Skin toxicity evaluation in patients treated with cetuximab for metastatic colorectal cancer: a new tool for more accurate comprehension of quality of life impacts

Michele De Tursi¹
Marinilla Zilli¹
Consiglia Carella¹
Matteo Auriero²
Maria Nadia Lisco²
Marta Di Nicola³
Giuseppe Di Martino⁴
Clara Natoli¹
Paolo Amerio²

¹Department of Medical, Oral and Biotechnological Sciences, Medical Oncology Unit, ²Department of Dermatology and Venereology, ³Department of Medical, Oral and Biotechnological Sciences, ⁴Department of Medicine and Aging Science, University degli Studi “G d’Annunzio” Chieti-Pescara, Chieti, Italy

Introduction

Epidermal growth factor receptor (EGFR) is a transmembrane tyrosine kinase receptor found on cells of epithelial origin, which has an important role in cell differentiation, proliferation, migration, apoptosis, angiogenesis, and cytokine regulation. EGFR signaling is commonly activated in various solid tumors, including colorectal cancer (CRC), and is associated with tumor progression and poor prognosis. Agents targeting EGFR are collectively described as EGFR inhibitors (EGFRIs).
Cetuximab and panitumumab, monoclonal antibody (mAb) EGFRIs, have proved efficient, both as single agents and in combination with chemotherapy, in the treatment of metastatic CRC without mutations in the RAS gene.\textsuperscript{3,4} The majority of patients treated with EGFR mAbs and also those treated with small molecule inhibitors of the EGFR tyrosine kinase domain, such as gefitinib and erlotinib, experience dermatological side effects.\textsuperscript{5,6} EGFR is highly expressed in the epidermis, especially in the basal cell layer, and in the epithelium of hair follicles. Although papulopustular skin rash is the most common skin toxicity associated with anti-EGFR mAbs, other cutaneous side effects are also observed, including xerosis, fissures, pruritus, paronychia, and blepharitis.\textsuperscript{7,8}

Cutaneous toxicity is rarely life-threatening; however, it can deeply impact quality of life (QoL) affecting the emotional, psychosocial, and physical well-being of patients. It is important to correctly assess the grade of cutaneous adverse events, since oncological treatment is often modulated based on their severity. Moreover, the adherence of patients to therapy may be severely affected; it has been reported that up to 30\% of patients stop therapy due to cutaneous adverse events.\textsuperscript{9,10} Interestingly, the occurrence and severity of skin toxicity is associated with improved clinical outcome in patients receiving EGFRIs.\textsuperscript{11,12} Therefore, optimal management of skin toxicity is crucial to maintain patient adherence and avoid treatment delay or interruption.

One of the problems hindering the effective management of EGFR-associated dermatological toxicities is the use of inaccurate and inconsistent toxicity evaluation criteria.\textsuperscript{13} Previously available tools were not designed for reporting EGFR-associated dermatological events, and this resulted in underreporting and poor grading of side effects. The National Cancer Institute’s Common Terminology Criteria for Adverse Events (NCI-CTCAE) was developed as a standardized method for use in oncology clinical trials to document and grade toxic effects of anticancer therapies, including dermatological adverse events.\textsuperscript{14} The most recent version (NCI-CTCAE v4.0) was published in 2009 and attempted to fill the gaps in the previous, more generic, criteria pertaining to rash, dry skin, and nail changes, by revisions and by grading these features separately.

The assessment and evaluation of the severity of cutaneous involvement in patients with drug-induced acneiform eruption (dAE) are controversial. Patients and physicians often disagree on the severity of dAEs, leading to inconsistencies in reporting of the grade of severity.\textsuperscript{15} Part of the difficulty could lie in the recognition and reporting of cutaneous signs, with which oncologists may not be familiar. Moreover, CTAE grading systems describe symptoms individually, while several symptoms and signs are often present simultaneously, with their combination contributing to the discomfort of the patient. In addition, it is difficult for healthcare providers to objectively measure the effect of a particular dAE on the health-related quality of life (HRQoL) of a patient. Therefore, it is crucial to develop a strategy to capture the patient’s understanding of the severity of dAEs and their effects on HRQoL.

It is sometimes difficult for both healthcare providers and patients to assess and communicate skin conditions, since oncologists are not always familiar with cutaneous lesions and their presentation, while patients do not always adequately express the distress associated with their symptoms. Monitoring, recognition, and early treatment may help to relieve the distress caused by symptoms and raise the adherence of patients to treatment, improving their QoL. The acneiform rash section of NCI-CTCAE v4.0 states that psychological impacts should be considered; however, it does not clearly define how to report these. Moreover, it does not take into account the patient’s subjective discomfort. Thus, underreporting the severity of skin toxicities may lead to under-adjustment, or inappropriate discontinuation, of anti-neoplastic therapy.

Recently, a measure of standardized patient-reported outcomes specific to EGFR-induced dermatological side effects was proposed, since severity scales, such as NCI-CTCAE v4.0, do not adequately assess the impact of EGFR on QoL.\textsuperscript{16} However, alternative grading tools, such as FACT-EGFRI-18 and DIEHL-24, require further evaluation. One alternative grading scale, specific for EGFR-associated skin toxicity, is the Multinational Association of Supportive Care in Cancer (MASCC) EGFR Inhibitor Skin Toxicity Tool (MESTT).\textsuperscript{17} Compared with the NCI-CTCAE scale, MESTT tends to report higher toxicity grades for some side effects, such as rash, xerosis, and paronychia.\textsuperscript{18} Although MESTT allows for more detailed reporting of cutaneous toxicity, there is evidence that it may not be able to efficiently capture the global impact of different cutaneous adverse events on the QoL of patients.\textsuperscript{19}

We recently constructed a composite tool, the Eruption Scoring System (ESS), the concept for which was derived from different acne scoring systems and incorporates several aspects of EGFI adverse events including rash, blepharitis, and paronychia, combined into subjective symptoms, such as pruritus and burning. Final scores generated by the ESS represent an assessment of severity of cutaneous involvement.
This study was designed to assess the correlation between QoL and severity of dermatological toxicity evaluated using the NCI-CTCAE v4.0 and our novel scale (ESS) in metastatic CRC patients treated with first-line chemotherapy combined with cetuximab. We also evaluated the differences between the assessments of cutaneous adverse events by oncologists and dermatologists using both tools.

**Patients and methods**

**Patient population**

Patients with histologically confirmed metastatic CRC, untreated with systemic therapy for metastatic disease (prior adjuvant chemotherapy for CRC was permitted), who were scheduled to start first-line chemotherapy (oxaliplatin- or irinotecan-based chemotherapy) combined with cetuximab, were eligible for inclusion in this study.

Other eligibility criteria included patients aged ≥18 years, an Eastern Cooperative Oncology Group performance status of 0–2, adequate hematologic, renal metabolic, and hepatic function, a life expectancy >3 months, and the ability to answer the QoL questionnaires.

**Ethics and consent**

This study was conducted in accordance with the guidelines from the International Conference on Harmonization-Good Clinical Practice and the Declaration of Helsinki. The study protocol was approved by the Ethical Committee of the University “G d’Annunzio”, Chieti-Pescara, Italy. All patients provided written informed consent before participating in the study.

**Procedures**

Cetuximab was administered on a weekly schedule, with a 400 mg/m² initial loading dose, followed by 250 mg/m² weekly infusions until disease progression. Chemotherapy was given according to the standard clinical practice at the investigator site.

Patients underwent clinical investigation, including skin toxicity assessment, at baseline (before starting antineoplastic treatment) and at every subsequent visit. Physical examinations were performed separately by a medical oncologist and a dermatologist. Cutaneous toxicity was evaluated using both the NCI-CTCAE v4.0 and ESS scales. Patients were monitored according to the study protocol until the scheduled date of study completion or up to recovery or stabilization of a followed-up cutaneous adverse event, whichever came last. The patient or physician could decide to withdraw from the study at any time for any reason. The QoL of each patient was assessed by two dermatological specific QoL questionnaires: Skindex-29 and its brief form, Skindex-17.

Skindex-29 and Skindex-17 are among the most commonly used dermatology-specific HRQoL instruments. Skindex-29 consists of 29 items, covering burden of symptoms, functioning, and emotional spheres. The questions refer to the previous 4-week period, and scores are provided on a five-point scale, from “never” to “all the time”. Greater scores indicate poorer QoL. Skindex-17, which is derived from Skindex-29, consists of 17, rather than 29, items, with answers given on a three-point, instead of a five-point, scale. These two questionnaires have demonstrated good correlation with other grading tools and with NCI-CTCAE. Moreover, although they are not specifically designed for targeted therapy-related cutaneous rashes, they were appropriate for the assessment of HRQoL in these patients.

No enrolled patients were permitted to receive any prophylactic skin toxicity treatment; however, all patients received treatment after any kind of cutaneous toxicity appeared. Skin treatments were chosen at the discretion of the investigator, according to the toxicity severity, beginning with mild-grade events. The primary objective of this study was to determine the correlation between patient QoL and skin toxicity severity, assessed using the NCI-CTCAE v4.0 and ESS grading systems. Secondary objectives included 1) determination of the incidence rates of skin toxicities and 2) evaluation of interinvestigator concordance rates in skin toxicity evaluation using NCI-CTCAE v4.0 (rash section) and the ESS score systems.

NCI-CTCAE v4.0 allows separate assessment of several aspects of targeted therapy-related adverse events, usually using five grades of severity; for example, the severity of a skin rash is established on the basis of clinical presentation and the reported distress of the patient. However, it has several limitations that may interfere with the optimal assessment of patients. First, dermatological presentations are very specific and their definitions are not always clear; second, the majority of the time patients present with multiple symptoms (eg, rash and pruritus or paronychia); hence, a classification of severity of each symptom may be misleading in global assessment of the patient; finally, the different grades are based on skin surface involvement, along with the development of psychosocial impact; however, these are not precise parameters, and this can lead to difficulties of interpretation.

The ESS is a composite objective and subjective scoring system, which considers the type of skin lesions, their localization, the presence of paronychia and blepharitis, and
takes into account self-reported symptoms, such as pruritus and burning.

ESS score

i) Local score

Four types of skin lesions have been described (Figure 1):

- Erythematous macule is defined as a flat, nonpalpable lesion that varies in pigmentation from the surrounding skin.
- Papule is defined as a circumscribed, solid elevated lesion, usually <1 cm in diameter.
- Pustule is defined as a small elevation of the skin containing purulent material.
- Nodule is defined as a circumscribed, solid elevation of the skin, usually >1 cm in diameter, extending in the dermis or subcutaneous fat. The depth of involvement is what differentiates a nodule from a papule.

Each type of a skin lesion is scored 0–4 as reported below:

<table>
<thead>
<tr>
<th>Score</th>
<th>Lesion Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No or one skin lesion</td>
</tr>
<tr>
<td>1</td>
<td>&gt;1 erythematous macule</td>
</tr>
<tr>
<td>2</td>
<td>&gt;1 papule</td>
</tr>
<tr>
<td>3</td>
<td>&gt;1 pustule</td>
</tr>
<tr>
<td>4</td>
<td>&gt;1 nodule</td>
</tr>
</tbody>
</table>

The maculo-papulopustular eruption appears with different frequencies on different anatomical sites. We supposed that, as it happens in other dermatological diseases such as psoriasis, the most socially visible areas may have a greater impact on psychological distress. Therefore, we arbitrarily decided to divide the body into seven different zones, with specific multiplication factors (indicating the most visible sites), as follows:

- I Forehead 2
- II Cheek 3
- III Chin 2
- IV Upper chest 2
- V Lower chest 1
- VI Upper back 2
- VII Lower back 1

Every multiplication factor will be multiplied by 0, 1, 2, 3, or 4, according to the particular type of skin lesion observed in the corresponding body area. The local score will be obtained by adding each of them, as follows: I+II+III+IV+V+VI+VII.

ii) Pruritus and Burning Visual Analogue Scale (VASP and VASB)

All patients assessed pruritus and burning intensity using the 10-point VAS.

iii) Blepharitis and paronychia

Inflammation of the eyelid and infection of the folds of tissue surrounding the nail of a finger are graded as follows:

<table>
<thead>
<tr>
<th>Blepharitis (Bs)</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild</td>
<td>+1</td>
</tr>
<tr>
<td>Moderate</td>
<td>+5</td>
</tr>
<tr>
<td>Severe</td>
<td>+7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Paronychia (Ps)</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild</td>
<td>+1</td>
</tr>
<tr>
<td>Moderate</td>
<td>+5</td>
</tr>
<tr>
<td>Severe</td>
<td>+7</td>
</tr>
</tbody>
</table>

Figure 1 Skin lesion score according to the Eruption Score System.
Final score:
The final score will define the severity of cutaneous involvement, employing the added score from the above:

Final score = (Local score I+II+III+IV+V+VI+VII) + VASP+VASB+Bs+Ps

The final score identifies the following categories of skin toxicity:

1  No toxicity
1–10 Very mild
16 Mild
30 Moderate
30–40 Severe
>41 Very severe

Sample size and statistical analyses
In reliability studies, the main aim is not to test but to generate accurate estimates. For this reason, sample size calculations focused on precision, rather than power. The minimum required sample size (n=29) for the calculation of a correlation coefficient was determined based on predicted $r=0.5$, with a 95% confidence interval (CI) and an alpha error rate of 5%. Qualitative variables are summarized as frequencies and percentages. Quantitative variables are summarized as means and standard deviation or medians and interquartile range, according to their distribution.

For ESS and NCI-CTCAE v4.0 assessments, median and IQR values were calculated for each measurement for every patient, and for each observer. Interoperator agreement for ESS and NCI-CTCAE v4.0 scoring was assessed by calculating interclass correlation coefficients (ICC) using a two-way random-effects mixed model. ICC values (based on 95% CIs) <0.5, 0.5–0.75, 0.75–0.9, and >0.90 are indicative of poor, moderate, good, and excellent reliability, respectively. Due to significant interobserver agreement, the mean values of all measurements by all observers using the ESS and NCI-CTCAE v4.0 systems were calculated for each patient.

For ESS and NCI-CTCAE v4.0 systems were calculated for each patient. Spearman’s rank tests were used to evaluate the correlation among mean ESS and mean NCI-CTCAE v4.0, and Skinindex-29 and Skinindex-17, scores. Analysis of the correlation between ESS and NCI-CTCAE v4.0 scores was also performed. All statistical tests were evaluated at an alpha level of 0.05. Statistical analysis was performed using SPSS® Advanced Statistical 11.0 software (SPSS, Inc., Chicago, IL).

Results
Study population
From May 2013 through May 2015, a total of 30 patients were enrolled in this study. The baseline demographic and clinical characteristics of the patients are reported in Table 1. Twenty patients (66.7%) were males. All enrolled patients were Caucasian, the median age was 59 years (range, 36–77 years), and 73.3% of patients had colon cancer and 26.7% had rectal cancer. Eleven patients (36.6%) had a medical history of rosacea, associated with an increased risk of cutaneous toxicity.

Primary end point
Five medical oncologists (three attending physicians and two resident physicians) and six dermatologists (three attending physicians and three resident physicians) evaluated all 30 patients enrolled in the study. Photographic documentation of each patient was available at baseline and during the treatment period.

A positive correlation was found between dermatology HRQoL using the Skindex-29 tool and the severity of dermatological toxicities, assessed using the NCI-CTCAE v4.0 scale for cutaneous papulopustular acneiform rash ($r=0.483$, $P=0.001$); however, a stronger correlation was documented between HRQoL and the ESS tool ($r=0.611$, $P=0.001$; Figure 2).

Similar results were observed using the Skindex-17 tool, including its psychosocial and symptom domains ($r=0.564$, $P=0.001$ for ESS vs $r=0.425$, $P=0.001$ for NCI-CTCAE v4.0; Figure 3).

Secondary end points
Whole skin toxicity occurred in 90% of patients and blepharitis occurred in 50% of patients, whereas hand skin toxicity occurred in 56%. The pruritus and/or burning were strong in 43% of patients and very strong in 13%.

Table 1 Baseline demographic and clinical characteristics of patients enrolled in the study

<table>
<thead>
<tr>
<th>Variables</th>
<th>Age (years), median (range)</th>
<th>Race, n (%)</th>
<th>Gender, n (%)</th>
<th>Primitive cancer site, n (%)</th>
<th>History of rosacea, n (%)</th>
<th>Skindex-17, median (IQR)</th>
<th>Skindex-29, median (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), median (range)</td>
<td>59.6 (36.0–77.0)</td>
<td>Caucasian</td>
<td>Male</td>
<td>Colon</td>
<td>Yes</td>
<td>56.7 (20.0–92.5)</td>
<td>68.1 (53.1–127.1)</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
<td>30 (100.0)</td>
<td>20 (66.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
<td>Caucasian</td>
<td>Female</td>
<td>Colon</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>30 (100.0)</td>
<td>10 (33.3)</td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primitive cancer site, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>Colon</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>History of rosacea, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Abbreviation: IQR, interquartile range.
Both NCI-CTCAE v4.0 and ESS tools demonstrated good interobserver agreement for grading of skin toxicity, with ICC coefficients of 0.729 (95% CI: 0.648–0.808; P<0.001) for NCI-CTCAE v4.0 and 0.925 (95% CI: 0.895–0.950; P<0.001) for ESS (Table 2), indicating a moderate interobserver agreement for NCI-CTCAE and an excellent agreement for ESS.

The statistical model also demonstrated an estimation of the effect linked to the observer, indicating the bias among
observers. This estimation was statistically significant ($P=0.002$) for NCI-CTCAE, indicating that measurements among observers are not reproducible. While for ESS, the estimation of the effect, indicating the bias between observers, is not statistically significant ($P=0.239$), indicating that the measurements done by the observers are reproducible for this method (Table 2).

In addition, a significant correlation was found between the two skin toxicity grading systems (NCI-CTCAE v4.0 and ESS; Figure 4).

The median duration of chemotherapy was 8 months (range, 3–29 months). Twenty-six percent of patients received anti-EGFR treatment with cetuximab for >12 months, whereas 40% of them received it for <6 months. No patients discontinued therapy nor reduced their dose of therapy due to toxicity.

**Discussion**

In this study, a strong correlation was identified between our novel ESS and the standard NCI-CTCAE v4.0 tool in grading cutaneous toxicity associated with the anti-EGFR mAb cetuximab. ESS can, therefore, be considered a valid instrument to identify and grade the severity of skin toxicity induced by cetuximab, with some advantages over the classical NCI-CTCAE v4.0 scoring system. First, unlike the standard NCI-CTCAE v4.0 grading system, the ESS tool allows for a more accurate evaluation of skin toxicity with crucial implications for clinical management of patients, taking into account all components of cutaneous adverse events and not evaluating them each individually. This may also explain the higher level of correlation found between dermatological HRQoL and the ESS, compared with the NCI-CTCAE v4.0, scoring system. The ESS score allows a better evaluation of the consequences of cutaneous toxicity on the lives of patients similar to other grading systems, such as FACT-EGFRI-18. Moreover, the greater emphasis of the ESS on psychosocial and symptomatic aspects of patient’s lives may lead to a greater level of empathy in the patient–physician relationship, with improved patient compliance with treatment.

Second, the ESS allows physicians to perform a more objective evaluation of dermatological toxicity induced by cetuximab, compared with the use of NCI-CTCAE v4.0. In particular, the specific and accurate assessment of each manifestation of cutaneous toxicity may decrease the impact of the experience of the physician in comprehensive evaluation. In fact, we found a strong interobserver correlation in grading the severity of skin toxicity associated with cetuximab, despite the different education and experience of the physicians involved in the study; only one young medical oncologist presented a very low ESS score (observer 7 in Table 2). This observer is an oncologist resident and this discrepancy might be explained by his inexperience. However, this occurrence did not alter the results of the interobserver reproducibility.

Although the ESS does not allow for a very rapid estimation of the severity of skin toxicity reactions, with a consequent potential negative impact on clinical practice, it does induce a deeper physician–patient relationship, resulting in improved clinical management of dermatological adverse events and improved patient compliance. We did not measure the time needed to complete the score. However, after a minor

| Table 2 Median and IQR of each observer for NCI-CTCAE v4.0 and ESS |
|------------------|------------------|------------------|------------------|------------------|------------------|
|                  | Dermatologists   | Oncologists      |                  |
|                  | Obs #1 | Obs #2 | Obs #3 | Obs #4 | Obs #5 | Obs #6 | Obs #7 | Obs #8 | Obs #9 | Obs #10 | Obs #11 |
| NCI-CTCAE v4.0   |         |         |         |         |         |         |         |         |         |         |         |
| Median           | 2.0     | 1.0     | 2.0     | 1.0     | 1.0     | 1.0     | 1.0     | 2.0     | 2.0     | 2.0     | 2.0     |
| IQR              | 0–2.0   | 0–2.0   | 0–3.0   | 1.0–2.0 | 1.0–2.0 | 1.0–2.0 | 0–2.0   | 1.0–2.0 | 1.0–2.0 | 1.0–3.0 | 0–3.0   |
| ESS              |         |         |         |         |         |         |         |         |         |         |         |
| Median           | 27.0    | 20.0    | 29.0    | 29.0    | 27.0    | 27.0    | 15.0    | 28.0    | 24.0    | 26.0    | 25.0    |
| IQR              | 0–35.0  | 0–36.0  | 5.0–43.0| 0–36.0  | 10.0–41.0| 3.0–42.0| 0–28.0  | 0–37.0  | 0–34.0  | 0–36.0  | 0–37.0  |

**Abbreviations**: ESS, Eruption Scoring System; IQR, interquartile range; NCI-CTCAE, National Cancer Institute’s Common Terminology Criteria for Adverse Events; Obs, Observer.
training in recognizing the different cutaneous lesions, the mean time to finalize the score was ~2 minutes per patient. Moreover, the simple structure of the ESS, in which the physician is asked to describe one lesion at a time, may increase the reporting accuracy of physicians that are not accustomed to describing cutaneous lesions and allows the completion of the clinical assessment in a few minutes.

We believe that the high concordance of ESS with HRQoL measurements such as Skindex-17 and Skindex-29 may improve the QoL assessment of patients. In addition, the ESS tool may become available as an app on smartphones and tablets, thereby overcoming its limitations, in terms of requirement for time spent on complex calculations. We are actually planning to add photo samples to improve the recognition and to fasten the ESS completion through a computer data sheet that automatically computes the score.

Conclusion

Here we show that the numerous limitations that the NCICTCAE grading system presents when applied to anti-EGFR rash may be overcome by the use of our grading system. Finally, it will be interesting to investigate the usefulness of the ESS for recognition and grading of the cutaneous toxicity associated with other EGFRIs, such as panitumumab, or the tyrosine kinase inhibitors, erlotinib and gefitinib.

Acknowledgment

This work was supported by the Consorzio Interuniversitario Nazionale per la Bio-Oncologia (CINBO).

Disclosure

The authors report no conflicts of interest in this work.

References
