An international validation study of the care thermometer: a tool supporting the quality of ergonomic policies in health care

Knibbe, J.J.[*], N.E.Knibbe
LOCOmotion, Research in Health Care, Brinkerpad 29, 6721 WJ Bennekom, The Netherlands.

Abstract. Occupational back pain among nurses leads to high costs and personal suffering for nurses. It is difficult to assess the success of such initiatives and to monitor results in a practical way. Such a practical monitoring and web-based instrument was developed. This Care Thermometer (CT) allows the users to assess the current situation in their facility today, and, with regular use, it can help to track progress over time. The Care Thermometer is a further step in the development of the TilThermometer, a validated assessment tool that is used on a large, national scale in The Netherlands. The claims of the newly developed Care Thermometer are ambitious and an international validation study was performed in four countries: the UK, USA, Germany and the Netherlands. The instrument appears to be sufficiently valid, useful and practical. There are however some points to keep in mind when interpreting the results of the CT. Especially a careful, punctual and stringent data collection phase is crucial for accurate and useful results. Some recommendations to further improve the practical use both for the design of the instrument and the process of data-collection and -entering are given.

Keywords: ergonomics, nursing, occupational back pain, prevention

1. Introduction

Occupational back pain among nurses still leads to high costs for health care facilities and personal suffering for nurses. Across the world preventive attempts are undertaken. It remains difficult to assess the success of such initiatives and to monitor results in a practical way without too much interference with daily routines.

Such a practical monitoring and web-based instrument was developed. This Care Thermometer (CT) allows the users to assess the current situation in their facility today, and, with regular use, it can help to track progress over time. The Care Thermometer is a further step in the development of the TilThermometer, a validated assessment tool [3,4]. This TilThermometer (in English Lift-Thermometer) was and is widely used on a national basis in the Netherlands. Currently about 80% of all clients in nursing homes and homes for the elderly have been assessed with the TilThermometer. National monitoring with the use of the results of the aggregated TilThermometer-data across the country show a steady improvement in the quality of preventive policies and a reduction of the back pain prevalence and sick leave in health care [4].

2. Method for validating the Care Thermometer

The claims of the new Care Thermometer are ambitious and therefore an international validation study was undertaken. The results of three different data sources ('triangulation') were compared and their degree of convergence was studied. One of the three was the Care Thermometer. For this purpose the results of the Care Thermometer were compared to the results of two other frequently used tools.

As the CT will, unlike the original tool of the TilThermometer, be used on an international level, the validation study would have to cover that scope as well. Therefore the study took place in four coun-

---

[*]Corresponding author. E-mail: j.j.knibbe@planet.nl .

tries (UK, USA, Germany and The Netherlands). From each country 4-5 facilities participated: two from a long stay, nursing home like setting and two from (a more) acute care setting.

The results of the following three data sources ('triangulation') were combined.

1. the results of the Care Thermometer (CT)

2. the results of the StaDyMeter (SDM): an activity log based on self-registration

3. the results of the RiskRadar (RR): a more general exposure assessment (not self-administered).

3. Results

A total of 17 facilities participated in de study and they were evenly spread among the four countries, half a more acute care setting and the other half a more residential care type setting. A total number of 301 nurses participated. The response rate ranged from 62%- 100%. The nurses registered 4919 activities on their activity logs (StaDyMeter): averaging out towards a response rate of 82%. Almost the same number and same group of nurses (295) replied by filling in a RiskRadar on their activities. This pointed to an overall response rate of 76%. And finally a total number of 1808 patients were assessed with the CT. This was a response rate of 96%. This data was collected from the same wards as the nurses worked on.

4. Conclusions

4.1. Are the parameters of the CT complete?

In all participating countries the parameters of the CT appeared to cover most, but not all, sources of potential physical overload. Although the sources not present in the CT do differ from country to country and from facility to facility the sources mostly missed, or not sufficiently covered across all countries were:

   a.   work load in general
   b.   pushing, pulling and maneuvering
   c.   long walking distances
   d.   prolonged standing

Although we cannot give a representative indication of the differences between countries it may be interesting to see that in Germany and the Netherlands pushing and pulling is considered a problem not sufficiently covered by the CT, whereas in the USA the general workload is often reported as a problem

and in the UK the long working hours, walking distances and 'being on your feet for hours on end' are mentioned as major problems.

However: the elements that are present in the CT were supported by all countries and, as a group, they do represent the major sources of physical overload as they were intended to do. This is an important conclusion for our study.

The conclusion is that the CT does cover the major, but not all sources of physical overload. The CT should therefore never be used as the sole check for work safety measures or to assess individual problems reported by individual nurses.

4.2. Do these parameters actually measure what they intend to measure?

They apparently do. They identify rather accurately the relative magnitude of each of the sources, but the extend with which overload occurs (absolute size) may differ from instrument to instrument. High seems to be consistently high, but the absolute degree varies. The StaDyMeter seems to structurally report a somewhat higher level of exposure than the CT does. In figure 1 an overview is given when it comes to classifying the care load in the 17 facilities. The higher levels (top line) are the aggregated results of the StaDyMeters from the nurses working on the same wards that the CT scores come from (lower line). The white line depicts the RiskRadar results that seem to correlate most with the scores from the StaDyMeter.

This effects seems to be rather constant across nations, so we do not expect this to be a problem of the CT, but more a result of the fact that subjective registrations may structurally indicate a higher level of exposure than the CT does, as has been reported before in other studies [3,4].

4.3. Does the CT produce valid and reliable results in real life conditions?

The CT uses 5 classes of patient dependency or mobility to identify the load resulting from patients. It appears that there was indeed some misclassification here, when the results of the CT were compared to the results of the StaDyMeter and also the observations performed. Nurses were sometimes too fast in their classificationprocess. The process is very visually oriented and they appeared to rely more on the pictures than on the actual criteria for the classifica-

tion. It appeared that the original instrument (the TilThermometer) is less sensitive to this issue as it uses a 3-class system: easier to use and to remember. This aspect is now recognized and a more clear and less visual differentiation has been developed and integrated in the software.

4.4. Are the parameters and the tool itself sensitive enough to highlight specific differences across health care sectors and across countries?

We can clearly state that this is indeed the case. We can see considerable differences between the facilities in the situation they are in, their risk levels, their patient population and the preventive precautions they have taken to promote safe working conditions and high quality of care. As can be seen in figure 1 the results differ widely between this group of 17 facilities. We could also not find a clustering within countries. Furthermore these differences are not only present, but also relevant for designing a preventive policy.

Nevertheless we do have to keep in mind that the more specific the results are analyzed (and therefore the smaller the subgroups will get) the less power the CT, or any other instrument for that matter, has to assess relevant differences in a valid and reliable way. We underline our earlier conclusion that the CT is not a stand alone tool and the conclusions should always be complemented with information from other sources and sound reasoning and observations.

5. Discussion

This leads us to the final conclusion that the Care Thermometer is indeed sufficiently valid for its purpose of policy and program evaluation, monitoring and tailoring. The data collection itself must be done under good conditions (good instruction, stable environment, no or not too much time pressure and preferably consensus on the data entered in the CT by at least three nurses deciding on the input together). Slight adaptations and reinforcements in the data collection protocol will improve the validity and reliability of the results by improving the sensitive first phase of data collection.

Also the results must not be specified into too much detail especially if the number of clients involved is low (f.e. smaller than 30). For larger scale and/or research purposes larger numbers are required to give the analyses more statistical power.

The results are more representative and valid in a more stable (f.e. long term care like) environment and less stable in a more acute environment (like an emergency setting or a post-operative unit in a hospital). This is not due to the tool itself, but to the very nature of the rapidly changing patient population and the resulting changes in the functional status of that population. We recommend that under those circumstances a more frequent assessment is made: more 'snapshots' will be required. This will increase the validity and reliability of the results by averaging out the fluctuations.

All in all the results show that the CT can, on the basis of this study, be considered to be a relevant, useful and valid tool, especially when the recommendations are followed and the procedures for data gathering are reinforced.
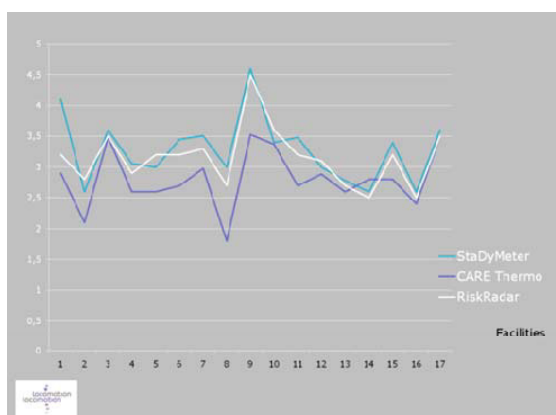

Figure 1: The aggregated results of the 17 facilities (x-axis) for the three instruments. All results were transformed into the same scale ranging from 0 (zero) to 5 (maximum load).

References

[1] Hignett, S., Crumpton, E., Ruszala, S., Alexander, P., Fray, M., Fletcher, B. (2003). Evidence based patient handling: systematic review. Nursing Standard, 17 (33), 33-36.

[2] Jansen, J., Morgenstern H., Burdorf, A. (2004). Dose-response relations between occupational exposures to physical and psychosocial factors and the risk of low back pain. Occupational Environmental Medicine, 61(12): 972–979.

[3] Knibbe, J.J., Friele, R.D. (1999). The use of logs to assess exposure to manual handling of patients, illustrated in an intervention study in home care nursing. International Journal of Industrial Ergonomics 4 (24), 445-454.

[4] Knibbe, J.J., Knibbe, N.E., Geuze, L. (2008). Een hap uit een gegroeide olifant, vierde nationale monitoring fysieke belasting, SOV&V, Den Haag (in Dutch). Fourth National Monitoring, Unions & Employers organizations, The Hague.