

Raising suspicion of maltreatment from burns: Derivation and validation of the BuRN-Tool

Alison M Kemp¹, MB BCh, MRCP, FRCPCH, Linda Hollén², BSc MSc PhD, Alan M Emond², MA MB MD FRCP FRCPCH, Diane Nuttall¹, RGN, Dip Res, David Rea³, RN BSc, Sabine Maguire¹, MB BCh, MRCPI, FRCPCH

Affiliations: The research team are all members of the Scar Free Foundation Centre for Children's Burns Research: ¹ School of Medicine, Cardiff University; ² School of Social and Community Medicine, University of Bristol; ³ University of the West of England, Bristol.

Corresponding Author: Alison M Kemp, (School of Medicine, Cardiff University, University Hospital of Wales, Cardiff CF14 4YS, kempam@cardiff.ac.uk)

Short Title: BuRN-Tool: Identify Maltreatment in Pediatric Burns

Funding Source: This study was supported by the Scar Free Foundation, and the National Institute for Social Care and Health Research (NISCHR). The Children's Burns Research Centre is part of the Burns Collective, a Scar Free Foundation initiative with additional funding from Vocational Training Charitable Trust and Health and Care Research Wales.

Financial Disclosure: The authors have indicated they have no financial relationships relevant to this article to disclose.

Conflict of Interest: The authors have no potential conflicts of interest to disclose.

Abbreviations:

AUC – Area under the curve

CPT – Clinical Prediction Tool

ED – Emergency Department

LR – Likelihood Ratios

LR+ – Positive Likelihood Ratio

LR- – Negative Likelihood Ratio

NICE – National Institute for Health and Care Excellence

ROC – Receiver operating curve

Table of Contents Summary

The BuRN-Tool: a novel clinical prediction tool for the emergency Room to identify cases of maltreatment in children who have sustained a burn.

What's Known on This Subject: A proportion of children with medically attended burns will have sustained their injuries from child neglect or physical abuse. These children are assessed by clinicians with varying pediatric experience and underlying maltreatment may go unrecognized.

What This Study Adds: A Clinical Prediction Tool, derived from research evidence and primary epidemiologic data, and validated prospectively on a novel dataset, has the potential to raise suspicion of maltreatment associated with pediatric burns and be an adjunct to clinical decision-making.

Contributors' Statement:

Alison Kemp and Sabine Maguire conceptualized, designed, managed and supervised the derivation and validation phase of the study, designed the data collection instruments, and coordinated and supervised data collection at the validation sites, wrote the study manuscript final versions.

Linda Hollén completed the statistical analyses, drafted the initial manuscript and coordinated the final version.

Alan Emond directed the research network, chaired and advised research team meetings, coordinated the design of database, gave editorial advice and contributed to writing the final manuscript.

Diane Nuttall and David Rea designed the data collection instruments, and coordinated and supervised data collection at the validation sites, coordinated the data collection, entered data onto the database at the UK validation sites.

All authors reviewed, revised and approved the final manuscript as submitted.

ABSTRACT

Background 10-25% of childhood burns arise from maltreatment. Aim: to derive and validate a clinical prediction tool to assist the recognition of suspected maltreatment.

Methods Prospectively collected data from 1327 children with burns were analyzed using logistic regression. Regression coefficients for variables associated with 'referral for child maltreatment investigation' (112 cases) in multivariable analyses were converted to integers to derive the BuRN-Tool, scoring each child on a continuous scale. A cut-off score for referral was established from receiver operating curve analysis and optimal sensitivity and specificity values. We validated the BuRN-Tool on 787 prospectively collected novel cases.

Results Variables associated with referral were: age <5 years, known to social care, concerning explanation, full thickness burn, uncommon body location, bilateral pattern and supervision concern. We established 3 as cut-off score, resulting in a sensitivity and specificity for scalds of 87.5% (95% CI:61.7-98.4) and 81.5% (95 CI:77.1-85.4) respectively and for non-scalds sensitivity was 82.4% (95%CI:65.5-93.2) and specificity 78.7% (95% CI:73.9-82.9) when applied to validation data. Area under the curve was 0.87 (95% CI:0.83-0.90) for scalds and 0.85 (95% CI:0.81-0.88) for non-scalds.

Conclusion The BuRN-Tool is a potential adjunct to clinical decision-making, predicting which children warrant investigation for child maltreatment. The score is simple and easy to complete in an emergency department setting.

INTRODUCTION

Burns are common pediatric injuries presenting to the Emergency Department (ED) with an estimated annual attendance of 50,000 children in the UK[1] and 500,000 in the USA.[2] While the majority are unintentional injuries, international estimates suggest 10-25% are the result of physical abuse or neglect.[3-7] In order to protect these children from future harm, early identification is essential.

Maltreatment is under-recognized in busy EDs,[8] due to high staff turnover, high volume of patients, pressure to manage and discharge patients and variable pediatric experience. Whilst child protection training is mandated in the UK for all staff in contact with children,[9] there is inconsistent uptake in ED where staff of different grades have different levels of training. The ED work force is often poorly equipped to identify children at risk of maltreatment.

Clinical Prediction Tools (CPTs) are commonly used in EDs to aid decision-making, and improve diagnostic accuracy.[10-13] These evidence-based tools use a combination of predictor variables from the history and clinical examination to determine the probability of a specified outcome. A number of generic screening tools for physical abuse or neglect are used in EDs, but systematic reviews show poor performance,[8, 14] with the exception of one recent study that suggests that the modified SPUTOVAMO, a screening tool used in the Netherlands, could identify children who were admitted to one Burns Centre who needed additional support for suspected abuse or neglect.[3] A small retrospective case-note review attempted to improve the identification of 'burn abuse' by introducing a 'checklist'[15] that resulted in a three-fold increase in the number of children being referred for assessment by social care. However, the 'checklist' was not validated.

This study aims to derive and externally validate a burns-specific CPT to aid the assessment of children who present to hospital and to identify cases where maltreatment (neglect or physical abuse) is suspected and a referral to hospital or Local Authority children's social care teams for maltreatment investigation is recommended.

METHODS

Derivation

To derive the CPT we used data from a prospective multicenter study of children presenting with a burn to two pediatric EDs, three general EDs and three burns units in the UK and Ireland during 2008-2010.[16] Children less than 16 years old were included; victims of household fires were excluded. A standardized data collection proforma, the Burns and Scalds Assessment Template (Supplementary Figure 1), was completed by the treating clinician, recording: age, gender, gross motor developmental stage, previous social care involvement, level of adult supervision, time to presentation for medical attention, the type of burn, its severity, bilateral and multiple burns, first aid, anatomical site, causal agent and mechanism of injury. The anatomical distribution of the injury was drawn on a body map (categorized into 12 anatomical sites; supplementary Table 1). These items were derived from the results of a systematic review of burns in maltreatment.[17]

Outcome and Predictor Variables

As outcome, children were classified into two groups: referred to the hospital or Local Authority children's social care team for investigation of suspected maltreatment or not. To derive the CPT, we analyzed ten potential predictor variables detailed in Table 1. The injury types were grouped into scalds (from hot liquids, food or steam) and non-scalds (contact,

chemical, electrical, radiation, and friction burns) and analyzed separately due to differences in causal agents, mechanisms, extent and pattern of injuries sustained.[16]

INSERT TABLE 1 HERE

Statistical Analyses

Analyses were performed in Stata v.14.[18] Descriptive statistics are reported using proportions, medians and interquartile range (IQR). Associations between the potential predictor variables (Table 1) and the outcome were conducted using univariate and multivariable logistic regressions. Odds ratios (ORs) and 95% confidence intervals (CIs) are reported. The Pearson goodness-of-fit test was used to ensure model fit. The analyses were repeated with bootstrapping using 10,000 replicates to internally validate the accuracy of estimates. We fitted receiver operating characteristic curves (ROC) and corresponding area under the curves (AUC) using the *roctab* command which performs nonparametric ROC analysis.

To create the CPT, we constructed a point scoring system by converting the regression coefficients of the independently predictive variables in the multivariable logistic models into integer values.[19, 20] Coefficients were rounded to the nearest integer. This enabled each case to be assigned a total score by summing the integer values for all variables. Cases were scored blinded to the outcome. We used ROC curve analysis for the full range of possible scores against the outcome to assign the best cut-off score based on the lowest score at or above which the sensitivity and specificity were optimal to maximize discovery of concerning cases while reducing unnecessary referrals to social care.[21-23] Using the *diagt* command, we extracted additional measures of diagnostic accuracy namely likelihood ratios, positive predictive values (PPV) and negative predictive values (NPV) for different

scores.[23] The resulting CPT is called the BuRN-Tool (**B**urns **R**isk assessment for **N**eglect or abuse **T**ool).

External Validation of BuRN-Tool

To externally validate the BuRN-Tool, a second study conducted during June 2013-April 2014, collected the same data using a refined version (supplementary Figure 1) of the original data collection proforma (BaSAT). Data were collected at multiple sites, including several outside the derivation setting (sites included two pediatric EDs, one general ED, two minor injury units and one regional children's Burns Centre in Bristol and Cardiff, UK). The same inclusion criteria were applied. Constant monitoring of the sites was provided to maximize completion of proformas; missing data fields were completed from case notes where possible. Data were entered onto a REDCap (Research Electronic Data Capture)[24] database using double data checking by two independent data clerks.

Classification of Cases

Cases with complete information on all the predictor variables were retrospectively assigned a total BuRN-Tool score according to the individual score for each variable and blinded to case outcome (referral to children's social care teams). ROC curve analysis was performed against the outcome and all diagnostic accuracy measures (sensitivity, specificity, LR, PPV, NPV) were obtained with the same methods as used for the derivation data. We identified the proportion of cases scoring greater than or equal to the chosen cut-off score.

Ethical approval for the derivation study was granted by Research Ethics Committee for England and Wales 08/H0504/133 Dublin REC 2009/02/02 and for the validation study

(MREC 13/WA/0003). Waived consent for the derivation study and the process of data recording were approved by the National Information Governance Board (PIAG4-05(i)2008) and for the external validation by the Confidentiality Advisory Group (CAG 1-06 (PR7)/2013).

RESULTS

Derivation of BuRN-Tool

Ascertainment and Demographics

A total of 1484 cases were identified, 157 were excluded due to double counting or failure to fulfil the inclusion criteria, leaving 1327 cases; 768 (58%) were scalds (median age; 1 year [IQR 1-3]) and 559 (42%) non-scalds (median age; 2 years [IQR 1-8]). Gender ratio (Male: Female) was 3:2. A total of 8.4% (112/1327; 7.7% for scalds and 9.5% for non-scalds) cases were referred to a children's social care team. Data completeness for scald cases included in the analyses was 99.6 % (765/768) and 100% for non-scalds.

Predictor Variables

Univariate and multivariable analysis showed that the strongest predictor of referral to a children's social care team was whether the child was previously known to social care at the point of ED attendance, followed by a concerning injury explanation and the injury being full thickness. Children with *scalds* were also more likely to be referred to a social care team if they were <5 years of age. A weaker association was found with referral if scalds showed bilateral patterning, or if the injury was in an uncommon body site (back, buttocks, groin, head within the hairline) (Supplementary Table 1, Table 2). The association with age and an uncommon body site was not seen for *non-scalds*. However, as age showed a strong association with referral for scalds, we kept it in the multivariable analysis to see whether

adjusting for the other variables would modify this relationship. Adjustment for the strongly influential variables did strengthen the association between age <5 and referral of *non-scalds*. Concerns about adult supervision showed a weaker association for both injury types (Table 2). No association was found with multiple burn sites or first aid provision. Although we had the data, late presentation was often inconsistently and unreliably recorded and was therefore omitted from further analyses. No change in odds ratios were found using bootstrapping (data not shown).

INSERT TABLE 2 HERE

Scoring

Converting the multivariable regression coefficient of the predictive variables to create the BuRN-Tool, resulted in integer scores ranging from 1-3 (Table 3). As age was less influential in non-scalds (score=1) compared to scalds (score=2), optimal cut-off scores in the ROC analyses differed by 1. For simplicity of application in clinical settings, every non-scald burn therefore received an additional score of 1 in order to use the same cut-off for both injury types. The maximum summed score obtainable for scalds was 12 and for non-scalds 10 (Table 3).

INSERT TABLE 3 HERE

Based on the ROC curve analyses (curves shown in supplementary Figures 2 and 3), a cut-off score of three corresponded to the optimum sensitivity and specificity for both injury types (Table 4), maximizing the discovery of suspected maltreatment cases while reducing

unnecessary referrals. Proportion of cases scoring three or above was 30.3% (232/765) for scalds and 28.8% (161/559) for non-scalds. Predictive accuracy, measured as the AUC, was 0.79 (95% CI: 0.75, 0.81) for scalds and 0.75 (95% CI: 0.71, 0.78) for non-scalds. For scalds, a score of three represented a positive likelihood ratio (LR+) of 2.7, and a score of four or more an LR+ ranging from 6.2-18.0 (Table 4). A score of 0-2 corresponded to a low negative likelihood ratio (LR-) of 0.2. For non-scalds, the values were very similar (Table 4). Supplementary Table 2 shows the contribution of the different variables to the overall total score.

INSERT TABLE 4

External Validation of the BuRN-Tool

Ascertainment and Demographics

In the validation study, data were recorded for 389 children with scalds and 398 with non-scalds. Median age was two years (IQR 1-6) for both scalds and non-scalds, and male:female ratio 1.2:1. Overall rate of referral to a children's social care team was 6.5% (50/769; 4.2% for scalds and 8.9% for non-scalds; 18 cases had missing data for referral status). Data completeness was 95.9% (373/389) for scalds and 93.5% (372/398) for non-scalds.

Scoring and Classification against Outcome

As in the derivation data, a cut-off score of three corresponded to the optimum sensitivity and specificity for both scalds and non-scalds (Table 5; ROC curves shown in supplementary Figure 4, 5); sensitivity and specificity for scalds were 87.5% (95% CI: 61.7-

98.4) and 81.5% (95% CI: 77.1-85.4) and for non-scalds 82.4% (95% CI: 65.5-93.2) and 78.7% (95% CI: 73.9-82.9). Proportion of cases scoring three or above was 21.4% (80/373) for scalds and 26.9% (100/372) for non-scalds. Predictive accuracy was greater than in the derivation setting; AUC was 0.87 (95% CI: 0.83-0.90) for scalds and 0.85 (95% CI: 0.81-0.88) for non-scalds. A score of ≥ 3 showed an incremental increase in LR+ for both *scalds* (range 4.7-22.3) and *non-scalds* (range 3.9-29.8) with a LR- of < 0.3 if the score was 0-2 (Table 5).

INSERT TABLE 5

Misclassification

There were seven *false-negative* cases that were referred to a children's social care team but scored 0-2. In three cases there were additional concerning factors, such as domestic abuse, and self-harm that triggered a referral. The potential reasons for referral in the other cases were unclear. The overall *false positive rate* (not referred but scored ≥ 3) was 19% (138/745), 59% (82/138) of these had a BuRN-Tool score of three. The distribution of predictor variables, stratified by age, amongst the false-positive cases are shown in supplementary Table 3. Distributions showed relatively equal spread for both age categories but there was a much larger proportion of false-positive cases previously known to social care in children ≥ 5 years old (64%) compared to in those < 5 years old (13%). With regards to supervision concerns, the reverse was true, as 52% of false positives in < 5 year olds had supervision concerns compared to 8% in over ≥ 5 year olds.

DISCUSSION

The BuRN-Tool is an easily completed CPT with the potential to identify which children with burns warrant further child maltreatment evaluation. The performance of the BuRN-Tool was prospectively validated against a novel dataset to that used for derivation. Although not perfect, the performance of the BuRN-Tool was satisfactory in both derivation and validation and performed slightly better in the validation study with a sensitivity and specificity around 80%.

The strengths of this study are that the standardized derivation methodology[11, 25] was followed in developing a novel evidence-based CPT. The clinical proforma had good data completion rates, especially within the validation study where ED staff received training and reminders to complete the data fields, and the sample was large enough to estimate diagnostic accuracy with reasonable certainty. The predictor variables are factors that should be recorded as part of any standard assessment of a child with a burn. A survey of 54 clinicians found that the research proforma was accepted as a quick and easy to complete clinical record and 90% of participants reported that they would use the CPT to support child protection decisions in the ED.[26]

In practice, the BuRN-Tool identified a number of children as being ‘of concern’ who were not referred for child protection assessment, especially infants and toddlers where there was concern about inadequate adult supervision. Any burn in a young child reflects some degree of inadequate adult supervision. This decision is arguably a subjective one, but potentially a key indicator of neglect. We attempted to make this an objective measure, as far as was possible. The BuRN-Tool recorded ‘supervision concern’ in young children who had no adult present at the time that they sustained a burn, or where the clinician was concerned about inadequate supervision. In either case, further assessment of the home situation in these young children and ongoing support would be warranted. It could be argued that children

previously or currently known to social care should not automatically be referred if they sustain a burn. However this factor was the strongest predictor variable in the derivation study, and has been identified in systematic reviews as a strong predictor of abuse.[8] Sharing of information about vulnerable children is an essential part of child protection to enable an increased level of parental support. Both level of appropriate supervision, and an assessment as to whether an explanation for the injury is of concern, are subjective decisions. The advantage of the BuRN-Tool is that it requires clinicians to consider these aspects and to make a clinical judgement. We have developed a short webinar resource for ED clinicians to use alongside the BuRN-Tool, to provide training for ED clinicians about the pre-determinants of childhood burns and their relationship to child development and parenting.

Whilst a good CPT should ensure that all concerning cases are recognized, this should not be falsely inflated at the expense of children of no concern being referred to a children's social care teams. The BuRN-Tool offers an incrementally high LR+ for scores ≥ 3 and low LR- for scores of 0-2, minimizing the risk of unnecessary referrals. A low score has the potential to reassure clinicians that maltreatment is unlikely. However, with only a few cases scoring at the top end of the range of scores, the precision of LRs for these is consequently low with wide confidence intervals and slight fluctuations in sample size could change the LRs substantially.[27]

There were additional variables included on the standardized clinical proforma (BaSAT) which are currently not in our BuRN-Tool. Late presentation to medical care has been suggested as an indicator for child protection concerns.[4] However, data for this variable were poorly recorded, and within the care pathways it was difficult to define either the time between injury and first presentation, or the reasons for late presentation, which were often related to the fact that the burn looked considerably worse some hours after it was

sustained and alerted the worried parent to seek medical advice after some delay. Other factors such as skin fold or central buttock sparing, or glove and stocking appearance of burns [17] have been described in physical abuse, however these features were rarely described in the study population and thus it was not possible to include them within the data analysis. Additionally, a child with a burn may have multiple factors unrelated to the burn that warrant a social care referral (e.g. parents under the influence of alcohol, domestic violence, and inflicted injuries other than a burn). These factors may have contributed to the false negative cases. The BuRN-Tool is designed to consider the burn itself within a holistic assessment of the child, and clinicians must be aware of the multitude of other signs and symptoms of child maltreatment that may co exist.

When compared to other child maltreatment CPTs, the BuRN-Tool is evidence based, and was derived and validated against a much larger sample than the previously published checklist for children's burns introduced by Clark in 1997.[15] The performance of the BuRN-Tool compares well with other CPTs developed for injuries - for example those designed to 'screen out' Abusive Head Trauma,[28] or identify potential abuse based on bruising patterns of young children admitted to intensive care.[29] Both of these are applicable to children admitted with serious injury, and were not designed to screen the large volume of children with the full spectrum of injury severity attending an ED. In the Netherlands, when a modified SPUTOVAMO was evaluated in a cohort of pediatric admissions to a regional burns centre, 12% screened positive for maltreatment concern.[3] Whilst it is difficult to make a direct comparison of performance with the BuRN-Tool due to a different population, different variables and outcome measure, the sensitivity (73.2%) and specificity (94.5%) based upon their outcome measure of children requiring additional support were comparable with those achieved by the BuRN-Tool.

The limitation of this study, in common with many studies in this field, is determining a categorical outcome of maltreatment. There is no available gold standard for diagnosis. As the over-riding purpose of this study was to identify cases that need to be referred to a children's social care team for further assessment, we used practitioners' referral to children's social care as the outcome measure in the knowledge that maltreatment is confirmed in an estimated 70-75% of children referred by clinicians to children's social care for suspected maltreatment.[30, 31] The BuRN-Tool in its current form relies upon features drawn from a standardized ED assessment. These features carry scores of different weighting that are added together to give a total score. The total score has the potential to identify children who warrant referral to children's social care teams for a comprehensive child protection assessment or to reassure ED staff of those children with a low score who carry a low risk of maltreatment. A recent study from a centre with 40 years of experience with pediatric burns identified a 7.1% prevalence of children referred to child protection services for possible 'non accidental injury', social care confirmed abuse in 75.7% of these cases. The authors identified similar findings to our study and recommended that 'Clinicians should approach all burn injuries in young children with a high index of suspicion, but in particular those with scalds, or injuries to the buttocks, perineum, or bilateral feet should provoke suspicion'.[31]

The BuRN-Tool confirms that there is a constellation of features that strongly predict which children are referred to social care for further assessment, many of them common to other reported literature.[31] Further work has now been funded to identify the effect of the BuRN-Tool when implemented into clinical practice, what action social and health care workers take when alerted to these children and what proportion of referred children are substantiated cases of maltreatment. This will enable us to establish whether our tool would work as a predictor of maltreatment and whether additional features could improve its sensitivity and

specificity.

CONCLUSION

The BuRN-Tool is a prospectively derived level 3 (highest evidence-based standard of validation) CPT,[11, 25] with satisfactory accuracy, that has been developed from research evidence and validated prospectively. The items are basic clinical features and the scoring is straight forward; together with the BaSAT it has the potential to act as an ‘aide memoire’ to standardize the assessment of children with burns across multiple professionals in busy emergency settings, and identify those who warrant a full child maltreatment evaluation and exclude those who do not.

ACKNOWLEDGEMENTS

The authors wish to thank everyone who contributed data to this study. The participating EDs and units and lead clinicians were:

Derivation study: Cardiff University / Cardiff and Vale University Hospital Board, Wales (Z Lawson, H Dowd, D Farrell), Birmingham Childrens Hospital, England (G DeBelle, C Thomas), Royal Gwent Hospital, Newport, Wales (A Rawlinson, S Jones, J Probert), Morriston Hospital, Swansea, Wales (T Potokar, P Thompson, A McNab), Our Lady's Hospital for Sick Children, Crumlin, Dublin, Ireland (R Sullivan, S Harty), Adelaide and Meath Hospital, Tallaght, Dublin, Ireland (E Curtis, O Callendar), Children's University Hospital, Temple Street, Dublin, Ireland (M McKay).

Validation study: Bristol Royal Hospital for Children Emergency Department (Dr Mark Lyttle), North Bristol NHS Trust Emergency Department (Dr Jason Kendall), North Bristol NHS Trust Minor Injury Unit (Gill Rodgers), South West UK Children's Burns Network (Dr Amber Young), Paediatric Emergency Department, University Hospital of Wales, (Dr Zoe Roberts), Minor Injuries Unit, Barry Community Hospital, (Melanie Noble).

Study data were collected and managed using REDCap (Research Electronic Data Capture) tools hosted at the University of Bristol. The Centre for Children's Burns Research is part of the Burns Collective, a Scar Free Foundation initiative with additional funding from the Vocational Training Charitable Trust VTCT and Health and Care Research Wales. The views expressed are those of the authors, and not necessarily those of the Scar Free Foundation or other funding bodies.

REFERENCES

- [1] The Royal Society for the Prevention of Accidents (RoSPA). Home and Leisure Accident Surveillance System (HASS and LASS) data. 2002.
- [2] American Burn Association. ABA National Burn Repository.
- [3] Bousema S, Stas H, van de Merwe M, Oen I, Baartmans M, van Baar M, et al. Epidemiology and screening of intentional burns in children in a Dutch burn centre. *Burns*. 2016.
- [4] Chester DL, Jose RM, Aldlyami E, King H, Moiemmen NS. Non-accidental burns in children - are we neglecting neglect? *Burns*. 2006;32:222-8.
- [5] Ojo P, Palmer J, Garvey R, Atweh N, Fidler P. Pattern of burns in child abuse. *Am Surg*. 2007;73:253-5.
- [6] Thombs B. Patient and injury characteristics, mortality risk, and length of stay related to child abuse by burning: evidence from a national sample of 15,802 pediatric admissions. *Ann Surg*. 2008;247:519-23.
- [7] Wibbenmeyer L, Liao J, Heard J, Kealey L, Kealey G, Oral R. Factors related to child maltreatment in children presenting with burn injuries. *J Burn Care Res*. 2014;35:374-81.
- [8] Woodman J, Pitt M, Wentz R, Taylor B, Hodes D, Gilbert R. Performance of screening tests for child physical abuse in accident and emergency departments. *Health Technol Assess*. 2008;12:1-95.
- [9] Great Britain. Children Act 2004 (c.31) London: The Stationery Office. 2004.
- [10] Wallace E, Smith S, Perera-Salazar R, Vaucher P, McCowan C, Collins G, et al. Framework for the impact analysis and implementation of Clinical Prediction Rules (CPRs). *BMC Med Inform Decis Mak*. 2011;11:1-7.
- [11] Stiell I, Bennett C. Implementation of clinical decision rules in the emergency department. *Acad Emerg Med*. 2007;14:955-9.
- [12] McGinn T, Wyer P, McCullagh L, Wisnivesky J, Devereaux P, Stiell I, et al. 19.4 Clinical Prediction Rules. In: Guyatt G, Rennie D, Meade M, Cook D, editors. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. 3rd ed. New York: McGraw-Hill Education; 2015.
- [13] Green S, Schriger D, Yealy D. Methodologic standards for interpreting clinical decision rules in emergency medicine: 2014 update. *Ann Emerg Med*. 2014;64:286-91.
- [14] Bailhache M, Leroy V, Pillet P, Salmi L. Is early detection of abused children possible? A systematic review of the diagnostic accuracy of the identification of abused children. *BMC Pediatr*. 2013;13:1-11.
- [15] Clark KD, Tepper D, Jenny C. Effect of a screening profile on the diagnosis of nonaccidental burns in children. *Pediatr Emerg Care*. 1997;13:259-61.
- [16] Kemp A, Jones S, Lawson Z, Maguire S. Patterns of burns and scalds in children. *Arch Dis Child*. 2014;99:316-21.
- [17] Maguire S, Moynihan S, Mann M, Potokar T, Kemp AM. A systematic review of the features that indicate intentional scalds in children. *Burns*. 2008;34:1072-81.
- [18] StataCorp. *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP; 2015.
- [19] Wong C, Khin L, Heng K, Tan K, Low C. The LRINEC (Laboratory Risk Indicator for Necrotizing Fasciitis) score: a tool for distinguishing necrotizing fasciitis from other soft tissue infections. *Crit Care Med*. 2004;32:1535-41.
- [20] Cole T. Algorithm AS 281: Scaling and rounding regression coefficients to integers. *Journal of the Royal Statistical Society Series C (Applied Statistics)*. 1993;42:261-8.

- [21] Brown M, Reeves M. Evidence-based emergency medicine/skills for evidence-based emergency care. Interval likelihood ratios: another advantage for the evidence-based diagnostician. *Ann Emerg Med.* 2003;42:292-7.
- [22] Deeks J, Altman D. Diagnostic tests 4: likelihood ratios. *BMJ.* 2004;329:168-9.
- [23] Florkowski C. Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. *Clin Biochem Rev.* 2008;29:s83-7.
- [24] Harris P, Taylor R, Thielke R, Payne J, Gonzalez N, Conde J. Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42:377-81.
- [25] Reilly B, Evans A. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. *Ann Intern Med.* 2006;144:201-9.
- [26] Johnson E, Hollén L, Kemp A, Maguire S. Exploring the acceptability of a clinical decision rule to identify paediatric burns due to child abuse or neglect. *Emergency Medicine Journal.* 2016;33:465-70.
- [27] Dujardin B, Van den Ende J, Van Gompel A, Unger J, Van der Stuyft P. Likelihood ratios: a real improvement for clinical decision making? *European Journal of Epidemiology.* 1994;10:29-36.
- [28] Hymel K, Armijo-Garcia V, Foster R, Frazier T, Stoiko M, Christie L, et al. Validation of a clinical prediction rule for pediatric abusive head trauma. *Pediatrics.* 2014;134:e1537-44.
- [29] Pierce MC, Kaczor K, Aldridge S, O'Flynn J, Lorenz DJ. Bruising characteristics discriminating physical child abuse from accidental trauma. *Pediatrics.* 2010;125:67-74.
- [30] Kemp A, Maguire S, Nuttall D, Collins P, Dunstan F. Bruising in children who are assessed for suspected physical abuse. *Arch Dis Child.* 2014;99:108-13.
- [31] Hodgman E, Pastorek R, Saeman M, Cripps M, Bernstein I, Wolf S, et al. The Parkland Burn Center experience with 297 cases of child abuse from 1974 to 2010. *Burns.* 2016;42:1121-7.
- [32] HM Government. Working together to safeguard children. A guide to inter-agency working to safeguard and promote the welfare of children. 2015.

Table 1. Definitions of predictor variables used in the *derivation* of the BuRN-Tool.

Predictor Variables	Categories	Explanation
Age	<5 years ≥5 years	Age of the child at the time of the scald/non-scald
Previously or currently known to social care	Yes/No	Yes if child/family member is: (a) child in need (b) child with protection plan (c) previously known to social care/social welfare (d) child/ family has/have had in past an allocated social worker[32]
Severity	Full thickness/not full thickness	Scald/non-scald classed as full thickness or not
Concern about the explanation for the burn injury *	Yes/No	Yes if: (a) concerns that explanation is not consistent with stage of development (b) concerns that explanation does not fit the burn pattern seen (c) injury is a bath scald
Body site location *	Common/uncommon	Uncommon if body location was any of those represented by <5% of cases in the derivation data (supplementary Table 1): For scalds : Uncommon if uppermost location included back, buttock, groin or head (within the hairline) For non-scalds : Uncommon if most severely affected area is anything other than face, hand, arm or leg
Supervision concern	Yes/No	Yes if: (a) clinician has concerns about appropriate supervision (b) child is <5 years of age and without an adult in the room at the time of injury
Pattern	Unilateral/Bilateral	Bilateral if the scald affects both sides of the body (scalds only)
Multiple burn sites	Yes/No	Yes if burns to several body sites. Contact burns only
First aid provision	Yes/No	Yes if first aid given before seeking medical attention
Late presentation	Yes/No	Yes if child presented > 24 hours post incident

* The body site and injury explanation variables were based upon the findings that common (typical) explanations for unintentional burns or scalds and predictable injury distributions are seen in children at different stages of motor development. For example, hot beverage scalds to the torso are typically seen in infants or toddlers who pull the beverage down on top of themselves as are contact burns to the hands in children who touch a hot iron or hair straighteners.¹⁵

Table 2. Univariate and multivariable associations with the outcome referral to child social care teams for children with scalds or non-scalds in the *derivation* data. Unadjusted odds ratios (ORs) and 95% confidence intervals (CIs) are from univariate logistic regression. Fully adjusted ORs are from multivariable logistic regressions.

Scalds <i>Predictor variables</i>	Referral				<i>Unadjusted ORs (95% CIs)</i>	<i>Fully adjusted ORs (95% CIs)*</i>
	<i>Yes N = 59</i>		<i>No N = 709</i>			
	N	%	N	%		
Previously known to Social Care (yes)	11	19	10	1	16.02 (6.48, 39.59)	24.23 (8.99, 65.30)
Severity (full thickness)	4	7	7	1	7.29 (2.07, 25.68)	8.71 (1.97, 38.58)
Concerning explanation (yes)	13	22	25	4	7.73 (3.71, 16.10)	4.90 (2.08, 11.57)
Age (<5 years)	56	95	554	78	5.22 (1.61, 16.91)	4.85 (1.42, 16.53)
Uncommon body site (yes)	9	15	42	6	2.66 (1.23, 5.74)	2.98 (1.24, 7.18)
Bilateral scald	19	32	112	16	2.53 (1.41, 4.53)	1.94 (0.96, 3.91)
Supervision concern (yes)	16	27	128	18	1.69 (0.92, 3.09)	2.41 (1.17, 4.95)
First aid provision (no)	6	10	133	19	0.49 (0.21, 1.16)	<i>Not included</i>

Non-scalds <i>Predictor variables</i>	Referral				<i>Unadjusted ORs (95% CIs)</i>	<i>Fully adjusted ORs (95% CIs)*</i>
	<i>Yes N = 53</i>		<i>No N = 506</i>			
	N	%	N	%		
Previously known to Social Care (yes)	10	19	9	2	12.84 (4.95, 33.30)	13.25 (4.56, 38.49)
Concerning explanation (yes)	14	26	23	5	7.54 (3.60, 15.80)	6.68 (2.87, 15.56)
Severity (full thickness)	14	26	35	7	4.83 (2.40, 9.73)	5.44 (2.50, 11.82)
Supervision concern (yes)	22	42	42	28	1.82 (1.02, 3.25)	1.71 (0.88, 3.29)
Uncommon body site (yes)	9	17	55	11	1.68 (0.78, 3.62)	<i>Not included</i>
Age (<5 years)	35	66	324	64	1.09 (0.60, 1.98)	2.08 (1.02, 4.26)
First aid provision (no)	16	30	156	31	0.97 (0.52, 1.80)	<i>Not included</i>
Multiple sites [†]	4/31	12	70/359	20	0.61 (0.21, 1.80)	<i>Not included</i>

*All variables with P<0.10 remained in fully adjusted model. Three scald cases omitted as missing data so total sample size for scalds in multivariable analysis is 765. Results remained the same after bootstrapping (not shown).

[†] Sample size different as multiple sites refer to contact burns only.

Table 3. Regression coefficients (Coef = ln OR) from the *derivation* variables predictive of referral to children’s social care teams included in the BuRN-Tool and their respective converted integer scores. Coefficients are from fully adjusted multivariable models.

Predictor variables	Scalds		Non-scalds	
	Coef	Score	Coef	Score
<i>Previously known to Social Care</i>				
No	0	0	0	0
Yes	3.2	3	2.6	3
<i>Severity</i>				
Not full thickness	0	0	0	0
Full thickness	2.2	2	1.7	2
<i>Concerning explanation</i>				
No	0	0	0	0
Yes	1.6	2	1.9	2
<i>Age</i>				
≥5 years	0	0	0	0
<5 years	1.6	2	0.7	1
<i>Concerns about supervision</i>				
No	0	0	0	0
Yes	0.9	1	0.5	1
<i>Uncommon body site (back, buttock, groin, within hairline)*</i>				
No	0	0		
Yes	1.09	1		
<i>Pattern*</i>				
Unilateral	0	0		

Bilateral scald	0.7	1
<i>Maximum score</i>	12	10**

* Uncommon body site ([supplementary table 1]) was not influential in non-scalds and consequently not scored. Pattern was not scored for non-scalds as was only relevant for scalds.

** As age was less influential in non-scalds, a score of 1 was added to every non-scald in order to keep the same clinical cut-off score (see methods).

Table 4. Measures of diagnostic accuracy for scalds and non-scalds in the derivation data. LR+ = Positive Likelihood Ratio, LR- = Negative Likelihood Ratio, PPV = Positive Predictive Value. NPV = Negative Predictive Value. Prevalence of referral to child social care teams was 7.7% for scalds and 9.5% for non-scalds.

Cut-point score	Sensitivity % (95% CI)	Specificity % (95% CI)	LR+ (95% CI)*	LR- (95% CI)*	PPV (95% CI)	NPV (95% CI)	N
<i>Derivation scalds</i>							
(>= 0)	100.0 (93.9, 100.0)	0.0 (0.0, 0.6)	1.0	-	8.5 (6.5, 10.8)	-	72
(>= 1)	100.0 (93.9, 100.0)	10.2 (8.1, 12.7)	1.1 (1.09, 1.14)	0	8.5 (6.5, 10.8)	100.0 (95.0, 100.0)	70
(>= 2)	96.6 (88.3, 99.6)	19.8 (16.9, 23.0)	1.2 (1.1, 1.3)	0.2 (0.04, 0.7)	9.2 (7.0, 11.7)	98.6 (95.0, 99.8)	391
(>= 3)	71.2 (57.9, 82.2)	73.1 (69.7, 76.3)	2.7 (2.2, 3.2)	0.4 (0.3, 0.6)	18.1 (13.4, 23.7)	96.8 (94.9, 98.1)	144
(>= 4)	50.8 (37.5, 64.1)	91.8 (89.5, 93.7)	6.2 (4.4, 8.8)	0.5 (0.4, 0.7)	34.1 (24.3, 45.0)	95.7 (93.9, 97.1)	42
(>= 5)	37.3 (25.0, 50.9)	96.6 (95.0, 97.8)	11.0 (6.6, 18.3)	0.7 (0.5, 0.8)	47.8 (32.9, 63.1)	94.9 (93.0, 96.4)	31
(>= 6)	15.3 (7.2, 27.0)	99.2 (98.2, 99.7)	18.0 (6.6, 48.7)	0.9 (0.8, 1.0)	60.0 (32.3, 83.7)	93.3 (91.3, 95.0)	10
(>= 7)	5.1 (1.1, 14.1)	99.7 (99.0, 100.0)	18.0 (3.1, 105.3)	1.0 (0.9, 1.0)	60.0 (14.7, 94.7)	92.6 (90.5, 94.4)	4
(>= 8)	1.7 (0.04, 9.1)	100.0 (99.5, 100.0)	-	1.0 (0.95, 1.02)	100.0 (2.5, 100.0)	92.4 (90.3, 94.2)	1
<i>Derivation non-scalds</i>							

(≥ 1)	100.0 (93.3, 100.0)	0.0 (0.0, 0.7)	1.0	-	9.5 (7.2, 12.2)	-	95
(≥ 2)	92.5 (81.8, 97.9)	18.0 (14.7, 21.6)	1.1 (1.0, 1.2)	0.4 (0.2, 1.1)	10.6 (7.9, 13.7)	95.8 (89.6, 98.8)	303
(≥ 3)	66.0 (51.7, 78.5)	75.1 (71.1, 78.8)	2.7 (2.1, 3.4)	0.5 (0.3, 0.7)	21.7 (15.6, 28.9)	95.5 (92.9, 97.3)	84
(≥ 4)	50.9 (36.8, 64.9)	90.5 (87.6, 92.9)	5.2 (3.6, 7.5)	0.5 (0.4, 0.7)	36.0 (25.2, 47.9)	94.6 (92.2, 96.5)	46
(≥ 5)	32.1 (19.9, 46.3)	97.2 (95.4, 98.5)	11.6 (6.1, 22.2)	0.7 (0.6, 0.8)	54.8 (36.0, 72.7)	93.2 (90.7, 95.2)	19
(≥ 6)	17.0 (8.1, 29.8)	99.4 (98.3, 99.9)	28.6 (8.0, 102.6)	0.8 (0.7, 0.9)	75.0 (42.8, 94.5)	92.0 (89.4, 94.1)	6
(≥ 7)	9.4 (3.1, 20.7)	99.8 (98.9, 100.0)	47.7 (5.7, 401.0)	0.9 (0.8, 1.0)	83.3 (35.9, 99.6)	91.3 (88.7, 93.5)	5
(≥ 8)	1.9 (0.1, 10.1)	100.0 (99.3, 100.0)	-	1.0 (0.95, 1.02)	100.0 (2.5, 100.0)	90.7 (88.0, 93.0)	1

*Two decimals only shown if upper and lower confidence limits appear to be the same when corrected to one decimal place

Table 5. Measures of diagnostic accuracy for scalds and non-scalds in the validation data. LR+ = Positive Likelihood Ratio, LR- = Negative Likelihood Ratio, PPV = Positive Predictive Value. NPV = Negative Predictive Value. Prevalence of referral to children’s social care teams was 4.2% for scalds and 8.9% for non-scalds.

Cut-point score	Sensitivity % (95% CI)	Specificity % (95% CI)	LR+ (95% CI)*	LR- (95% CI)*	PPV (95% CI)	NPV (95% CI)	N
<i>Validation scalds</i>							
(≥ 0)	100.0 (79.4, 100.0)	0.0 (0.0, 1.0)	1.0	-	4.3 (2.5, 6.9)	-	80
(≥ 1)	93.8 (69.8, 99.8)	22.1 (17.9, 26.8)	1.2 (1.1, 1.4)	0.3 (0.04, 1.9)	5.1 (2.9, 8.3)	98.8 (93.2, 100.0)	9
(≥ 2)	93.8 (69.8, 99.8)	24.6 (20.3, 29.5)	1.2 (1.1, 1.5)	0.3 (0.04, 1.6)	5.3 (3.0, 8.6)	98.9 (93.9, 100.0)	204
(≥ 3)	87.5 (61.7, 98.4)	81.5 (77.1, 85.4)	4.7 (3.6, 6.3)	0.2 (0.04, 1.6)	17.5 (9.9, 27.6)	99.3 (97.6, 99.9)	45
(≥ 4)	62.5 (35.4, 84.8)	93.0 (89.8, 95.4)	8.9 (5.2, 15.3)	0.4 (0.2, 0.8)	28.6 (14.6, 46.3)	98.2 (96.2, 99.3)	13
(≥ 5)	50.0 (24.7, 75.3)	96.1 (93.5, 97.8)	12.8 (6.3, 25.9)	0.5 (0.3, 0.9)	36.4 (17.2, 59.3)	97.7 (95.6, 99.0)	14
(≥ 6)	25.0 (7.3, 52.4)	98.9 (97.2, 99.7)	22.3 (6.1, 81.2)	0.8 (0.6, 1.0)	50.0 (15.7, 84.3)	96.7 (94.3, 98.2)	3
(≥ 7)	12.5 (1.6, 38.3)	99.2 (97.6, 99.8)	14.9 (2.7, 82.9)	0.9 (0.7, 1.1)	40.0 (5.3, 85.3)	96.2 (93.7, 97.9)	3
(≥ 8)	6.3 (0.2, 30.2)	99.7 (98.4, 100.0)	22.3 (1.5, 340.8)	0.9 (0.8, 1.1)	50.0 (1.3, 98.7)	96.0 (93.4, 97.7)	2

<i>Validation non-scalds</i>							
(≥ 1)	100.0 (89.7, 100.0)	0.0 (0.0, 1.1)	1.0	-	9.1 (6.4, 12.5)	-	93
(≥ 2)	94.1 (80.3, 99.3)	26.9 (22.3, 32.0)	1.3 (1.2, 1.4)	0.2 (0.1, 0.9)	11.5 (8.0, 15.8)	97.8 (92.4, 99.7)	179
(≥ 3)	82.4 (65.5, 93.2)	78.7 (73.9, 82.9)	3.9 (3.0, 5.0)	0.2 (0.1, 0.5)	28.0 (19.5, 37.9)	97.8 (95.3, 99.2)	47
(≥ 4)	64.7 (46.5, 80.3)	90.8 (87.2, 93.7)	7.1 (4.7, 10.7)	0.4 (0.3, 0.6)	41.5 (28.1, 55.9)	96.2 (93.5, 98.0)	24
(≥ 5)	47.1 (29.8, 64.9)	96.2 (93.5, 97.9)	12.2 (6.4, 23.2)	0.6 (0.4, 0.8)	55.2 (35.7, 73.6)	94.8 (91.8, 96.9)	16
(≥ 6)	29.4 (15.1, 47.5)	99.1 (97.4, 99.8)	33.1 (9.6, 114.6)	0.7 (0.6, 0.9)	76.9 (46.2, 95.0)	93.3 (90.2, 95.7)	8
(≥ 7)	8.8 (1.9, 23.7)	99.4 (97.9, 99.9)	14.9 (2.6, 86.2)	0.9 (0.8, 1.0)	60 (14.7, 94.7)	91.6 (88.2, 94.2)	1
(≥ 8)	8.8 (1.9, 23.7)	99.7 (98.4, 100.0)	29.8 (3.2, 278.9)	0.9 (0.8, 1.0)	75.0 (19.4, 99.4)	91.6 (88.3, 94.2)	3
(≥ 10)	2.9 (0.1, 15.3)	100.0 (98.9, 100.0)	-	1.0 (0.9, 1.0)	100.0 (2.5, 100.0)	91.1 (87.7, 93.8)	1

* Two decimals only shown if upper and lower confidence limits appear to be the same when corrected to one decimal place

FIGURES

Supplementary figure 1. Burns and Scalds Assessment Template (BaSAT)

Supplementary figure 2. Receiver operating curve for scalds in the derivation data set.

Supplementary figure 3. Receiver operating curve for non-scalds in the derivation data set.

Supplementary figure 4. Receiver operating curve for scalds in the validation data set.

Supplementary figure 5. Receiver operating curve for non-scalds in the validation data set.