Is speed of healing a good predictor of eventual healing of pyoderma gangrenosum?

**Short title:** Predictors of healing for pyoderma gangrenosum

Sally R. Wilkes PhD¹, Hywel C Williams DSc¹ Anthony D Ormerod MD², Fiona E Craig³, Nicola Greenlaw⁴, John Norrie⁵, Eleanor J. Mitchell BA (Hons)⁶, James M Mason PhD⁷, and Kim S Thomas PhD¹* on behalf of the UK Dermatology Clinical Trials Network’s STOP GAP team.

¹Centre of Evidence Based Dermatology, University of Nottingham, NG7 2NR, UK
²Division of Applied Medicine, Aberdeen University, Aberdeen, AB24 2ZD, UK
³Department of Dermatology, Aberdeen Royal Infirmary, Aberdeen, AB24 2ZD, UK
⁴Robertson Centre for Biostatistics, University of Glasgow, Glasgow, G12 8QQ, UK
⁵Centre for Healthcare Randomised Trials (CHaRT), Aberdeen University, Aberdeen, G12 8QQ, UK
⁶Nottingham Clinical Trials Unit, University of Nottingham, Nottingham, NG7 2UH, UK
⁷Warwick Medical School, University of Warwick, Coventry, CV4 7AL

*Corresponding author:
Professor Kim Thomas
Professor of Applied Dermatology Research and Co-Director of the Centre of Evidence Based Dermatology
Division of Rheumatology, Orthopaedics and Dermatology
School of Medicine
University of Nottingham
Room A103, King’s Meadow Campus, Lenton Lane
Nottingham, NG7 2NR
t: +44 (0) 115 84 68632
f: +44 (0) 115 84 68618
Kim.Thomas@nottingham.ac.uk

Conflict of Interest Disclosure: None Declared.

Funding:
The STOP GAP trial and STOP GAP cohort study were funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research funding scheme (RP-PG-0407-10177).

Word counts:
Abstract – 189 words
Capsule summary – 50 words
Text – 2098 words
Number of Tables - 3
Number of Figures - 2
ABSTRACT

Background

Pyoderma gangrenosum is a rare inflammatory skin condition. The STOPGAP studies compared treatments for pyoderma gangrenosum using a primary outcome of healing speed at 6 weeks.

Objective

Using data from both studies we assessed the predictive value of three early predictors for healing at 6 months - speed of healing, Investigator Global Assessment and resolution of inflammation, recorded at 2 and 6 weeks.

Methods

Logistic regression models were applied and the effectiveness of the three measures was assessed through estimating the positive (PPV) and negative predictive values (NPV) and the area under the receiver operating characteristic curve (AUC).

Results

The PPV and NPV at 6 weeks were 63.5% (95% CI:52.4%, 73.7%) and 74.6% (95% CI:62.5%, 84.5%) respectively for speed of healing; 80% (95% CI:68.7%, 88.6%) and 74.2% (95% CI:64.1%, 82.7%) for IGA; and 72.1% (95% CI:59.9%, 82.3%) and 68.1% (95% CI:57.7%, 77.3%) for resolution of inflammation. Investigator Global Assessment had the best combined PPV, NPV and AUC at 2 and 6 weeks.

Limitations

We were limited by data available from the STOP GAP trial and cohort study.

Conclusion

Speed of healing, Investigator Global Assessment and resolution of inflammation were all shown to be good predictors of eventual healing.
KEY WORDS
pyoderma gangrenosum; speed of healing; lesion improvement; resolution of inflammation; predictors; clinical trials; clinical practice

CAPSULE SUMMARY
- Speed of healing has been shown to be a good predictor of eventual healing for leg ulcers.
- Here, speed of healing, Investigator Global Assessment and resolution of inflammation are all good predictors of eventual healing for pyoderma gangrenosum.
- This finding is helpful for informing future trial design and clinical decision-making.
ABBREVIATIONS

STOP GAP - Study of Treatments fOr Pyoderma GAngrenosum Patients
RCT – randomised controlled trial
PPV – positive predictive value
NPV - negative predictive values
AUC - area under the receiver operating characteristic curve
INTRODUCTION

Pyoderma gangrenosum is a rare inflammatory skin condition that causes tissue to become necrotic, leaving deep ulcerative lesions. These ulcers can be painful, rapidly spread, and may take many months to heal.\(^1\) There is a paucity of evidence for pyoderma gangrenosum treatments.\(^2\) Most evidence is based on observational studies and only two randomised controlled trials (RCTs) have been conducted to date.\(^1,3\) One of the challenges of conducting research into rare skin conditions such as pyoderma gangrenous, is the lack of validated outcome measures for assessing treatment response.

The primary outcome for two recently completed studies (STOP GAP randomised controlled trial\(^3\) and STOP GAP prospective cohort study\(^4\) was speed of healing over the first 6 weeks of treatment. Initial treatment response was used as a surrogate measure for time to healing; which is more clinically-relevant in that it influences patient satisfaction, cumulative drug exposure and drug safety.

Speed of healing, if valid, could become a useful surrogates for eventual healing and could be used to guide early treatment decisions in clinical practice.

Although speed of healing has been shown to be a good predictor of healing in patients with leg ulcers caused by venous disease\(^5,6\), it is unclear whether the same applies to patients with an inflammatory condition such as pyoderma gangrenosum.

Using data from the STOP GAP trial and cohort study, we investigated whether speed of healing in the first 6 weeks of treatment was a good indicator of subsequent healing in patients with pyoderma gangrenosum, or whether other measures, such as Investigator Global Assessment for lesion improvement, or resolution of inflammation, were more useful.
METHODS

This work involved secondary data from previous studies and as such did not require specific approval from an Institutional Review Board.

Study conduct

Ethics and regulatory approvals were obtained for the STOP GAP trial and cohort studies (ethics: 09/H0903/5, Medicines and Healthcare Products Regulatory Agency: 19162/0213/001); all participants gave written informed consent. Oversight of the study was performed by independent Trial Steering Committee and Data Monitoring Committee.

Specific ethical approval for this study was not required.

Summary of the STOP GAP trial and STOP GAP cohort study

Both the RCT and the cohort study included adults with a clinical diagnosis of pyoderma gangrenosum (as confirmed by a dermatologist, with biopsy as required), and followed participants for a maximum of 6-months. For the STOP GAP trial, participants were randomised to receive either ciclosporin or prednisolone, and in the cohort study, participants received topical therapy according to local practice (49 / 74% received clobetasol propionate 0.05%).

For participants with multiple lesions, a target lesion was chosen for study. This was defined as being the largest lesion on a single plane (i.e. not around the curvature of a limb). Lesions were measured by physical measurements taken by the clinician. Grade for lesion improvement was also measured by the clinician using an Investigator Global Assessment (IGA) and resolution of inflammation was measured using the scale reported by Foss. Details of each of these scales are given in Supplementary File 1.

For patients participating in the RCT, lesion size, grade for lesion improvement (IGA) and resolution of inflammation were also assessed by an independent assessor using digital
images. For lesion size the measurements were taken from the digital images using VEV computerised planimetry. An example of measurements being taken from a digital image is shown in Supplementary File 2. These measurements were used in the analyses of the primary and secondary outcomes in the RCT. Where digital images were not available or were of poor quality, the physical measurements recorded by the clinician were used instead. These physical measurements approximated lesion area through the formula: length x width x 0.785.

Outcomes were captured at baseline, 2 weeks, 6 weeks and when the ulcer had healed (up to a maximum of 6 months). Lesions were considered to have healed when sterile dressings were no longer required as reported by patients. If this information was missing, then healing as confirmed by a clinician at the next clinic visit was used instead. Further details of the STOPGAP trial and cohort study are described elsewhere.3, 4

**Patient populations**

The sample size for this study was based on available data. We analysed data from 112 patients who participated in the STOP GAP trial4 and 67 patients from the cohort study.4

**Methods for assessing predictors of healing**

We assessed three possible early indicators for healing or non-healing by 6 months. The first, speed of healing at 2 and 6 weeks, was estimated as follows:

\[
\text{Speed of healing} = \frac{\text{Lesion area at 2 or 6 weeks} - \text{Lesion area at baseline}}{\text{Time between visits (~2 or 6 weeks)}}
\]

Investigator Global Assessment as reported by the clinician at 2 weeks and 6 weeks, as well as resolution of inflammation using the scale reported by Foss7 at 2 weeks and 6 weeks were
also considered as possible early indicators for healing or non-healing by 6 months.

Investigator Global Assessment was treated as a categorical variable (1 "Completely/almost clear", 2 "Marked improvement", 3 "Moderate improvement", 4 "Slight improvement", 5 "No change/worse"). Resolution of inflammation was treated as a binary variable (successful/not successful), with success defined as erythema and border elevation reduced to “none”.

Healing status by 6 months was treated as a binary outcome; healed or not healed. Logistic regression models were used to test the effectiveness of each of the three measures as indicators for healing or non-healing by 6 months. The models were adjusted for age, gender, baseline lesion area, underlying systemic disease and lesion location.

A logistic regression model was fitted in order to estimate the positive (PPV) and negative predictive values (NPV) along with the area under the receiver operating characteristic curve (AUC). The cut-off point for predicted probabilities was set at 0.5. An AUC value of 0.5 demonstrates that the measures are non-predictive of healing or non-healing and a value of 1 is be considered a perfect prediction (i.e. the measures discriminate perfectly between those who heal and those who don’t heal).^8

In terms of lesion area, for the purposes of this study, the physical measurements recorded by the clinician were used throughout. However, a sensitivity analysis was carried out just on the RCT data to establish whether the method of measurement (i.e. physical measurements or digital images) had an impact on the results.

Statistical analyses were conducted using Stata v13 (Stata Corporation, TX, U.S.A).
RESULTS

Participant characteristics and missing data

A total of 179 patients were available for analysis - 112 patients who participated in the STOP GAP trial and 67 patients from the cohort study. The baseline characteristics of the 179 patients are given in Table 1. One patient was missing a baseline lesion measurement and so was excluded from all analyses. At the 2 week visit, 18 patients were missing all three measurements for lesion size, Investigator Global Assessment and resolution of inflammation and so were excluded from all 2 week analyses. At the 6 week visit, 15 patients were missing all three measurements and so were excluded from all 6 week analyses. One patient was missing a measurement for resolution of inflammation and so was excluded solely from the 6 week analysis for resolution of inflammation. Ten patients were missing lesion size at 6 weeks and so were excluded solely from the analysis for speed of healing.

Assessment of predictors of healing

The PPV, NPV and AUC were calculated for the three different measures of early treatment response (Table 2). Figure 1 shows the receiver operating characteristic (ROC) curves at 2 and 6 weeks for each of the three measures that were considered as predictors for healing or non-healing at 6 months.

All three measures demonstrated an AUC greater than 0.7 at both 2 and 6 weeks. Investigator Global Assessment for grade of lesion improvement had the best combined PPV, NPV and AUC at 2 weeks and at 6 weeks.

Physical measurements vs. digital images for speed of healing

Of the 112 patients (104 after excluded missing data) who participated in the RCT, 86 (82.7%) had their lesion size measurements based on digital images in addition to physical measurements. A sensitivity analysis was carried out to assess whether there were any
differences in terms of predictive value between speed of healing estimated using 100% physical measurements and speed of healing estimated using 82.7% digital images and 17.3% physical measurements. Table 3 gives the PPV, NPV and AUC for each of these as a predictor of healing at 6 months. In terms of predicting healing at 6 months, there were no significant differences between the digital images and physical measurements for speed of healing (Table 3).
DISCUSSION

Main findings

Speed of healing, Investigator Global Assessment and resolution of inflammation were all shown to be good predictors of eventual healing. The Investigator Global Assessment was marginally the best of the three measures. In terms of the timing of assessments, the 6-week measurements were better predictors of eventual healing than assessments at 2 weeks, and would be the most advisable time-point to use in future trials. However, the 2-week measurements were reasonably predictive and could still be useful for clinical practice.

Speed of healing estimated through physical measurements or digital images yielded no differences in terms of predicting eventual healing. This indicates that the digital images may be just as good as other clinical indicators. As such, if a blinded outcome is needed in future trials of pyoderma gangrenosum then digital images could be considered for this.

These findings support the choice of primary outcome in the STOP GAP trial (speed of healing at 6 weeks, assessed by blinded assessors using digital images), and suggest that important clinical differences were not missed as a result of this focus on early treatment response.

In addition, the Investigator Global Assessment and the resolution of inflammation scale were both shown to be good early predictors of healing. Both of these are relatively simple tools to use that could prove useful in clinical practice when making decisions on whether to stop, switch or alter doses of treatment.

Relevance to other studies
Several other studies have investigated early predictors of wound healing in venous and diabetic foot and leg ulcers. These studies reported early response at week 4 to be a good predictor of healing at 12 to 24 weeks \(^5,9,10\).

**Strengths and limitations**

This study is the first to assess the utility of early predictors of healing in patients with pyoderma gangrenosum and represents efficient re-use of data to inform clinical practice and trial design. Limitations of this study include the difficulty of defining the reference standard for eventual healing. Lesions were considered to have healed when sterile dressings were no longer required, which is a patient-orientated definition of healing. An alternative definition could have been complete healing of the lesion, but this would have required more frequent clinic assessments than were possible in the clinical trial. We were also limited by the data available from the STOP GAP trial and cohort study in that measurements were only taken at 2 and 6 weeks after start of treatment. It is possible that other time points could have been equally good predictors of eventual healing.

**Conclusion**

Early treatment response appears to be a good indicator of eventual healing, regardless of how it is measured. This finding is helpful for informing future clinical trial design and clinical decision-making.

**ACKNOWLEDGEMENTS**

The STOP GAP trial and STOP GAP cohort study were funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research funding scheme (RP-PG-0407-10177). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.
REFERENCES


10. Sheehan P, Jones P, Caselli A, Giurini JM, Veves A. Percent change in wound area of diabetic foot ulcers over a 4-week period is a robust predictor of complete healing in a 12-week prospective trial. Diabetes Care 2003;26:1879-82.
Figure 1: Pyoderma gangrenosum. Receiver operating characteristic (ROC) curves at 2 weeks (a) and 6 weeks (b) for each of the three measures considered as predictors of healing or non-healing at 6 months.

* Adjusted for age, gender, baseline lesion size, underlying systemic disease and lesion location.
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>N=179</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>55.54 (16.66)</td>
</tr>
<tr>
<td>Female</td>
<td>118 (65.92%)</td>
</tr>
<tr>
<td>Location of target lesion:</td>
<td></td>
</tr>
<tr>
<td>Upper limbs</td>
<td>10 (5.59%)</td>
</tr>
<tr>
<td>Lower limbs</td>
<td>115 (64.25%)</td>
</tr>
<tr>
<td>Not limb</td>
<td>54 (30.17%)</td>
</tr>
<tr>
<td>Underlying systemic disease</td>
<td>59 (32.96%)</td>
</tr>
<tr>
<td>Baseline lesion area</td>
<td>7.64 (2.81 to 18.84)*</td>
</tr>
</tbody>
</table>

*Based on 178 patients.*

Table 1: Baseline characteristics of patients in the STOP GAP trial and the observational study. Values are number (%), mean (standard deviation) or median (interquartile range).
<table>
<thead>
<tr>
<th>Time Point</th>
<th>No. patients in analysis</th>
<th>Positive Predictive Value</th>
<th>Negative Predictive Value</th>
<th>Area under ROC curve (AUC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed of healing</td>
<td>2 weeks</td>
<td>159</td>
<td>68.3% (55.3% to 79.4%)</td>
<td>67.7% (57.4% to 76.9%)</td>
</tr>
<tr>
<td></td>
<td>6 weeks</td>
<td>152</td>
<td>63.5% (52.4% to 73.7%)</td>
<td>74.6% (62.5% to 84.5%)</td>
</tr>
<tr>
<td>Investigator Global Assessment</td>
<td>2 weeks</td>
<td>159</td>
<td>73.2% (59.7% to 84.2%)</td>
<td>68.0% (58.0% to 76.8%)</td>
</tr>
<tr>
<td></td>
<td>6 weeks</td>
<td>163</td>
<td>80.0% (68.7% to 88.6%)</td>
<td>74.2% (64.1% to 82.7%)</td>
</tr>
<tr>
<td>Resolution of inflammation</td>
<td>2 weeks</td>
<td>159</td>
<td>66.1% (53.0% to 77.7%)</td>
<td>66.0% (55.7% to 75.3%)</td>
</tr>
<tr>
<td></td>
<td>6 weeks</td>
<td>162</td>
<td>72.1% (59.9% to 82.3%)</td>
<td>68.1% (57.7% to 77.3%)</td>
</tr>
</tbody>
</table>

Table 2: Predictive values and area under the receiver operating characteristic curve (AUC) (95% confidence intervals) at 2 weeks and 6 weeks for the three measures considered as predictors of healing or non-healing at 6 months.

* Adjusted for age, gender, baseline lesion size, underlying systemic disease and lesion location.
Table 3: Sensitivity analysis to compare results at 6 weeks using physical measurements alone, or a mixture of physical measurements and digital images. Analyses only carried out on RCT data (n=104).

* Adjusted for age, gender, baseline lesion size, underlying systemic disease and lesion location.

<table>
<thead>
<tr>
<th>Speed of healing at 6 weeks</th>
<th>Method of measurement</th>
<th>Positive Predictive Value</th>
<th>Negative Predictive Value</th>
<th>Area under ROC curve (AUC)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Physical measurements only</td>
<td>69.0% (55.5% to 80.5%)</td>
<td>80.4% (66.1% to 90.6%)</td>
<td>0.8434 (0.7701 to 0.9168)</td>
</tr>
<tr>
<td></td>
<td>Mixture of physical measurements (17.3%) &amp; digital images (82.7%)</td>
<td>72.9% (59.7% to 83.6%)</td>
<td>86.7% (73.2% to 94.9%)</td>
<td>0.8623 (0.7936 to 0.9311)</td>
</tr>
</tbody>
</table>
Grade for lesion improvement was measured by the clinician using the Investigator Global Assessment

<table>
<thead>
<tr>
<th>INVESTIGATOR GLOBAL ASSESSMENT OF EFFICACY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
</tbody>
</table>

Resolution of inflammation was measured using the scale reported by Foss¹

<table>
<thead>
<tr>
<th>INFLAMMATION ASSESSMENT OF THE TARGET LESION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please tick one box only for each section</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Erythema</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
</tr>
<tr>
<td>Slight</td>
</tr>
<tr>
<td>Moderate</td>
</tr>
<tr>
<td>Severe</td>
</tr>
<tr>
<td>Very severe</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Border elevation</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
</tr>
<tr>
<td>Slight</td>
</tr>
<tr>
<td>Moderate</td>
</tr>
<tr>
<td>Severe</td>
</tr>
<tr>
<td>Very severe</td>
</tr>
</tbody>
</table>
