessments of the evaluators was calculated. In addition, Kappa values were calculated for the individual scores of the DTC and the final score of the KIM. For the interpretation of Kappa results the following cut-off points were used: ‘low’ (0–0.20); ‘moderate’ (0.21–0.40); ‘reasonable’ (0.41–0.60); ‘sufficient to good’ (0.61–0.80) and ‘almost perfect’ (0.81–1.00) [11].

2.4 Results

Face-Validity

Table 1 shows the results of the KIM and DTC, and expert judgments for the ten push/pull tasks. When comparing the results from the KIM with the expert judgments, they are in agreement for four tasks, more strict (orange instead of green) for three tasks and less strict (orange instead of red) for three tasks. The ratings of the DTC are in agreement with the expert judgments for five tasks, more strict for four tasks (twice red instead of green and once red instead of orange) and less strict for one task (green instead of red). Overall, KIM reviews were orange for seven out of ten tasks; DTC reviews were red for seven out of ten tasks. Also notable is the fact that only one of the tasks (task 2) was judged the same by the KIM and DTC.

Inter-rater Reliability

In Table 2, the percentage of agreement in categorical scores (red, orange, green) between the experts is displayed for the DTC and the KIM. The overall percentage of agreement for the KIM is 81%. Since the DTC does not present an overall score, the level of agreement is presented for hand force, shoulder load and low back load separately. The lowest level of agreement for DTC (62%) was found for the shoulder load (sustained) and the highest level of agreement (99%) was found for the low back load. To give a final judgment for DTC the experts used a ‘worst case’ assessment; the agreement percentage of this judgment was 91%. In addition to the percentage of agreement, the kappa values and their classifications are presented for both tools. The kappa rating for the KIM risk score is ‘sufficient to good’ (0.705) and for the DTC it varies from ‘reasonable’ to ‘almost perfect’ (0.447–0.967), but is ‘good’ (0.833) for the DTC worst case assessment.

Table 1. Average reviews and corresponding red-orange-green judgements according to the KIM, most common judging according to the DTC and consensus judgments from experts

Tasks	KIM: mean final score KIM (sd)	DTC: most common score* (% agreement)	Consensus expert judgment
Pushing a small trolley with food in the train	44 (14)	Green (63%)	Green
Postal expedition: pushing/ pulling carts in distribution centre	54 (14)	Red (100%)	Red
Postal distribution: pushing/ pulling carts in distribution centre	43 (8)	Red (100%)	Green
Level out concrete floor with an electrical (vibrating) rei	85 (22)	Green (86%)**	Red
Move a hand pallet truck in a warehouse	35 (12)	Red (100%)**	Orange
Move trolleys to and in trucks	29 (8)	Red (100%)	Red
Move garbage containers along tiles with 1 or 2 persons	39 (13)	Red (100%)	Red
Move garbage containers along tiles with helping device	30 (13)	Red (100%)**	Green
Move garbage containers along tiles without helping device	38 (17)	Red (100%)**	Red
Pull money carts along carpet in casino	7 (1)	Orange (57%)**	Green

* The DTC does not present a 'overall score '. Therefore, the highest subscore per evaluator and most prevalent score among all evaluators was used; **N = 7.

2.5 Discussion

Face Validity

Both the DTC and KIM resulted in a large number of assessments that did not equal the expert assessment: six out of ten tasks for KIM and five out of ten for the DTC. Overall, the assessments of the KIM were more in line with the expert assessments because the

Table 2. Percentage of agreement and Kappa-score to indicate the agreement between the eight raters, for ten tasks, both for DTC and KIM categorical results

Evaluation	% agreement	Kappa	Kappa classification
DTC, initial hand force (to set in motion)	92%	0.853	Good
DTC, sustained hand force (to keep in motion)	79%	0.680	Sufficient to good
DTC, back strain at onset	97%	0.933	Almost perfect
DTC, back strain – sustained	99%	0.967	Almost perfect
DTC, shoulder strain at onset	78%	0.567	Reasonable
DTC, shoulder strain – sustained	62%	0.447	Reasonable
DTC, worst case	91%	0.833	Good
KIM, risk score	81%	0.705	Sufficient to good

number of overestimations was equal to the number of underestimations, whereas the DTC overestimated more severely and more frequently. However, some of the tasks that scored ‘red’ by the experts were not scored ‘red’ when using the KIM. In those cases, the KIM underestimated the physical load according to the experts. Moreover, both methods have less responsive character than the experts: seven out of ten reviews with the KIM are orange, nine out of ten reviews with the DTC are red.
The face-validity of an instrument also depends on its scientific background. The DTC is based on Hoozemans et al. [12] for the estimation of exerted hand forces, low back load and shoulder load, and based on Mital et al. [3] for the evaluation of the hand force; on Jäger [13] for the evaluation of the low back load and on Chaffin et al. [14] for the evaluation of the shoulder load. At the time of this study, the authors of the KIM could not provide scientific background for the KIM. Moreover, the KIM results do not appear to be associated with the Mital tables [3]. The KIM evaluates pushing and pulling with a low frequency as less hazardous and with a high frequency as (much) more hazardous than the Mital tables.

A possible explanation for the differences between assessments with both instruments on one hand and expert reviews on the other hand is that the KIM takes environmental factors into consideration while the DTC does not. Moreover, there is variation in the expert judgments that were used as a ‘reference standard’ in this study. This is probably due to different opinions of the experts about the risk of exposure for work-related musculoskeletal complaints.

Inter-rater Reliability. De kappa-classification for the DTC varied from ‘reasonable’ to ‘almost perfect’(0.45–0.97), and was ‘good’ (0.83) for the ‘worst case’ assessment, which is the most important result of the DTC. Mutual differences in results can be explained by difference in interpretation of the information on the task provided for using KIM and DTC. Examples are differences in determining the male/female distribution in the population, working posture and working conditions; these factors are difficult to estimate and generalize per task.

2.6 Limitations and Conclusions

This research has limitations. Firstly, there was little variation between the experts' assessments of the tasks, which were 'red' relatively often. This may give a too one-sided view. Secondly, there was no need for an interpretation by the evaluator for some pre-provided data. However, in practice variation may arise for these factors, which would result in a lower reliability than we found in this study. Thirdly, this study was carried out with experts trained in the application of similar methods. Application of the KIM and DTC by users without prior knowledge is expected to provide lower reliability as they will have more difficulty in determining the requested input data. To apply a tool as we envision, with little or no prior expertise on physical workload assessment or ergonomics training is probably a major constraint.

Both the KIM and the DTC have a moderate face-validity and inter-rater reliability to assess the work-related risk of pushing and pulling. Therefore, and due to the lack of scientific evidence for the KIM, it was concluded that a new tool should be developed that is not based on either the KIM or the DTC but should incorporate the strengths of both methods. In addition, the new tool should be evidence-based, present a clear overall final judgement and provide insight in the main hazards, include both the average and peak load in the assessment to avoid 'means' of extreme situations, mention conditions for which the tool is not applicable (e.g. sliding of objects) and include recommendations for risk reduction and give insight into the effect of small improvements, to encourage workers and donors to take and use measures.

3 Development of DUTCH

3.1 Maximum Acceptable Push and Pull Forces

In theory, limits for an acceptable workload should be deduced from epidemiological literature, biomechanical models and/or psychophysical experiments. Through a concise literature review, which we do not discuss here, recent relevant epidemiological literature was studied. In the literature a strong relationship is described between pushing/pulling and the prevalence of shoulder symptoms [2]. However, the literature does not provide sufficient guidance to define clear limits above which the risk of shoulder symptoms strongly increases. Moreover, biomechanical shoulder models are insufficient for a scientific criterion. Available shoulder models, which include muscle load around the shoulder joint, indicate limited health limits. Therefore, it was decided to use the psychophysical Snook tables [15] for the determination of the maximum acceptable push and pull hand forces. These tables correspond to those of Mital et al. [3] but provide more extensive data on the frequency of pushing and pulling and population percentiles.

3.2 Snook Tables

Based on self-reporting, the Snook tables present the maximum acceptable horizontal hand force (Newton) for pushing or pulling rolling stock under different conditions, if that task would last all day. The acceptable hand force depends on the direction of force (pushing or pulling), the frequency (number of activities per day), the distance per activity, the hand height and the gender of the employees. There is also a distinction between the initial force (to start the motion) and the sustained force required (to maintain the motion of the load). To use these data in the DUTCH we made the following calculations and choices:

- based on the normal-distributed percentile values per push-pull situation, curves were assessed for maximal acceptable hand forces, as a percentage of the working population (see Fig. 1);
- in addition to separate values for men and women, limit values were calculated for an equal distribution of men and women (dotted line in Fig. 1);

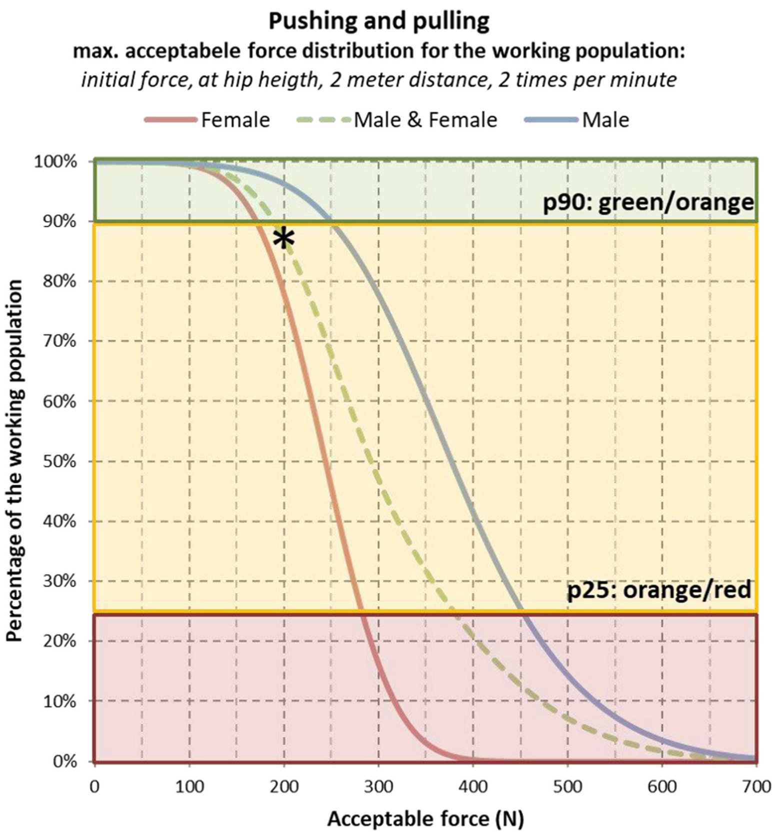

Fig. 1. Maximal acceptable hand forces for different percentages of the working population, for a specific push/pull condition (initial force, hip height, 2 m, 2 times/min). The green dotted line shows values for a population of 50% men/50% women. The limits between the coloured planes show limits for p90 (green/orange) and p25 (orange/red). * For example: a force of 200 N is acceptable for 88% of the workforce (men and women), when pushing or pulling carts at waist level twice per minute. (Color figure online)

- a traffic light model was added (coloured planes in Fig. 1); The orange/red border represents the point where 90% of the workforce is protected, which is common in ergonomic guidelines and instruments. With the orange/red border 25% is protected, which is a boundary based on consensus by the expert group;
- because limit values for the initial motion of a load are lower than those for the sustained motion of the same load, limit values for the initial motion are used;
- because in practice there is almost always a combination of pushing and pulling, the average acceptable hand forces for pushing and pulling are used.

3.3 Convert Hand Force to Cart Weight

Because force measurements are difficult to perform in practice and often lead to errors, maximum hand forces from the Snook tables were converted to cart weights. For this purpose a formula was distracted from measurements of Hoozemans et al. [12] at 3 cart weights (85, 135 and 320 kg), 2 hand heights (hip and shoulder height), push and pull with 1 and 2 hand(s) and initial and sustained hand forces. The number of measurements and the different push and pull conditions in this study were limited. Therefore, the formula must be validated in future. To get an impression of the 'face-validity' of the formula, hand forces from the Snook table were converted to cart weights using the formula. These cart weights seemed to be realistic.

3.4 Influence of the Surface and Material (Qualitative Part of the Method)

The force required for pushing and pulling a cart also depends on factors such as the type of surface, wheel diameter, material of the wheels and state of maintenance. Because these factors are often difficult to determine for the user and because of insufficient scientific support for the impact of these features on the hand forces, it was decided to process them in a 'qualitative' way and not in a 'quantitative' way. This means that we indicate whether these characteristics are *favourable* or *unfavourable* to the required strength, without calculating the effect on an acceptable cart weight.

4 The Result: DUTCH - Description of the Tool

4.1 Structure of the Tool

Figure 2 presents an overview of (1) required input for using the tool (2) calculations that are being made using the input data, and (3) the results that are being presented to the user. The figure also shows the difference between the quantitative (upper part in blue) and qualitative part (lower part in green) of the method.

4.2 Results of an Assessment

Based on the quantitative input data, the DUTCH calculates the average and maximum cart weights (cart and load together) in a specific situation. For the evaluation of this score the traffic light model is used: 'green' means 'low work demands, with minimal

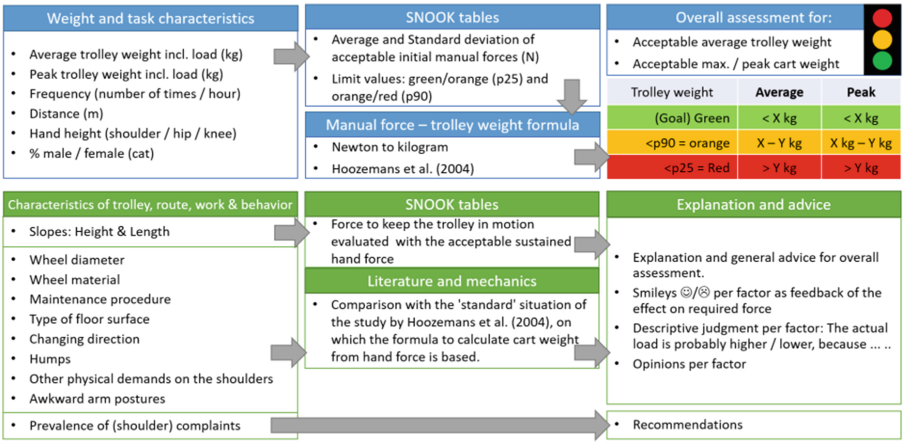

Fig. 2. Input, calculations and results of the DUTCH for the quantitative and qualitative assessment of pushing and pulling (Color figure online)

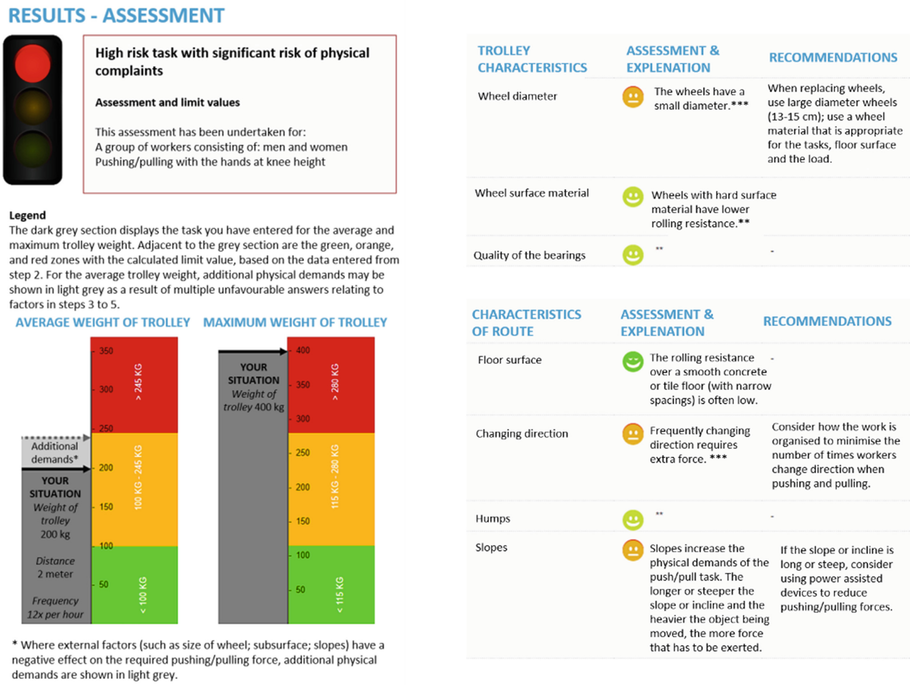

Fig. 3. Example of the quantitative assessment (left) and qualitative assessment (right) (Color figure online)

risk of physical symptoms; 'orange' means 'demanding task: there is a risk of physical symptoms and 'red' means 'high work symptoms, with high risk of physical symptoms' (Fig. 3, left side).

In addition, information is presented on the effect of the characteristics of the environment, material and behaviour. Emoticons and explanatory text show for each factor if they have a favourable or unfavourable effect on the evaluation. This is the result of the qualitative assessment (Fig. 3, right side). For example, a smooth and firm surface has a positive effect on the assessment (☺) because pushing a cart or trolley on a smooth firm surface requires less force than if the surface is rough (☹).

In case of an orange or red assessment, the DUTCH presents measures to reduce the physical work demands and it refers to a 5 steps risk reduction approach.

5 Conclusions

The DUTCH is an evidence and expert based, simple webtool for a quick evaluation of push and pull activities at work, providing insight into the presence of risk factors and potential measures to reduce the work demands. The English version of the tool is available at https://www.fysiekebelasting.tno.nl/en/ and is targeted at occupational health officers. The DUTCH has been tested on a small scale for usability by companies and experts. The reliability and validity of the evaluation method have not been studied extensively yet.

References

1. Dutch Health Council (Gezondheidsraad) (2012) Kracht zetten, duwen en trekken in werksituaties. Gezond-heidsraad, Den Haag. http://gr.nl/nl/adviezen/gezonde-arbeidsomstandigheden/kracht-zetten-duwen-en-trekken-werksituaties. (in Dutch)
2. Hoozemans MJ, Knelange EB, Frings-Dresen MH, Veeger HE, Kuijer PP (2014) Are pushing and pulling work-related risk factors for upper extremity symptoms? A systematic review of observational studies. Occup Environ Med 71(11):788–795
3. Mital A, Nicholson AS, Ayoub MMA (1997) Guide to manual materials handling. Taylor & Francis, London
4. Jürgens WW, Mohr D, Pangert R, Pernack E, Schultz K, Steinberg U (2002) Handlungsanleitung zur Beurteilung der Arbeitsbedingungen beim Ziehen und Schieben von Lasten. LASI Ver-öffentlichung LV29. Hrsg. Länderausschuss für Arbeitsschutz und Sicherheitstechnik
5. Steinberg U, Caffier G, Liebers F (2006) Assessment of manual material handling based on key indicators: German guidelines. In: Karwowski W (ed) Handbook on standards and guidelines in ergonomics and human factors, Chap 18. Lawrence Erlbaum Associates, Mahwah, pp 317–335
6. Hoozemans MJM, Visser B, Van Dieën JH (2010) Evaluation of pushing and pulling at the work-place using an web-based PushPullCalculator. In: Seventh international scientific conference on prevention of work-related musculoskeletal disorders PREMUS 2010, Angers, France

7. Kuijer PPFM, Hoozemans MJM, Frings-Dresen MHWA (2007) A different approach for the ergonomic evaluation of pushing and pulling in practice. Int J Ind Ergon 37:855–862
8. HSL (2003) Pulling and pushing operations risk assessment tool, draft, 24 June 2013
9. Hoozemans MJ, van der Beek AJ, Frings-Dresen MH, van Dijk FJ, van der Woude LH (1998) Pushing and pulling in relation to musculoskeletal disorders: a review of risk factors. Ergonomics 41(6):757–781 Review
10. Weston EB, Aurand A, Dufour JS, Knapik GG, Marras WS (2018) Biomechanically determined hand force limits protecting the low back during occupational pushing and pulling tasks. Ergonomics 61(6):853–865. https://doi.org/10.1080/00140139.2017.1417643
11. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33(1):159–174
12. Hoozemans MJM, Kuijer PPFM, Kingma I, Van Dieën JH, De Vries WHK, Van der Woude LHV, Veeger HEJ, Van der Beek AJ, Frings-Dresen MHW (2004) Mechanical loading of the low back and shoulders during pushing and pulling activities. Ergonomics 47 (1):1–18
13. Jäger M (2001) Belastung und Belastbarkeit der Lendenwirbelsäule im Berufsalltag. Ein interdisziplinärer Ansatz für eine ergonomische Arbeitsgestaltung. Fortschr.-Ber. VDI Reihe 17 Nr. 208. VDI Verlag, Düsseldorf
14. Chaffin DB, Andersson GBJ, Martin BJ (1999) Occupational biomechanics, 3rd edn. Wiley, New York
15. Snook SH, Ciriello VM (1991) The design of manual handling tasks: revised tables of maximum acceptable weights and forces. Ergonomics 34:9