Response to Letter to the Editor

Comparison of clinical outcomes in diabetic and non-diabetic burns patients in a national burns referral center in Southeast Asia: A 3-year retrospective review; Methodological issues

Dear Editor-in-Chief, Dr. Wolf

We gladly welcome the discussion on the methodological issues brought about by Reza Pakzad and colleagues.

Allow us to point out that the main objective of our study was to compare the complications between DM and non-DM burns patients and compare the patients characteristics in these 2 groups, which was straightforward. The methodological issues related to logistic regression modelling mentioned in their letter were encountered to some extent in our study. These issues are all related to the challenge of sample size and insufficient power to detect all the potential risk factors for the complications studied. These issues will be addressed in this letter.

We used all the data that were available and up-to-date in our Burns registry and extracted additional variables needed for the study from our hospital medical records. The number of complications was quite rare, for instance, 29 (4.9%) patients in our dataset had unplanned readmission. According to the rule of thumb for sample size for carrying out a logistic regression of unplanned readmission, we can have up to 3 factors in the model, one of the factor of interest being DM. Given our low rate of complications for all the 3 types of complications, we contemplated presenting only results of univariate analyses of DM with complications. However, we felt that it was justified attempting to adjust for potential confounders in a multivariate regression of complication on DM. Despite the potential for inadequate power to detect significant factors, we felt that factors with large effect would be detected. In addition, our study is considered a descriptive study, mostly exploratory in nature, not so much as confirmatory.

We initially started with multivariate analysis with stepdown variable selection method but the final model building strategy adopted was user-controlled. Factors considered for entry in the model were those found significant at $p<0.05$ in the univariate analysis as well as factors known to be clinically important from domain knowledge (identified with input from the clinicians in the study). Taking into account the rule of thumb of 10 events for each predictor for sample size requirement, logistic regression presents other challenges as well. We did encounter statistical issues such as quasi-separation (near perfect prediction) and collinearity in the initial multivariate modelling with all potential factors. These issues were addressed by removing some of the factors that were correlated with each other. Whenever possible, we attempted exact logistic regression as well. In summary, the final model building for our multivariate models was manual selection of variables that gave the best fit, using statistical techniques such as adding factors that brought substantial resulting changes to the regression coefficients. The models were built in discussion with the team, with identified clinical domain knowledge in mind.

We acknowledge that with our present sample size, only factors with large impact can be detected and have their parameters estimated with precision. We intend to validate our results when our revamped Burns registry has captured additional pertinent parameters and become more mature. Only then, will the research studies that need very large sample sizes be more effective and less labour-intensive.

Zhao-Kai Low
Yong Loo Lin School of Medicine, National University of Singapore, NUHS Tower Block, Level 11, 1E Kent Ridge Road, Singapore 119228, Singapore

Wai-Yee Ng
Division of Research, Singapore General Hospital, 226 Outram Road Blk A Level 2, Singapore 169039, Singapore

Stephanie Fook-Chong[a,b]
[a]*Division of Research, Singapore General Hospital, 226 Outram Road Blk A Level 2, Singapore 169039, Singapore*

[b]*Centre for Quantitative Medicine, Duke-NUS Graduate Medical School, Singapore*

Bien-Keem Tan
Si-Jack Chong
Jolie Hwee
Department of Plastic, Reconstructive & Aesthetic Surgery, Singapore General Hospital, Outram Road, Singapore 169608, Singapore

Sook-Muay Tay*
Department of Anaesthesiology, Singapore General Hospital, Outram Road, Singapore 169608, Singapore

* *Corresponding author*. Fax: +65 6226 1492.
E-mail address: tay.sook.muay@sgh.com.sg (S. Tay).

http://dx.doi.org/10.1016/j.burns.2017.03.019

Letter to the Editor

Pain behavior observation scales in young children with burns

Dear Editor,

With great interest we recently read the paper reporting on the accuracy of the Face, Legs, Activity, Cry, Consolability Scale (FLACC), a structured behavior observation scale, as a potential instrument for the measurement of pain in children with burns [1]. We welcome this growing attention for pain

DOI of original article: http://dx.doi.org/10.1016/j.burns.2017.02.020
DOI of original article: http://dx.doi.org/10.1016/j.burns.2017.04.010

Table 1 – Linear regression analyses estimating the effect on nurses' estimation on the total pain score for Comfort-B, POCIS and VAS obs respectively for background and procedural pain.

	Nurse perception	Comfort-B β	POCIS β	VAS obs β
Background pain (n=2532 observations)	Pain	0.73	0.05	0.00
	Anxiety	0.00	0.00	0.01
	R square	13%	32%	23%
Procedural pain (n=1299 observations)	Pain	0.38	0.34	0.00
	Anxiety	0.00	0.00	0.12
	R square	28%	28%	27%

measurement in children admitted to a burn center. Particularly in young children, unable to provide self-reports, pain should be assessed by structured behavior observation [2–4]. This requires rigorously developed and tested instruments to obtain data of the highest possible quality. This high quality can be obtained by using scales with sufficient psychometric properties, which is a prerequisite for research into pain interventions and wound treatment. Instruments with reliable cut-off points such as the Comfort-B [5] have surplus value as it assists adequate pain management. In older children (5–18 years old), usually able to provide self-reports, self-reports should be the first choice of pain measurement.

The study of Shen et al. [1] basically has several strong points such as the assessment of different pain levels and investigating whether personal nursing characteristics affect the pain rating. This is interesting information. However, we would like to add some comments that may enable the readers to interpret the results from another perspective. First, the authors concluded that pain intensity affected the accuracy of the pain ratings when using the FLACC. To investigate accuracy of pain measurement, a pain score established by an expert panel including three persons, was used as the gold standard. Accuracy was defined as not deviating statistically significantly from this panel score when rating four video fragments with different levels of pain. Two videos judged to reflect different pain levels, were rated highly similar by the 24 nurses involved in the study. It was concluded that nurses using the FLACC could not distinguish between these pain levels. However, it was not established to what extent the group of 24 nurses reached mutual agreement, e.g., expressed by an intraclass correlation coefficient as recommended by Mokkink et al. [6,7] or whether there may be a learning curve when using the scale more frequently. If there is high agreement among the nurses, this could also indicate the lack of discernment among the video fragments, or it may question the panel's pain score, rather than concluding it was the nurses' inability to distinguish pain levels or being an artifact of the scale. Second, the authors concluded that accuracy of the pain rating was influenced by nurses' clinical experience and thus can be regarded as a biased pain rating when using the FLACC. The reference group was formed by the nurses having 0–5 years of clinical experience (n=12). It was concluded that their score showed higher accuracy (i.e., similar to the panel score) as compared to the nurses having 11–15 years experience (n=3), 21–25 years experience (n=3) and >25 year experience (n=2). When a correction for multiple testing was applied, only the highest experienced group would reach significance. However, when taking into account the extremely small subsample size (n=2), this conclusion is questionable. Does the effect hold when comparing two groups (e.g., 0-5 versus >5 years)?

The study of Shen et al. triggered us to conduct secondary analyses in our study to better understand the possible influence of biasing factors. Our study investigated the reliability and validity of two behavioral observation instruments and the visual analogue scale used as an observation measure (VAS obs) in children with burns [8–10]. The Comfort-B and POCIS, two instruments that are comparable to the FLACC, were found valid in the burns population whereas the VAS obs was not. Unfortunately, we did not examine the influence of personal characteristics of the nurses, but the nurses in our study were asked to report their perception of what caused the child's behavior: pain, anxiety, anger, etc. A regression analysis (using the total scale score as the dependent variable and the nurses' perceptions pain and anxiety as the predictors) revealed that both behavioral instruments were not influenced by the nurses' perception that the child was in pain, whereas the VAS obs did show to be influenced by the perception that the child was in pain (Table 1). Importantly, as reported earlier [8,9], the paired observations (two nurses observed the same child), did not reach sufficient agreement when using the VAS obs but they did when using the POCIS or Comfort-B. This suggests the behavioral observation scales were more resistant to a possible bias caused by the nurses' estimation of the child's pain. This suggests that the two behavioral observation scales minimize the effect of external factors.

In conclusion, a pain scale should preferably be resistant to external factors that may bias the observation. Behavioral observation scales have been shown to meet this criterion but they are not perfect yet. Therefore, a study into possible biasing factors such as Shen et al. did, remains relevant. Notwithstanding the high relevance, we trust our comments and additional analyses enable your readers to value the usefulness of observational behavioral pain measures in burns centers and to put into perspective the interpretation of the biasing factors. More research may be required.

Conflict of interest

The authors declare to have no conflict of interest.

REFERENCES

[1] Shen J, Giles SA, Kurtovic K, Fabia R, Besner GE, Wheeler KK, et al. Evaluation of nurse accuracy in rating procedural pain among pediatric burn patients using the Face, Legs, Activity, Cry, Consolability (FLACC) Scale. Burns 2017;43(1):114–20.
[2] Blount RL, Loiselle KA. Behavioural assessment of pediatric pain. Pain Res Manag 2009;14(1):47–52.
[3] Henry D, Foster R. Burn pain management in children. Pediatr Clin North Am 2000;47(3):681–98.
[4] Martin-Herz S, Patterson D, Honari S, Gibbons J, Gibran N, Heimbach D. Pediatric pain control practices of North American burn centers. J Burn Care Rehabil 2003;24(1):26–36.
[5] De Jong AEE, Bremer M, Van Komen R, Vanbrabant L, Schuurmans MJ, Middelkoop E, et al. Pain in young children with burns: Extent, course and influencing factors. Burns 2014;40(1):38–47.
[6] Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. BMC Med Res Methodol 2010;18:10–22.
[7] Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010;63(7):737–45.
[8] De Jong AEE, Bremer M, Schouten M, Tuinebreijer WE, Faber AW. Reliability and validity of the pain observation scale for young children and the visual analogue scale in children with burns. Burns 2005;31(2):198–204.
[9] De Jong AEE, Baartmans M, Bremer M, van Komen R, Middelkoop E, Tuinebreijer W, et al. Reliability, validity and clinical utility of three types of pain behavioural observation scales for young children with burns aged 0 to 5 years. Pain 2010;150(3):561–7.
[10] De Jong AEE, Tuinebreijer WE, Bremer M, van Komen R, Middelkoop E, van Loey N. Construct validity of two pain behaviour observation measurement instruments for young children with burns by Rasch analysis. Pain 2012;153(11):2260–6.

A.E.E. de Jong[a,b,*]
[a]Burn Centre, Red Cross Hospital, Beverwijk, the Netherlands

[b]Association of Dutch Burn Centres, Beverwijk, the Netherlands

N.E.E. van Loey[a,b]
[a]Association of Dutch Burn Centres, Beverwijk, the Netherlands

[b]Department of Clinical Psychology, Utrecht University, Utrecht, the Netherlands

* Corresponding author at: Burn Centre, Red Cross Hospital, Vondellaan 13, 1942 LE Beverwijk, the Netherlands. Fax: +31 251264948.
E-mail address: aeedejong@rkz.nl (A.E.E. de Jong).

http://dx.doi.org/10.1016/j.burns.2017.03.032

Response to Letter to the Editor

Evaluation of nurse accuracy in rating procedural pain among pediatric burn patients using the Face, Legs, Activity, Cry, Consolability (FLACC) Scale

Dear Editor,

Thank you for providing us the opportunity to respond to the Letter to Editor regarding our paper on nurses' rating accuracy using the FLACC scale [1]. Accurate pain assessment is critical in pediatric burn care practice, and we greatly appreciate the letter authors' interest and time in evaluating our study findings and initiating this discussion.

To respond to authors' concerns regarding our data analysis, we have conducted additional analyses. In response to authors' suggestion to report the mutual agreement among the 24 nurses, we went back to our raw data and conducted intraclass correlation analyses. These results showed that there was relatively low mutual agreement among nurses for each of the four videos in our study. Specifically, the intraclass correlation coefficients were 0.65, 0.50, 0.74, and 0.7 for the "mild", "moderate", "high", and "severe" videos, respectively among all nurses (N=24). Based on the fixed effects estimate for the variable "Round" (−0.08) reported in Table 3 of our original paper, there did seem to be a learning curve, where nurses tended to score more accurately (narrower distance from expert scores) as they rated from Round 1 through Round 3. However, this learning effect was not significant in our model estimation (p=0.71).

The authors (de Jong and van Loey) also raised a great point that some groups in our variable "Years of nursing experience" had small sample sizes. Therefore, we combined all the groups that reported more than five years of experience into a single group (n=12) and compared it to the rating accuracy of nurses who had less than five years' experience (n=12), as suggested by the authors. We then re-conducted the repeated-measures ANOVA and found that our conclusion held. Nurses who had more than five years of nursing experience had significantly less rating accuracy than those with less than five years of experience (B=−3.37, p=0.02). As discussed in our paper, more research is needed to explore the possible mechanism for this seemingly counterintuitive finding.

Finally, the authors' secondary data analyses comparing Comfort-B, POCIS, and VAS obs were very interesting, and we thank authors for these and their important insights. We agree with the authors that behavioral-based measurement of pediatric pain such as FLACC, Comfort-B, and POCIS would be less influenced by raters' perceived level of pain than VAS obs. Part of the reason might be that VAS obs by design measures a child's pain based on the rater's perception and is thus more likely (as it should be) influenced by subjective factors within individual raters. In terms of measurement reliability and 'objectivity', we fully agree with the authors that observational behavioral pain assessment tools are a better choice, and indeed more research is needed in pediatric burn

DOI of original article: http://dx.doi.org/10.1016/j.burns.2017.03.032