What faces reveal: a novel method to identify patients at risk of deterioration using facial expressions

MADRIGAL-GARCIA, M, RODRIGUES, Marcos <http://orcid.org/0000-0002-6083-1303>, SHENFIELD, Alex <http://orcid.org/0000-0002-2931-8077>, SINGER, Mervyn and MORENO-CUESTA, Jeronimo

Available from Sheffield Hallam University Research Archive (SHURA) at:
http://shura.shu.ac.uk/19049/

This document is the author deposited version. You are advised to consult the publisher's version if you wish to cite from it.

Published version


Copyright and re-use policy

See http://shura.shu.ac.uk/information.html
Title

What faces reveal: a novel method to identify patients at risk of deterioration using facial expressions

Maria Isabel Madrigal-Garcia (1) MSc, Marcos Rodrigues (2) PhD, Alex Shenfield (2) PhD, Mervyn Singer (3) MD, FRCP, FFICM, Jeronimo Cuesta (1) MD

Department of Intensive Care, North Middlesex Hospital, London (1); GMPR group Department of Engineering and Mathematics, Sheffield Hallam University, Sheffield (2), and Bloomsbury Institute of Intensive Care Medicine, Division of Medicine, University College London (3), United Kingdom.

Corresponding author

Dr Jeronimo Moreno Cuesta

North Middlesex University Hospital, Sterling Way N18 1QX,
London, United Kingdom.

Phone: 0208 887 2414; Fax: 0208 887 2669

Email: Jeronimo.Moreno-Cuesta@nhs.net

Address for reprints

Dr Jeronimo Moreno Cuesta

North Middlesex University Hospital, Sterling Way N18 1QX,
London, United Kingdom.

Phone: 0208 887 2414; Fax: 0208 887 2669

Email: Jeronimo.Moreno-Cuesta@nhs.net
Conflict of interest

None

Funding/support

No financial support was requested

Key words

Stress, decision-making, facial recognition, facial expression, intuition
Abstract

Objectives: To identify facial expressions occurring in patients at risk of deterioration in hospital wards.

Design: Prospective observational feasibility study

Setting: General ward patients in a London Community Hospital, United Kingdom

Patients: Thirty-four patients at risk of clinical deterioration

Interventions: A 5 minute video (25 frames per second, 7,500 images) was recorded, encrypted and subsequently analysed for Action Units (AU) by a trained Facial Action Coding System (FACS) psychologist blinded to outcome.

Measurements and Main Results: AU of the upper face (UF), head position (HP), eyes position (EP), lips and jaw position (LJ) and lower face (LF) were analysed in conjunction with clinical measures collected within the National Early Warning Score (NEWS). The most frequently detected AU were AU43 (73%) for UF, AU51 (11.7%) for HP, AU62 (5.8%) for EP, AU25 (44.1%) for LJ and AU15 (67.6%) for LF. The presence of certain combined face displays (FD) was increased in patients requiring admission to intensive care, namely, AU 43+15+25 (FD1, p <0.013), AU 43+15+51/52 (FD2, p <0.003) and AU 43+15+51+25 (FD3, p <0.002). Having FD1, FD2 and FD3 increased the risk of being admitted to intensive care 8-fold, 18-fold and as a sure event, respectively. A logistic regression model with FD1, FD2, FD3 and NEWS as independent covariates described admission to intensive care with an average concordance statistic (c-index) of 0.71 (p = 0.009).

Conclusion: Patterned facial expressions can be identified in deteriorating general ward patients. This tool may potentially augment risk prediction of current scoring systems.
Introduction

The stress response is an evolutionary, non-specific and compensatory allostatic reaction to any aggression. It serves to adapt the body, both behaviourally and physiologically, to deal with any stressful situation. The point where the stress response is triggered is individual-specific (resilience point). Unfortunately, the body, as compared with any elastic material, has a limit where deformation will start to occur, inducing organ failure. The point where deformation happens is the allostatic overload point. The brain controls the stress response via the central stress system (CSS) based in the diencephalon and brainstem. The CSS has been described as a monitor and regulator of the stress response\(^1\). It receives inputs from within and outside the body through the somato-sensorial system and from cognitive-related areas located in the prefrontal cortex, limbic system and hippocampus. All play a major role in information processing, emotions and behaviour\(^1\).

Changes in biology induced by the stress response are associated with changes in the cognitive domain\(^2\). The internal state can be communicated through changes in facial expressions. Medical decisions using heuristic reasoning (also known as “gut feeling” or intuition) have been made by healthcare professionals looking at the patient’s face to gather information regarding their physiological state\(^3\). Traditional track and trigger systems were developed in the assumption that clinical signs appear early before patients become critically unwell. The first system was created in Liverpool Hospital, Australia as part of an effort to implement a Medical Emergency Team to prevent cardiac arrest in the wards. The original system evaluated (but was not limited to) changes in the physiology, namely consciousness, blood pressure, respiratory rate, pulse rate and temperature\(^4\). Subsequently other research groups advocated the use of different scoring systems to identify patients at risk of
deterioration, among them was the Early Warning Score published as part of an aggregate weighted system\textsuperscript{5}. The National Health System (NHS) in the United Kingdom started to adopt a common system across hospitals in 2012, the National Early Warning Score (NEWS), recently updated\textsuperscript{6}. Perhaps surprisingly, no early warning system utilizes assessment of facial expressions. However, these can be quantified using anatomic-based score systems, such as the Facial Action Coding System (FACS)\textsuperscript{7}. Pain research on patients’ displays has been well developed in neonates using a Neonatal Facial Coding System (NFACS) inspired originally in FACS and a computer enable “point-pair method” by means of anatomical matching references\textsuperscript{8}.

Preliminary studies using FACS suggested that outreach nurses identified mostly sadness and fear in patients at risk of deterioration\textsuperscript{9}. Moreover, facial photographs from the Longitudinal Study of Aging Danish Twins could predict short-term survival and this correlated with cognitive and biological markers such as leukocyte telomere length\textsuperscript{10}. We thus undertook this pilot study on analysing facial expressions in ward patients at risk of deterioration, investigating muscle movements or Action Units (AU in FACS terminology).

This research was presented in part in abstract form at the Intensive Care Society State of the Art Congress 2016, London, United Kingdom\textsuperscript{11}.

**Methods**

This single center descriptive study collected facial data through video sequences to detect AUs and correlate them with physiological states. The study was approved by the UK Health Research Authority (IRAS 165739, REC 16/LO/0365). Patients included in the study were deteriorating patients reviewed by an intensive care outreach nurse at a 420 bed Community Hospital in London, UK. Patients were invited to participate if they were (i) awake (AVPU score
zero) and able to consent in accordance with NHS England Research Ethics Committee regulations (ii) adult (≥18 years old), (iii) had a National Early Warning Score (NEWS) ≥5 or scored 3 points in one single NEWS variable (abnormalities in respiratory rate, oxygen saturation, temperature, systolic blood pressure, heart rate and level of consciousness, and use of oxygen. Total score range between 0-20), or the nurse was concerned about the patient, and (iv) Patients were on room air, oxygen cannula or face-mask.

The minimum frequency of NEWS monitoring in the ward is adjusted to the severity of the patients and hence to the actual value of NEWS score. Patients with low scores can be monitored from 12 to 4-6 hourly and patients with high scores are monitored every hour or even continuously if the aggregated value is above 7 or more.

Patients were excluded if they were <18 years old, receiving sedatives or drugs in doses high enough to interfere with consciousness level, had anatomical face impediments, such as previous facial surgery, prominent beards, and eye abnormalities, or were unable to give written informed consent.

A 5 minute colour video of the patient’s face was taken only after the referral was done from the outreach nurse to the clinical investigator, inclusion/exclusion criteria was ascertained, outreach nurse had left (or was not clinically involved in the surroundings) and the patient’s consent was obtained. The camera EOS 7D with an EFS 18-135 lens (Canon, Town, Japan) was placed on a tripod (Vinten Vision 100, Town, England) placed at the foot of the patient’s bed (91.4 cm width x 203 cm length) to avoid intrusion into their close environment. A lateral position of the camera was used when a better view of the lips and jaw was required in patients with face-mask. Patients were instructed to look forward. The camera recorded a 5 minute video at a speed of 25 frames/sec (7,500 images per patient). AU are captured for
study purposes in videos lasting few seconds^7 and a usual length of recording for research is at around 30 secs^8. The five minutes length of the video was selected to generate enough data for the coding interpretation integrity in a never before attempted setting.

Outcomes were critical care referral rate for the study (number of patients referred to the clinical investigator per total patients assessed by the critical care outreach nurse), approach rate, consent rate, drop-out rate of research participants, proportion of AU identified (in each anatomic area and study group) and admission to intensive care.

The videos were coded and stored under password protection and encryption (Veracrypt™, IDRIX, Paris, France) within a dedicated laptop computer. The patient’s name was anonymized and a code number assigned. Original videos were deleted from the video recorder. All recorded images were analysed frame by frame (7,500 images per patient) by a trained FACS psychologist blinded to outcome. The coder identified AU in Upper Face (UF), Head Position (HP), Eyes Position (EP), Lips and Jaw position (LJ) and Lower Face (LF). Three pre-established rules were used, namely: recording of spontaneous displays and not those related to external stimuli, identifying AU only on first seeing the face (first impression), and selection of AUs that were predominant for each part of the face. For example, eye closure^7 was scored as such if the eyes were closed for more than ½ sec during the 5 min video, but only the dominant AU for each region was considered for the study. A double-coding in two separate sessions for all patients by the trained psychologist was used to ascertain intra-rater variability. The formula used was the FACS reliability formula:  (number of AU on each round the psychologist agreed \( \times 2 \)) divided by the sum of the number of AUs scored in each round^12. Descriptive analysis was undertaken to characterize the study sample, using means and proportions, as appropriate. Measurements were made of sensitivity, specificity, odds ratio for contingency tables,
Pearson chi-squared for categorical comparisons and Student’s t test for numerical comparisons, a binary logistic regression and clustering algorithms for inferential analysis. (Systat, Inc. v13.1, San Jose, CA, USA). Variables are presented as mean ± SD and percentages. As this was a pilot study designed to collect preliminary information on feasibility and effects, formal sample size calculations were not undertaken.

Results

The outreach team reviewed 747 patients during the time enrolment was active. Sixty-three (8.4%) patients were referred for study, and 58 (92%) were approached, thirteen (22.4%) refused to participate while 11 (18.9%) were excluded pre-filming for not meeting the inclusion criteria and 34 (58.6%) gave consent. Once enrolled, the drop-out rate was zero. (Supplemental Figure 1).

The intra-rater reliability index was 0.77 (typical inter-coder agreement for FACS is between 0.82-0.75)\(^\text{12}\).

Mean age of the 34 enrolled patients was 62.6 ± 13.9 years (range 27-83), of whom 21 were female, 24 Caucasian, six Afro-Caribbean and four Asian. The NEWS score was 5.2±3.1. A table of the individual components of the NEWS score is presented in the supplemental digital content (Supplemental Table 1). We did not find significant differences neither in NEWS score nor in its physiological component variables between the ICU and non-ICU group. Nevertheless, 100% patients of the ICU admission group had oxygen compared with 60% from the non-ICU group (p < 0.05). The NEWS-6 (mean for the six previous NEWS determinations) was 5.4±2.6 and a table of the individual physiological components is showed in the supplemental digital content (Supplemental Table 2). This table showed that patients
admitted to ICU had significant higher mean respiratory rate than non-admitted (23.0±7.7 vs 20.5±4.5, p < 0.05).

The Numerical Rate Pain score was 0.3±1.1 (out of 10) during recording the video. There were not significant different pain scores between patients requiring ICU transfer and those that did not (0.18± 0.6 vs 0.45±1.3). During the critical care outreach review seventeen patients had a provisional diagnosis of sepsis and the others had initial diagnosis grouped into nine conditions (ventricular failure, postsurgical intervention, pancreatitis, Crohn’s disease, severe anaemia, pulmonary embolism, pneumothorax, asthma, intra-cardiac thrombus). There were 9 patients without oxygen, 22 patients with oxygen delivered by a nasal cannula and 3 patients with oxygen delivered by a face-mask. 100% patients were scored by the coder irrespective of the use an oxygen device or none.

The most frequently detected AU on the upper face (UF) was AU43 (eyes closed, 73%), for Head Position (HP) AU51 (11.7%, head turned left), and for Eye Position (EP) AU62 (5.8%, eyes turned right). The commonest AU for Lips and Jaw (LJ) was AU25 (44.1%, lips parted) and for Lower Face (LF) AU15 (67.6%, lip corner depressor) (Supplemental Figure 2). There was no association between AU and age, sex, ethnicity or diagnosis.

Demographics of patients who were (n=11) and were not (n=23) admitted to intensive care are presented in Table 1. AUs in both groups are shown in Supplemental Figure 2. There were no significant differences in the proportions of AU in the regions studied in the face (UF, HP, EP, LJ, and LF) between groups. Cluster analysis demonstrated a grouping classification of three AU, namely AU15, AU25 and AU43 (Supplemental digital content Figure 3). Furthermore, manual overlapping of these AU with AU51 or AU52 identified differences on certain combined facial displays (FD) in patients who were/were not admitted to intensive care.
care, namely FD1 [presence of AU 43 (eye closure) + 15 (lip corner depressor) + 25 (lips parted)] \( (p<0.013) \); FD2 [presence of AU43 (eye closure) +15 (lip corner depressor) +51 (head turned left) or 52 (head turned right)] \( (p<0.003) \), and FD3 [(presence of AU43 (eye closure) + 15 (lip corner depressor) + 25 (lips parted) + 51 (head turned left)] \( (p<0.002) \).

Table 2 shows these facial displays (FD1, 2, 3) and their corresponding sensitivity, specificity and odds ratios (OR) related to the need for ICU admission. The presence of FD1, FD2 and FD3 increased the risk (odds ratio) of being admitted to intensive care 8-fold, 18-fold and as a sure event respectively. In accordance with the OR a new score (VIEWS) was created in which the presence of FD1, FD2 and FD3 were weighed as 1, 2 and 3 points, respectively.

The NEWS score did not predict admission to intensive care (c-index 0.57, \( p = 0.443 \)), but a logistic regression model with VIEWS and NEWS as independent covariates could describe admission to intensive care with an average concordance statistic (c-index) of 0.71 \( (p = 0.008) \) (Figure 1).

**Discussion**

The presence of prototypic facial expressions, namely FD1-3, in patients at risk of deterioration was associated with an increased risk of being admitted to intensive care. Furthermore, these facial displays improved the performance of the NEWS score in predicting admissions to intensive care. The overall recruitment rate was slow because of the need to include patients able to consent, not availability of the photographer to record round-the-clock and the inability to recruit during patients’ rest on the ward (midday to 2 pm, 5 – 6 pm and 8 pm-8 am). A study amendment (approved by the Health Research Authority) was implemented to include patients with oxygen (cannula or face mask) to improve referral and
recruitment rates. The feasibility variables highlight the need of additional resources to increase the recruitment rate in future research. Furthermore, new studies should consider, after discussions with their Ethics Committee, to include adults not able to consent for themselves approaching legal representatives, consultees, close relatives or friends. Out of the 34 patients enrolled in the study 32% (11 patients) were admitted into intensive care. This figure is twice the proportion of patients admitted to intensive care after a routine outreach nurse visit in our hospital and reflects the selection criteria of the recruitment process (NEWS \( \geq 5 \)).

Painters have been capturing face expressions since antiquity. The painting 'The Dead Christ Mourned' (Figure 2, left) by Annibale Carracci (1560-1609) is striking in its composition. Carracci showed the same facial expression in the dead Christ and the Madonna, clearly displaying FD3, namely AU15 (lip corner depression), AU43 (eye closure), AU51 (lateral position of head) and AU25 (lips parted). A collection of 20 photographs of patients on Palliative Care taken by the photographer Andrew George at Los Angeles’ Providence Holy Cross Hospital\(^{13}\) display similar facial expressions (Figure 2, centre). The most frequent AUs identified in Andrew George’s group were also AU43 (eye closure, 70%) and AU15 (lip corner depressor, 65%)\(^{11}\) (Supplemental digital content Figure 4). The normal facial expression in hospitalised non-deteriorating patients have not been systematically evaluated. Future studies with a control group of hospitalised non-deteriorating patients is warranted.

Sex, age or ethnicity was not associated with a particular facial expression. This finding is in accordance with Ekman et al\(^{14}\) who stated that facial expressions are universal and transcultural, an observation first proposed by Darwin\(^ {15}\). AU43 (eye closure) and AU15 (lip corner depressor) have been associated with pain and sadness, respectively\(^ {7}\). The prototypic pain face (PPF) is better described using the Prkachin-Salomon formula\(^ {16}\) when brow lowering
orbital tightening (AU6 and AU7), nose wrinkling (AU9) or upper lip raised (AU10), and eye closure (AU43) are present. The prototypic sadness face (PSF) corresponds to the combination of three muscles, namely inner brow raiser (AU1), brow lowerer (AU4) and lip corner depressor (AU15)⁷. Of note, neither PPF nor PSF were clearly seen in patients at risk of deterioration who later required admission to intensive care. Furthermore, pain scores were low overall and not significant different between patients requiring ICU transfer and those that did not. These ratings could be different in samples with a distinct case-mix.

A combination of facial displays (FD1-3) was identified instead. These displays could be named ‘prototypic serious illness faces’ (PSIF) (Figure 2, right with FD3). Noteworthy, eye closure might be related with primitive reflexes of protection as it happens in the startle response or painful stimuli. Continued work with adequate sample size and quantitative analyses are needed to support the conclusions for a proposed prototypic facial expression of deterioration and its association with admission to intensive care.

A serious illness will trigger a stress response with biological and psychological dimensions (Figure 3). Appraisal¹⁷ of the situation will activate changes in emotional state and facial displays. Facial displays (face expressions) are most likely a paralanguage – i.e. a way to communicate emotions or needs¹⁸. Furthermore, the CSS may be able to elicit reflexes as part of the biological response. Some of these reflexes will also be displayed in the face, for example associated with systemic responses (e.g. the startle response) or directly displayed for a few fractions of a second (also known as micro-expressions)¹⁴. Connections between the facial nerve nuclei with the reticular system and hypothalamus explain the link between the emotional and limbic systems and facial displays¹⁹. FACS uses anatomic landmarks to describe all possible expressions, thus it should be able to report any changes in facial displays,
regardless of whether their origin was coming from emotional states, communication needs or reflexes.

We have previously demonstrated that patients on Palliative Care and patients at risk of deterioration showed similar facial displays, with a predominance of AU43 and AU15. According to the General Unified Theory of Stress there is a point (resilience, $\rho$) where the stress response will be elicited by the injury. We can assume that $\rho$ has been reached by both palliative and acutely ill patients. The similar AU found in the Andrew George group and patients at risk of deterioration emphasises the need to value facial expressions as an important clue of distress, enhancing the potential utility of new scores that include visual clues for patients in hospital wards.

Decision-making in critical care requires a cognitive process. In acute care management, visualization is a main source of this information. Remarkably, visual information derived from facial expression has received little attention yet is integral to initial assessment. No early warning system score, implemented to identify early signs of clinical deterioration, includes any facial information. Complications in intensive care are not always deduced by the likelihood of risks assessed by current scoring systems. For example, nurses who rely on visual assessments are more likely to be alerted by their intuitive judgment. Intuition is recognised as a key element in how decisions are made by critical care nurses, and the patient’s face can offer important clues. Subjective feelings, and not just objective measures used in early warning scores, should be thus taken into consideration.

This study is a pilot study to assess feasibility and is underpowered to confirm a relationship between the presence of a particular AU and subsequent deterioration. Nevertheless, we are encouraged by the increased risk found between some facial displays.
and subsequent deterioration with consequent admission to intensive care. This finding is supported by the coincidental evidence of similar AUs in terminally ill patients and by the representation of stress in paintings by old masters.

Further investigation is required to confirm the predictive value of facial displays in conjunction with traditional early warning scores. A newly emerging technique in this field is automatic face feature detection and recognition by advanced computer vision algorithms. Continuous video recording of faces with appropriate automatic software to identify triggers of concern could be a reality in the near future.

**Conclusion**

Patterned facial expressions can be identified in deteriorating general ward patients. This tool may potentially augment risk prediction of current scoring systems.

**Acknowledgements**

The Critical Care Outreach team at North Middlesex Hospital for its invaluable help during the study
REFERENCES


Table legends

**Table 1:** Demographic data in full sample by admission or not to intensive care. Figures are means and standard deviation unless stated otherwise. A= Asian, AC=Afro-Caribbean, C = Caucasian. S = Sepsis. NS = Non-sepsis diagnosis.

**Table 2:** Sensitivity, Specificity and Odds Ratio and level of significance for facial displays (FD) 1 to 3 in regard to ICU admission
Figure legends

Figure 1. Logistic regression to ascertain the risk of being admitted to intensive care when VIEWS and NEWS were used as covariates.

Figure 2. Annibale Carracci painting (left, The Dead Christ Mourned -'The Three Maries'- © The National Gallery, London. Presented by Rosalind Countess of Carlisle, 1913), a patient at risk of deterioration of the study (right) and a patient deemed to die (centre, Andrew George©. Right, before I die. [Online] www.rightbeforeidie.com).

Figure 3. The stressor (physical, psychological or social) elicits a stress response with biological and cognitive dimensions that are both associated with facial displays.
Supplemental Digital Content

Supplemental Table 1. Individual physiological components of the NEWS for groups. *p < 0.05

Supplemental Table 2. Individual physiological components of the NEWS-6. N = total observations available per group in the ward chart. * p < 0.05.

Supplemental Figure 1. Flow diagram for the progress of the study

Supplemental Figure 2. Distribution of AU in Upper Face (UF), HP (Head Position), Eyes Position (EP), Lips and Jaw (LP) and Lower Face (LF) in ICU (red) and non-ICU (blue) admission groups. No statistical significance between groups was seen. The total count (N) of all AU identified is represented in black.

Supplemental Figure 3. Hierarchical cluster tree (dendrogram) for dominant AU (N = 34).

Supplemental Figure 4. Distribution of Action Units (AU) in a collection of 20 photographs by Andrew George of patients on palliative care.