Study of Treatments of Pyoderma Gangrenosum Patients – the STOP GAP Trial

Final report for REC and Competent Authority (MHRA / IMB)

EudraCT Number: 2008-008291-14

REC Reference: 09/H0903/5 (UK only)

Summary

What was already known about this topic?

- PG is a severe, painful ulcerative condition of the skin, with a significant mortality.
- The evidence base for currently treatments is very weak, with a single published RCT involving 30 participants.
- In more severe cases, prednisolone has historically been the main systemic therapy, but many clinicians have since switched to another systemic therapy, ciclosporin, in the belief it is more effective and has fewer side effects.
- Both prednisolone and ciclosporin have significant predictable side effects.

What did this study add?

- The hypothesis that ciclosporin is more efficacious than prednisolone was not supported by our trial evidence. Both agents were of similar efficacy, and neither were outstanding, with only about 50% of ulcers healed by 6 months.
- We found that adverse events were very common (around two thirds of patients) in both groups, and the nature of the adverse event profile (serious infections with prednisolone and hypertension and renal dysfunction with ciclosporin) may help to inform which treatment could be considered according to underlying patient risk factors
- Recurrence of PG is common – approximately one third of cases will suffer a further episode within 2 to 3 years.
- Current therapeutic strategies are inadequate and further research into more effective treatment options in required.

Introduction

PG is an inflammatory ulcerative skin disease, which is frequently painful, and often occurs in association with conditions such as inflammatory bowel disease (IBD), arthritis or

haematological malignancy[701]. A 2012 retrospective cohort study in the UK has reported an adjusted incidence rate (standardised to the European standard population) of 0.63 per 100,000 person-years. The development of PG was associated with a three-fold increased risk of death compared to general population controls, and a 72% increased mortality over controls with IBD[702].

There are currently no national or international guidelines covering the management of PG. Patient information issued by the British Association of Dermatologists (BAD) describes topical and systemic treatment options, as well as lesser used options such as intravenous steroids or biologics. Topical treatments for PG include potent steroid preparations or calcineurin inhibitors, and commonly prescribed systemic treatments comprise of antibiotics, steroids and immunosupressants[703]. Only one RCT in patients with PG is reported in the literature[704]. This was a small study of 30 patients compared the anti-tumour necrosis factor α (TNF-α) monoclonal antibody infliximab® (Remicade, Schering-Plough) (5 mg/kg) to placebo. Significantly more patients in the infliximab® group demonstrated clinical improvement at 2 weeks compared to placebo (the primary endpoint; 46 versus 6%, respectively, p = 0.025). However, due to its cost, infliximab® is not currently considered a first-line treatment for PG.

Consistent with the lack of good quality RCT evidence, systematic reviews of treatments for PG have primarily relied upon anecdotal reports or retrospective case series[705]. Based on the available evidence, systemic corticosteroids such as prednisolone are generally considered to be the most predictable, effective medications for PG when delivered in adequate doses[706]. However, retrospective data also lend support to the use of ciclosporin. A number of case reports document complete remission of steroid refractory, IBD-related PG lesions with ciclosporin. Complete response was reported for all participants in a study of five patients receiving oral ciclosporin 4-5 mg/kg/day[707], and 11 patients receiving the drug at an initial concentration of 4 mg/kg/day intravenously[708]. Other case series have reported encouraging proportions of patients, with a range of underlying diseases, achieving complete responses to ciclosporin (10 of 11[709] and 3 of 7[710]). Vidal and colleagues performed a review of 26 cases of classical PG[711, 712]. Ciclosporin 3-6 mg/kg/day was used in 22 of these patients, with 51 episodes between them. Among these episodes, a complete response was recorded for 96% and a partial response for 3%. The second most commonly

used treatment was prednisone, used in 15 patients and 26 episodes. Complete responses were recorded in 61% episodes, partial responses for 26% and no response for 11%. However, data from these studies are challenging to interpret as the majority of patients were receiving concomitant steroids rather than ciclosporin alone.

Given the complete absence of high-quality evidence on the first-line treatment of PG, an RCT (STOP GAP) was conducted in order to test the hypothesis that ciclosporin was superior to prednisolone in the treatment of PG.

Methods

A summary of the trial methods is presented here; a more detailed account of the trial protocol has been published [713], and the protocol and statistical analysis plan are appended (See *Appendix 21* and *22* and *23*).

Trial Design and Oversight

STOP GAP was a multicentre, parallel-group, observer-blind RCT, to compare the efficacy and safety of ciclosporin with prednisolone. It was a pragmatic trial that reflected current practice as far as possible. Patients were assessed at baseline, 2 weeks, 6 weeks and when the ulcer had healed (maximum of 6 months). Appropriate national ethics and regulatory approvals (ethics: 09/H0903/5, Medicines and Healthcare Products Regulatory Agency: 19162/0213/001, EudraCT: 2008-008291-14) were obtained; all participants gave written informed consent. The trial was co-ordinated from the NCTU at the University of Nottingham. Oversight of the trial was performed by monthly Trial Management Group meetings and an independent Trial Steering Committee which met twice a year. All data issues, including safety, were overseen by progress reports presented to an independent Data Monitoring Committee. The trial was registered at Controlled-Trials.com (ISRCTN35898459) prior to start of recruitment.

Participants

Recruitment took place at 39 hospitals in the United Kingdom; inclusion and exclusion criteria are listed in Table 1.

Table 1: Inclusion and Exclusion criteria

Inclusion Criteria

- PG as diagnosed by the recruiting dermatologist. (An ulcerative lesion may have mixed aetiology, but provided the investigator has confidence that a clinical diagnosis of PG is appropriate then they are eligible. Other contributing factors and atypical features will be captured in the case report form).
- Must have a measurable ulceration (e.g. not pustular pyoderma gangrenosum)
- Age over 18 years.
- Able to provide written, informed consent.

Exclusion Criteria

- Granulomatous PG – this condition is very rare and may respond differently to treatment.
- Ciclosporin or prednisolone or IVIG therapy in the previous month.
- Already participating in another clinical trial.
- Pregnant, lactating or at risk of pregnancy.
- Hypersensitivity to prednisolone or ciclosporin
- Biopsy consistent with a different diagnosis.
- Biopsies will be used to exclude alternative aetiologies (e.g. malignancy, granulomatous PG, arteritis) rather than to confirm the diagnosis of PG, since histology is supportive rather than pathognomic. Ideally, the biopsy will be a 1.5cm rectangular biopsy taken through the edge of the ulcer and left to granulate and heal by secondary intention. Alternatively, 2 separate punch biopsies done at the edge of the ulcer and at the extending margin may be used. It is not normal practice to await histological confirmation before initiating therapy, so patients will be randomised prior to receiving histological results. If the histology indicates an alternative aetiology, the participant will be excluded at that time.
- Clinically significant renal impairment that would result in the investigator not normally treating with either study drug.
- Any pre-treatment investigations, the results of which would prompt the investigator not to use either study drug.

- A diagnosis of malignancy or pre-malignant disease where treatments might interfere with ongoing therapy or might cause harm (e.g. history of lymphoma, multiple lymphoma, leukaemia, cervical epithelial neoplasia – CIN, systemic cytotoxic therapy)
- The patient has a concurrent medical condition that means the investigator would not normally treat the patient with either of the study drugs (for example: a degree of hypertension that would not lead to using either of the study drugs, advanced heart failure, poorly-controlled diabetes, history of peptic ulcer, malignancy in previous years).
- Administration of a live vaccine (BCG, Measles, Mumps, Rubella, Yellow Fever, Oral Polio, Oral Typhoid) within the last 2 weeks
- The patient is currently taking Rosuvastatin (Crestor®) for the treatment of hypercholesterolaemia, since this is contra-indicated when taking Neoral® (ciclosporin).

Participants were asked not to use any topical therapy (for example, corticosteroids or calcineurin inhibitors) after randomisation. Patients who required first-line topical therapy were invited to enter a parallel observational study (See **Error! Reference source not found.** - **Error! Reference source not found.**).

Interventions

Participants were randomised to receive either oral prednisolone 0.75 mg/kg/day in a single dose or ciclosporin (Neoral®, Novartis Pharmaceuticals) 4 mg/kg/day in two divided doses. The dose of study drug could be adjusted (up or down) according to normal practice, though clinicians were encouraged not to alter the dose until week two if possible. The maximum increase permitted per day was 1 mg/kg/day for prednisolone and 5 mg/kg/day for ciclosporin.

A change to the protocol was made in August 2011 after recruitment of 82 participants. A patient with a very high BMI who was randomised to prednisolone experienced bowel perforation on a dose of 110 mg/day. As a result of this serious adverse event, ceiling doses

of 75 mg/day of prednisolone and 400 mg/day of ciclosporin were implemented from thereon, regardless of body weight.

Randomization and blinding

Participants were randomised (1:1) to treatment allocation using a computer-generated pseudorandom list, using permuted blocks of randomly varying size between two and six (Using the RALLOC add-on[714] for Stata, Stata Corporation, Texas, USA). Randomisation was stratified by lesion size (≥20 cm^2 versus lesions <20 cm^2) and presence or absence of underlying systemic disease. For the purposes of randomisation, lesion size was estimated based on the maximum longitudinal length and maximum perpendicular length and converted to approximate area by the formula (1), which approximates to an ellipse.

$$\textbf{Length} \times \textbf{Width} \times \textbf{0.785} \qquad (1)$$

Treatment allocation was concealed until interventions were all assigned, and recruitment, data collection, data cleaning and analysis using dummy treatment codes were complete except for the purpose of DMC analyses.

This was an observer-blind study; the primary outcome (velocity of healing) and global treatment response were assessed from digital images of the target lesion by assessors blind to the allocated treatment; clinicians and participants were however aware of their treatment allocation. Full blinding was not possible due to logistic and methodological difficulties in blinding treatment allocation. For example, the two drugs require different dosing regimens and different arrangements for monitoring of side-effects. Blinding the trial interventions using placebo medications was beyond the scope of this pragmatic trial, and may have had a detrimental effect on treatment adherence resulting from need for additional tablets. Nevertheless, treatment allocation was only revealed to the recruiting physician once participants' details and key stratification variables had been irrevocably entered by the physician onto the web-based randomisation system maintained by NCTU.

Assessments

Clinic visits were conducted by a dermatologist and took place at baseline, week 2, week 6 (primary outcome) and when the ulcer had healed (up to a maximum of 6 months post-randomisation). Clinic visits consisted of standard clinical tests, medical history taking,

assessment of side effects, measurement of the ulcer, and clinician's evaluation of the target lesion. A digital image of the target lesion was also taken at baseline, week 6 and final visit. The same clinician saw the participants at each clinic visit whenever possible.

Participants assessed the severity of their PG and quality-of-life at baseline, 6 weeks and on healing (or 6 months if not healed) (See *Appendix 24* and *25* for baseline CRF and week 6 CRF). In addition, they completed a study diary that captured daily pain scores and use of analgesics for the first 6 weeks, plus impact on daily activities, use of dressings, adverse events and the use of health services throughout the trial (See *Appendix 27*). Adherence to trial medication was assessed using patient diaries. These data were categorized as using medication every day, most days, some days or never.

At the end of the trial, investigators obtained hospital records of those for whom the target lesion had healed to ascertain the recurrence of PG and time to recurrence.

Digital images were used to assess the blinded outcomes of velocity of healing and global treatment response. If digital images were not available, then physical measurements of the lesion taken during clinic visits, and global response by the treating clinician were used.

A standardised template was photographed alongside the target ulcer in order to calibrate the image in the image analysis software (See *Figure 1*). Images were stored electronically and transferred in an anonymised fashion to the coordinating centre at NCTU. Each image was loaded into the image analysis software and the circumference of the lesion was manually drawn by two trained assessors (BE and JP) using Verge Videometry VEV MD software (Vista Medical, Winnipeg, Canada).

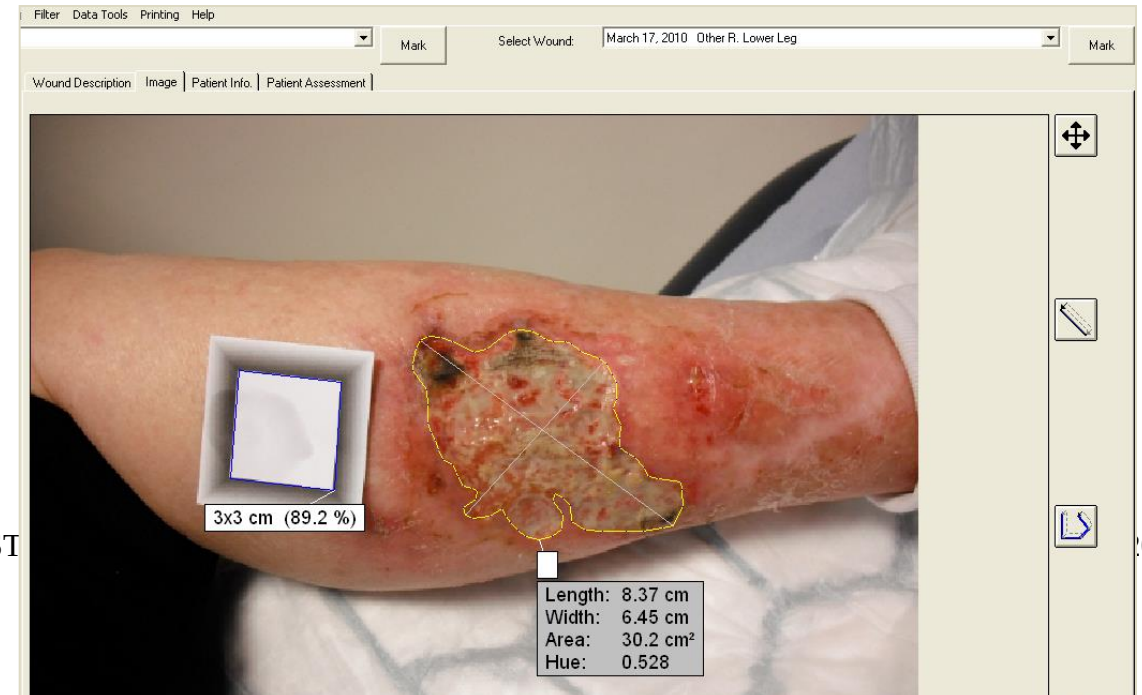


Figure 1: Pyroderma gangrenosum ulcer measurement using image analysis software

For the global treatment response, an independent dermatologist (SO) scored patients' treatment response using a pair of images from baseline and final visit.

Quality control of digital image assessments

All images were independently reviewed by two dermatologists to ensure that the lesions were consistent with a diagnosis of PG, and that the measurements taken by the trained assessors were an accurate representation of the ulcer size. In cases where discrepancies were observed the following rules were applied:

Image assessments	Action
• Both measurements map the lesion appropriately and measurements agree within a ratio of 1:1.1	Use mean of two measurements
• Both measurements map the lesion appropriately but measurements disagree by a ratio of greater than 1:1.1	Two dermatologists to re-measure and use the mean of the new measurements
• One measurement has mapped the lesion incorrectly (e.g. where healed areas have	Discard the image that was mapped incorrectly and use the second measurement

mistakenly been included due to residual erythema or surface changes)	
• Both measurements have mapped the lesion incorrectly	Dermatologists to re-draw the measurement and use this for analysis

Outcomes

Primary outcome

Velocity of healing at 6 weeks measured used the formula below (2). Date *X* is the earliest of either the date at which the lesion stopped requiring dressings if this occurred prior to 6 week visit, or the date of the 6 week visit. Healing was captured for a single target lesion per patient. If multiple lesions were present, the lesion that could be photographed on a single plane (i.e. not around the curvature of a limb) was designated the target lesion.

Change in area (cm^2) / (Date *X* – randomisation date (days)) (2)

Velocity of healing was chosen for the primary outcome as it has been shown in previous studies to be a good predictor of healing in patients with leg ulcers[715, 716], and because blinded assessment was possible using digital images and independent assessors, being able to assess velocity of healing at 6 weeks also minimised the risk of missing data. Nevertheless, it was always our intention that time to healing be considered the most important secondary outcome, as it is more clinically relevant and is easier to interpret. Time to healing also gives an indication of the duration of treatment, and therefore the potential for cumulative drug toxicity.

Secondary outcomes

Time to healing. Assessed by participants based on the time at which sterile dressings were no longer required for the wound, and confirmed using digital photography at the first opportunity. If the date the lesion stopped requiring dressings was not recorded, the date of the clinic visit was used.

PG-specific global treatment response. A seven-point Likert scale ranging from completely clear through to worse (assessed by clinicians, participants and by digital images).

Resolution of inflammation. This is a previously published PG assessment scale including erythema and border elevation[717]. Erythema and border elevation are each scored from zero to four (representing none through to very severe). Resolution of inflammation was taken to have been achieved if both items were scored as zero (none) as per the original scale[717]. This score was recorded by clinicians for each clinic visit and participants by completing a postal-return questionnaire. In response to feedback from patients during development of the trial, an additional question probing the degree of exudate was also included. This information is presented separately and was not incorporated into the "resolution of inflammation" score.

Self-reported pain. For the first 6 weeks, participants reported daily pain severity in their study diary on a scale from 0 to 4 (none, mild, moderate, severe or extreme), and whether or not painkillers were taken that day.

Health-related quality of life. Assessed at baseline, 6 weeks and 6 months (or healed), using validated questionnaires (Dermatology Life Quality Index, (DLQI)[467], European Quality of Life-5 Dimensions (EQ-5D-3L)[718] and EQ-5D visual analogue scale (EQ-5D VAS)[345]). DLQI is a disease specific quality of life measure including 10 questions each with 4 levels (0-3) scored from 0 (no effect) to 30 (extremely large effect on quality of life). EQ-5D-3L is a generic quality of life measure with 5 domains each scored at 3 levels: findings are mapped onto societal health state preference values referenced to scores of 0 (dead) and 1 (perfect health). Negative scores are possible for some heath states considered worse than death.

Cost--analysis. Costs and health service resource use were compared from a health service perspective. Patient diaries were used to capture health service contacts related to the treatment of PG. These were then returned to the trial team during scheduled contacts (at 6 weeks and 6 months). National unit costs for 2012 were applied to resource use providing a cost of care for each patient during follow-up, unit costs included: outpatient visits (£139); community nurse visits (£39); practice nurse contacts (£14); GP consultations (£43) and GP home visits (£110)[348]. Hospital consultations were calculated by identifying the commonest HRG codes for PG hospital admissions from national Hospital Episodes Statistics (HES) data and calculating a weighted average per diem (£323/day) for these codes using national reference costs [719]. Ciclosporin and prednisolone treatment costs were estimated in two

steps. The product of the daily dose prescribed and treatment duration were used to calculate a total quantity of treatment (mg). National prescribing data were accessed to calculate the average cost by weight of these drugs given the current national prescribing pattern [720]. These average costs were then applied to the weight of active drug provided and the patient drug cost added to their cost of care.

Time to recurrence. A recurrence was defined as the occurrence of a further episode of PG (at any site) that appeared after the target lesion was confirmed as being healed by a physician or nurse. The period of follow-up available varied depending on the time at which the participant was randomised into the trial.

Number of treatment failures. Treatment failures were defined as being participants who withdrew (or were withdrawn) from their randomised treatment because of treatment intolerance or worsening of the PG, or those whose target lesion remained unhealed after 6 months of follow-up.

Adverse reactions to study medications. Defined as adverse events that were possibly, probably or definitely related to the study medication.

Sample size

This was a superiority trial, with prednisolone as the control intervention. In order to provide 80% power (5% level of significance) to detect a difference in means of 0.5 standard deviations in the primary outcome of velocity of healing at 6 weeks, the total target sample size was 140 participants, assuming a loss to follow-up of 10%.

Statistical Analysis

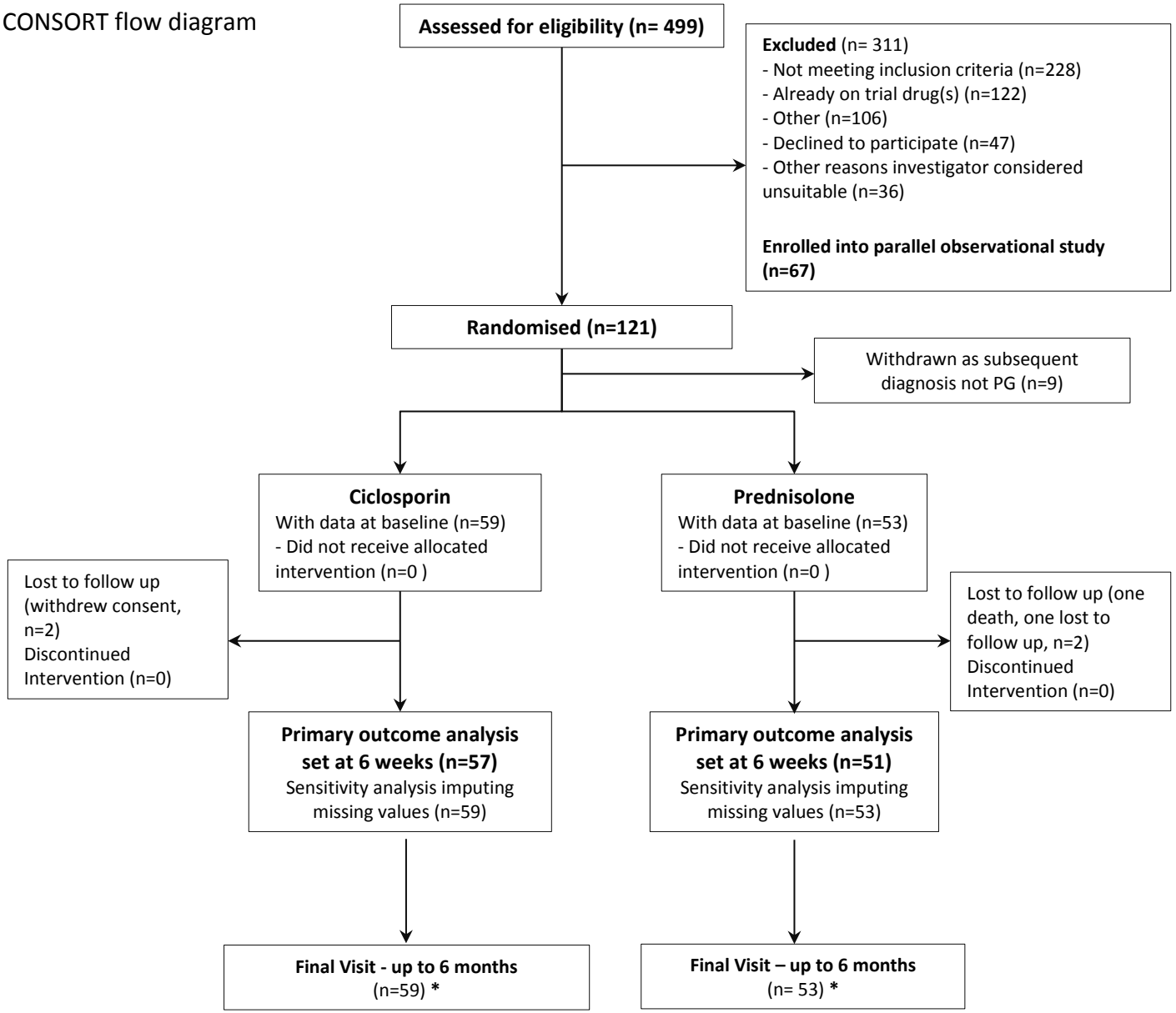
The primary analysis was conducted according to the ITT principle. The ITT population was defined as all randomised patients, excluding those whose later diagnosis was determined to be something other than PG. All patients with available data at both the baseline and the six week visit were included in the primary analysis. By way of sensitivity analyses, if neither a digital image, nor physical measurements taken during clinic visits were available at 6 weeks, multiple imputation was used based on the assumption that the data were missing at random. We also used the date of visit at which the lesion was declared as healed, in the event that the date the lesion stopped requiring dressings was not available. Differences

between treatment groups for the primary outcome at 6 weeks were analysed using a linear regression model.

Secondary outcomes were analysed as follows: Cox regression models for the time to healing of the target lesion and the time to recurrence; linear regression models for DLQI, EQ-5D and EQ-VAS scores (adjusted for baseline values), and for self-reported pain (which were summarized using AUC); proportional odds models for the categorical secondary outcomes, including global assessment of improvement; logistic regression models for the resolution of inflammation (by clinician and patient). Comparisons between the average number of times painkillers were used between the treatment groups were made using an adjusted non-parametric test. Analyses were adjusted for the stratification variables of lesion size and presence or absence of underlying autoimmune disease. In addition, sensitivity analysis of the primary outcome further adjusted for additional baseline variables including: age, sex, weight, size of recruiting centre and geographical region.

Other sensitivity analyses were conducted in which participants who switched randomised treatments, or who received both trial drugs in combination during the period of the trial were either excluded from the analysis for the primary outcome of velocity of healing, or included for the secondary outcome of time to healing analysis, but censored at the time of change. Resource and cost data were highly skewed, and thus parameter uncertainty was estimated by method of bootstrap using 10,000 replications. All statistical analyses were conducted with the use of SAS software, version 9.2 and R version 2.10.1.

Figure 2: CONSORT flow diagram



Results

Study population

Recruitment took place from May 2009 to November 2012.

Of 499 patients screened, 121 were eligible for the trial and gave written informed consent (86% of target). Fifty-nine were assigned to the prednisolone group, and 62 to the ciclosporin group (Figure 2). Of these patients, nine (six in the prednisolone group, three in the ciclosporin group) were subsequently found not to have PG and so were withdrawn after randomisation; making the analysable population 112 participants. In addition, there were two losses to follow up in each arm before the primary endpoint at 6 weeks was reached. Two-thirds of participants (75/112; 67.0%) were recruited from dermatology clinics and the remainder from a wide range of other disciplines including gastroenterology, general medicine, surgery, rheumatology and tissue viability.

The median number of participants per recruiting centre was 3 (minimum 1, max 20).

Baseline characteristics of the participants were well balanced between the groups (See Table 2). On entry into the trial, methotrexate was being taken by one patient in each group; azathioprine was being taken by three patients in the ciclosporin group and one in the prednisolone group; and tetracycline was being taken by three patients in the prednisolone group.

Table 2: Baseline Characteristics

		Ciclosporin (n=59)	Prednisolone (n=53)
Demographics			
Age: years	Mean (SD)	57.2 (16.9)	51.3 (15.2)
Sex: n (%)	Female	42 (71.2)	31 (58.5)
Ethnicity: n (%)	White	55 (93.2)	53 (100)
Weight: kg	Mean (SD)	88.4 (24.5)	93.2 (27.2)
	Min; max	50.0, 171.0	50.6, 151.0
Medical History			

Underlying co-morbidities: n (%)	Crohns Disease	5 (8.5)	3 (5.7)
	Ulcerative colitis	7 (11.9)	8 (15.1)
	Rheumatoid arthritis	4 (6.8)	4 (7.5)
	Other inflammatory arthritis	3 (5.1)	3 (5.7)
	Monoclonal gammopathy	0 (0.0)	0 (0.0)
	Myeloma	0 (0.0)	0 (0.0)
	Haematological malignancy	0 (0.0)	0 (0.0)
	Other malignancy	4 (6.8)	0 (0.0)
	Diabetes	4 (6.8)	9 (17.0)
	Renal impairment	2 (3.4)	0 (0.0)
	Epilepsy	0 (0.0)	1 (1.9)
Characteristics of PG			
Type of PG: n (%)	Classical	50 (84.7)	47 (88.7)
	Cribriform	4 (6.8)	2 (3.8)
	Peristomal	2 (3.4)	2 (3.8)
	Bullous	0 (0.0)	1 (1.9)
	Unsure	3 (5.1)	1 (1.9)
Previous episode of PG: n (%)		17 (28.0)	14 (26.4)
Area of target lesion: cm^2	Median (Q1; Q3)	9.1 (3.6; 24.7)	8.1 (2.4; 20.2)
Location of lesion: n (%)	Upper limb	2 (3.4)	1 (1.9)
	Lower limb	41 (69.5)	34 (64.2)
	Other	16 (27.1)	18 (34.0)
Number of lesions	Number	(n=59)	(n=51)
	Mean (SD)	2.2 (1.8)	2.6 (2.4)
	Min; max	(1, 10)	(1, 12)
Erytherma n (%)	None	4 (6.8)	2 (3.8)
	Slight	2 (3.4)	3 (5.7)
	Moderate	21 (35.6)	15 (28.3)

	Severe	22 (37.3)	17 (32.1)
	Very Severe	10 (16.9)	16 (30.2)
Border Elevation n (%)	None	4 (6.8)	1 (1.9)
	Slight	24 (40.7)	29 (54.7)
	Moderate	19 (32.2)	17 (32.1)
	Severe	8 (13.6)	5 (9.4)
	Very Severe	4 (6.8)	1 (1.9)
Exudate n (%)	None	3 (5.1)	1 (1.9)
	Slight	7 (11.9)	9 (17.0)
	Moderate	32 (54.2)	27 (50.9)
	Severe	12 (20.3)	3 (5.7)
	Very Severe	5 (8.5)	13 (24.5)

During the trial, 16/112 (14.3%) of participants either switched to the alternative trial drug or received the two drugs concurrently (8; 15.1% for prednisolone versus 8; 13.6% for ciclosporin). Of these events, five occurred prior to the primary outcome assessment at 6 weeks (one in the prednisolone group and four in the ciclosporin group).

Data on adherence to study medication were available from 68/112 (60.7%) participants. Of these, 36/37 (97.3%) in the ciclosporin group and 29/31 (93.5%) in the prednisolone group took their treatment every day in the first 6 weeks of the trial.

Primary Outcome

In total, 108 (96.4%) participants had data at both baseline and week 6. Of these, 86 (79.6%) had blinded outcome data on the basis of digital images. For 22 (20.4%) participants, velocity of healing was assessed on the basis of 'unblinded' physical measurements taken by investigators during clinic visits, as digital images were either unavailable or of insufficient quality to allow assessment.

For participants with data at both baseline and 6 weeks, the median (Q1; Q3) lesion area at baseline was 9.1 (4.5; 24.7) cm^2 in the ciclosporin group, versus 6.8 (2.4; 20.2) cm^2 in the prednisolone group. The median (Q1; Q3) changes in lesion area from baseline to week 6

were –1.7 (-6.9; 0.7) cm^2 and –2.0 (-5.7; 0.4) cm^2, respectively. The calculated unadjusted mean (SD) for the velocity of healing at 6 weeks was -0.21 (1.00) cm^2/day for ciclosporin and -0.14 (0.42) cm^2/day for prednisolone.

The primary analysis showed no significant difference between the two treatments for velocity of healing at 6 weeks (See *Table 3*).

Sensitivity analyses for velocity of healing

Similar results were observed for the sensitivity analyses in which missing data were imputed [adjusted mean difference: 0.001 cm^2 /day (95% CI -0.204, 0.206); p=0.994], and separately, after adjusting for additional baseline covariates [adjusted mean difference: -0.100 cm^2 /day (95% CI -0.328, 0.127); p=0.382].

Excluding the five patients who either swapped to the alternative trial drug, or used both drugs in combination prior to the 6 week visit did not change the overall treatment effect: adjusted mean difference -0.036 (95% CI -0.211, 0.139), p=0.685.

Table 3 Velocity of healing at 6 weeks, time to healing by 6 months and time to recurrence subsequent to initial healing

Primary Outcome					
Velocity of healing at 6 weeks (cm^2 per day)	**Mean (SD)**	**Mean difference (ciclosporin – prednisolone)**	**Adjusted mean difference[#]**	**95% CI**	**p**
Ciclosporin (n=57)	-0.213 (0.998)	-0.074	0.003	-0.204, 0.211	0.975
Prednisolone (n=51)	-0.139 (0.417)				
Secondary outcomes					
Time to healing[^]	**Number healed by 6 months (%)[#]**	**Median time to healing in days (IQR)**	**Hazard ratio for healing [#]**	**95% CI**	**p**
Ciclosporin (n= 59)	28 (47.5%)	134.0 (60.0, 183.0)	0.94	0.55, 1.63	0.839
Prednisolone (n= 53)	25 (47.2%)	112.0 (46.0, 182.0)			
Time to recurrence	**Number with PG recurrence**	**Median time to recurrence in days (IQR)**	**Hazard ratio for healing [#]**	**95% CI**	**P**

	(%)[$]				
Ciclosporin (n=27)	8 (29.6)	582.0 (172.0, 932.0)	1.43	0.50, 4.07	0.501
Prednisolone (n=25)	7 (28.0)	612.0 (148.0, 934.0)			

adjusted for stratification variables (lesion size and presence of underlying disease). ^ Healed defined as the date that dressings were no longer required, or if this was missing (n=3) the date of the clinic visit at which healing was confirmed. $ in those who had healed by 6 months

Secondary Outcomes

All secondary outcomes were consistent with the primary outcome in showing no significant difference between the two treatments.

Time to healing

At 6 weeks, 9 (15.3%) in the ciclosporin group and 11 (20.8%) in the prednisolone group had healed. By 6 months, the proportion healed had increased to 47.5% (28) and 47.2% (25) in the ciclosporin and prednisolone groups respectively. The median time to healing was 134 days for ciclosporin compared to 112 days for prednisolone. The Cox regression model for time to healing showed no significant difference between the interventions (See Table 3 and Figure 4).

Sensitivity analyses for time to healing

Adjusting for additional baseline covariates was consistent with the main result (HR 1.01 (95% CI 0.57, 1.79); p=0.985).

Sixteen participants swapped to the alternative trial drug, or used both drugs in combination during the follow-up period (8 randomised to ciclosporin, 8 to prednisolone). Sensitivity analysis excluding these participants was also consistent with the main result (HR 0.861 (95% CI 0.49, 1.52), p = 0.604).

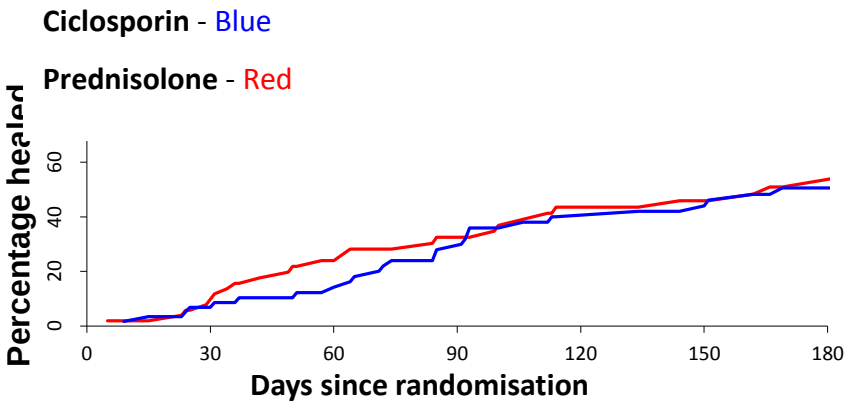


Figure 3: Kaplan Meier to time to healing by treatment group

Global assessment of efficacy

There were no significant differences between the treatments in global assessments of efficacy at final visit; whether based on data from physicians, patients or blinded assessments using digital images (See *Figure 4* and *Figure 5* and *Figure 6*).

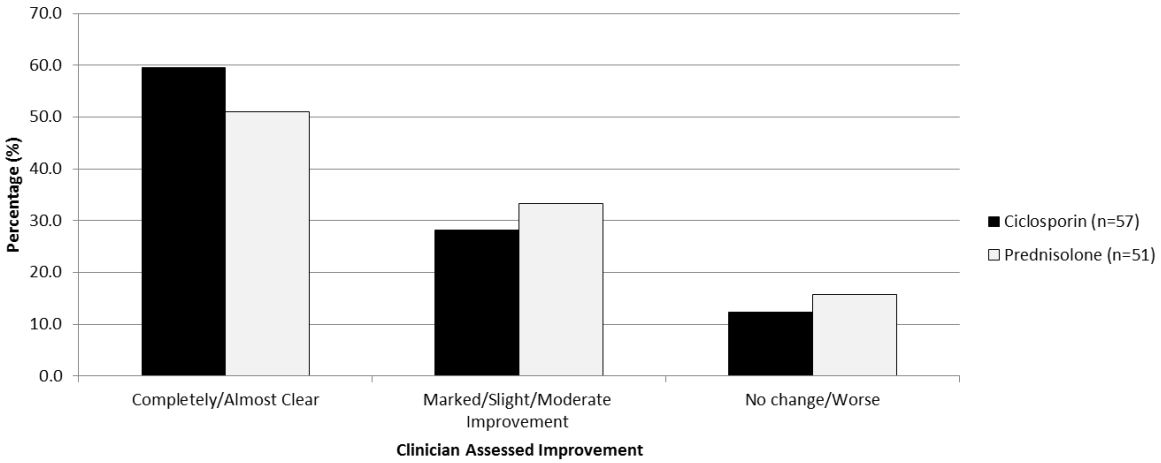


Figure 4: Global treatment response (by clinician)
p = 0.3285; Odds ratio: 1.457 (0.685, 3.098)

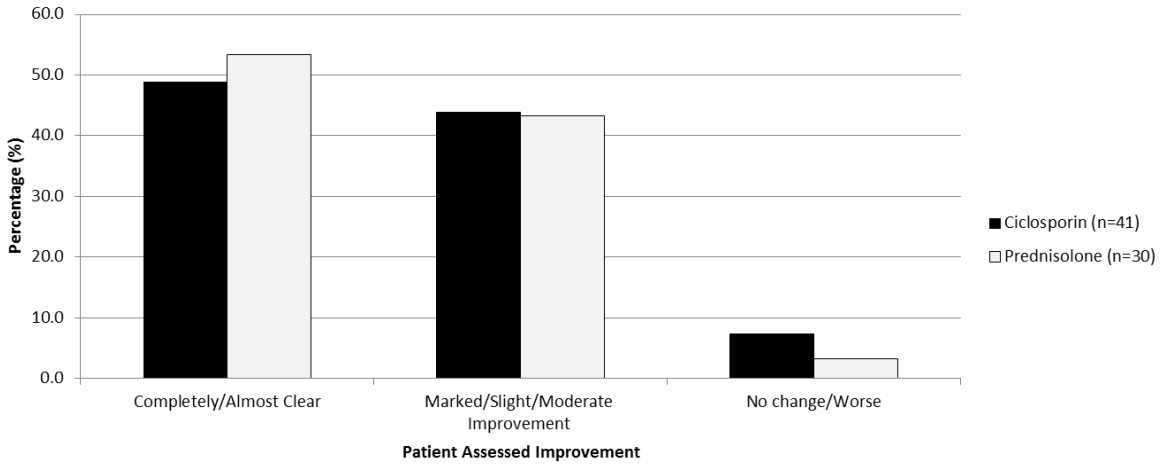


Figure 5: Global treatment response (by patient)
p = 0.6702; Odds Ratio: 0.814 (0.315, 2.103)

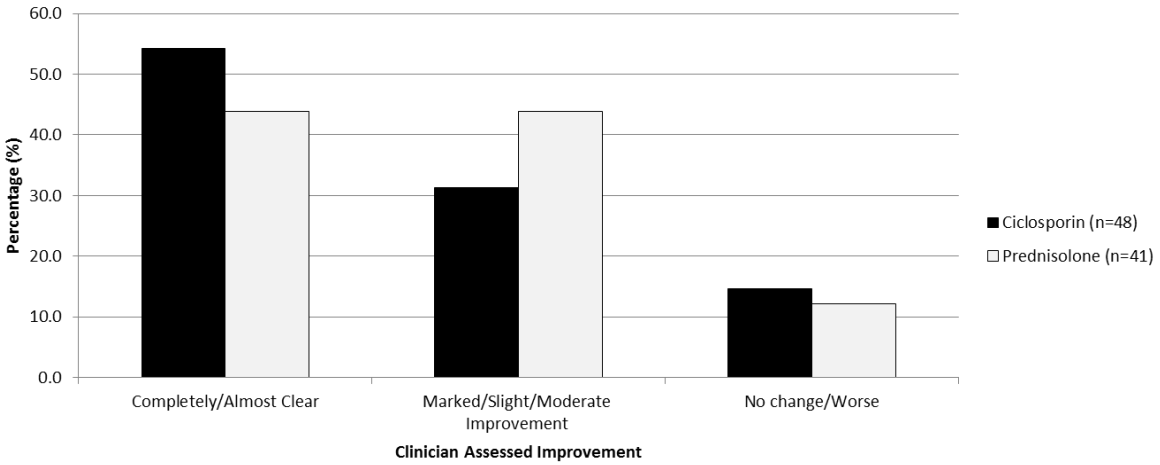


Figure 6: Global treatment response (by independent clinician from digital images)
p= 0.4199; Odds Ratio: 1.393 (0.623, 3.114)

Resolution of Inflammation

Full details of inflammation assessment at baseline (including change in exudate) are tabulated in Table 2. There were no between group differences in the resolution of inflammation as assessed by clinicians at either 6 weeks or final visit (See Table 6).

Table 4: Characteristics of changes in target lesions (erythema, border elevation and exudate) as assessed by investigator at week 6

Parameter	Information	Ciclosporin	Oral Prednisolone

Erythema		(n=56)	(n=51)
	Worse	8 (14.3)	6 (11.8)
	Same	13 (23.2)	13 (25.5)
	Improved	35 (62.5)	32 (62.7)
Border Elevation		(n=57)	(n=51)
	Worse	7 (12.3)	7 (13.7)
	Same	16 (28.1)	11 (21.6)
	Improved	34 (59.6)	33 (64.7)
Exudate		(n=57)	(n=51)
	Worse	6 (10.5)	6 (11.8)
	Same	16 (28.1)	14 (27.5)
	Improved	35 (61.4)	31 (60.8)

Table 5: Characteristics of changes in target lesions (erythema, border elevation and exudate) as assessed by investigator at final visit

Parameter	Information	Ciclosporin	Oral Prednisolone
Erythema		(n=57)	(n=51)
	Worse	6 (10.5)	3 (5.9)
	Same	11 (19.3)	10 (19.6)
	Improved	40 (70.2)	38 (74.5)
Border Elevation		(n=57)	(n=51)
	Worse	2 (3.5)	8 (15.7)
	Same	15 (26.3)	9 (17.6)
	Improved	40 (70.2)	34 (66.7)
Exudate		(n=57)	(n=51)
	Worse	5 (8.8)	4 (7.8)
	Same	7 (12.3)	8 (15.7)
	Improved	45 (78.9)	39 (76.5)

Table 6: Resolution of inflammation at 6 weeks and by final visit[1]

	n	Week 6 n (%)	Odds ratio[2]	95% CI	p
Ciclosporin	56	5 (8.9)	1.03	0.27, 3.97	0.964
Prednisolone	51	6 (11.8)			
	n	**Final visit (up to 6 months) n (%)**	**Odds ratio[2]**	**95% CI**	**p**
Ciclosporin	57	10 (17.5)	1.11	0.39, 3.12	0.849
Prednisolone	51	10 (19.6)			

[1]Based on border elevation and erythema reduced to "none"[717]
[2]Adjusted for stratification variables (lesion size and presence of underlying disease).

Self-reported pain

The mean (SD) self-reported pain score reduced from 1.92 (1.06) in week 1, to 1.26 (1.15) by week 6. There was no difference between ciclosporin and prednisolone groups in area under the curve (AUC) for the average weekly pain scores over the first 6 weeks (Table 7).

The median (Q1; Q3) number of days on which painkillers were used in the first 6 weeks was 14.0 days (0.0, 38.0) in the ciclosporin group and 20.5 days (1.0, 40.0) in the prednisolone group, with a non-significant treatment effect (p = 0.782).

Table 7: Self-reported pain during first 6 weeks of treatment and health related quality of life at final visit

		Ciclosporin	Prednisolone	Mean difference (ciclosporin – prednisolone)	Adjusted mean difference[#]	95% CI	p
Pain scores (range 0-4)							
Week 1	n	47	38				
	Mean (SD)	1.98 (1.0)	1.84 (1.2)				
Week 2	n	46	37				
	Mean (SD)	1.74 (1.1)	1.69 (1.3)				
Week 3	n	46	36				
	Mean (SD)	1.59 (1.0)	1.48 (1.2)				
Week 4	n	45	35				
	Mean (SD)	1.34 (1.2)	1.50 (1.2)				
Week 5	n	46	34				
	Mean (SD)	1.22 (1.1)	1.49 (1.3)				
Week 6	n	45	32				
	Mean (SD)	1.10 (1.0)	1.49 (1.3)				
AUC weeks 1-6	n	45	32				
(0 to 20)	Mean (SD)	7.5 (4.8)	7.9 (5.6)	-0.40	-0.48	-2.82, 1.87	0.685

DLQI (range 0 – 30) (high score = worse)							
Baseline	n	58	53				
	Mean (SD)	10.3 (7.3)	13.2 (9.0)				
6 weeks	n	43	38				
	Mean (SD)	6.2 (6.1)	9.1 (8.2)				
Final visit	n	38	28				
	Mean (SD)	4.8 (6.8)	6.3 (7.6)	-1.5	-0.45	-3.46, 2.56	0.767
EQ-5D-3L (range-0.594 to 1.000) (low scores = worse)							
Baseline	n	56	52				
	Mean (SD)	0.51 (0.35)	0.44 (0.38)				
6 weeks	n	45	40				
	Mean (SD)	0.65 (0.30)	0.54 (0.38)				
Final visit	n	42	27				
	Mean (SD)	0.76 (0.30)	0.63 (0.41)	0.13	0.13	-0.02, 0.28	0.095
EQ-5D VAS (range 0 to 100) (low scores = worse)							
Baseline	n	57	53				
	Mean (SD)	62.6 (22.2)	61.4 (21.5)				
6 weeks	n	45	41				
	Mean (SD)	70.9 (16.0)	66.2 (25.1)				
Final visit	n	41	29				

	Mean (SD)	73.2 (20.5)	70.6 (22.3)	2.6	0.48	-9.32, 10.29	0.922

adjusted for baseline values and stratification variables (lesion size and presence of underlying disease).

Health-related quality of life

All health related quality of life scores improved during the period of the trial. No significant between-group differences were identified in DLQI, EQ-5D or EQ-5D VAS (See Table 8)

Time to recurrence

Analysis of recurrence was based on 52/53 (98%) of the lesions that had healed by 6 months. Follow-up for these patients ranged from 0 to 40.3 months depending on when they were recruited into the trial. Of those receiving ciclosporin, 8 patients (29.6%) had a recurrence, compared with 7 (28.0%) of the prednisolone-treated patients. There was no significant treatment effect in the time to the first recurrence (See *Table 3*)

Number of treatment failures

Treatment failure was documented in approximately half of the patients in each group (29/59, 49.2% in the ciclosporin group; 26/53, 49.1% in the prednisolone group; p = 0.88).

Adverse reactions

Overall 40 (67.8%) of participants in the ciclosporin group and 35 (66.0%) in the prednisolone group experienced at least one adverse reaction. Specific events that occurred in at least 3% of patients in either treatment group are presented in Table 8 (See *Appendix 28* full details of all adverse reactions).

Table 8: Specific adverse reactions occurring in ≥ 3% participants in either treatment group

Upper level classification	**Lower level classification**	**Ciclosporin (n = 59) n (%)**	**Prednisolone (n = 53) n (%)**
Blood and the lymphatic system disorders	Anaemia	2 (3.4)	0 (0.00)
	Leucocytosis	0 (0.0)	5 (9.4)
Endocrine disorders	Diabetes	0 (0.0)	3 (5.7)
Metabolism and nutrition disorders	Hyperglycaemia	0 (0.0)	5 (9.4)

Nervous system disorders	Tremor	5 (8.5)	2 (3.8)
	Headache	5 (8.5)	0 (0.0)
	Paraethesia	2 (3.4)	0 (0.0)
	Euphoria	0 (0.0)	3 (5.7)
	Depression	1 (1.7)	2 (3.8)
Gastrointestinal disorders	Nausea	12 (20.3)	1 (1.9)
	Vomiting	4 (6.8)	0 (0.0)
	Diarrhoea	2 (3.4)	0 (0.0)
	Candidiasis	1 (1.7)	2 (3.8)
Cardiovascular disorders	Hypertension	10 (16.9)	4 (7.5)
	Oedema	0 (0.0)	2 (3.8)
Heptatobiliary disorders	Hepatic dysfunction	2 (3.4)	1 (1.9)
Skin and subcutaneous tissue disorders	Hypertrichosis	2 (3.4)	0 (0.0)
Musculoskeletal, connective tissue and bone disorders	Muscle cramps	2 (3.4)	0 (0.0)
	Myalgia	2 (3.4)	1 (1.9)
	Arthralgia	2 (3.4)	0 (0.0)
Renal and urinary disorders	Renal dysfunction	18 (30.5)	1 (1.9)
General disorders	Serious infection (requiring hospitalisation or parenteral antibiotic)	0 (0.0)	6 (11.3)
	Other infection	4 (6.8)	5 (9.4)
	Fatigue	2 (3.4)	4 (7.5)
	Weight increase	1 (1.7)	4 (7.5)

Adverse reactions differed between the treatments as would be expected based on each drug's recognised tolerability profile. Differences of note include 5.7% patients developing diabetes and 9.4% developing hyperglycaemia in the prednisolone group versus none for either condition in the ciclosporin group. A higher number of participants in the prednisolone group developed a serious infection (11.3%), with no occurrence in the ciclosporin group, and disorders of the lymphatic system (9.4%), whilst this was less prevalent in the ciclosporin group (3.4%). Headache was reported by 8.5% patients in the ciclosporin group but none treated with prednisolone. Nausea, vomiting and diarrhoea were all more common in the ciclosporin group (20.3%, 6.8% and 3.4%, respectively) than in the prednisolone group (1.9%, 0.0% and 0.0%, respectively). Renal dysfunction was also notably more common in the ciclosporin than the prednisolone group (30.5% versus 1.9%, respectively).

There were nine serious adverse reactions (SARs) recorded throughout the trial; two in the ciclosporin group and seven in the prednisolone group. The SARs in the ciclosporin-treated patients were a ruptured abdominal aortic aneurysm and a case of acute kidney injury with elevated serum creatinine (212 µmol/L). Both of these events were considered 'possibly related' to study treatment. In the prednisolone group, the SARs were one case of bowel perforation (probably related), five serious infections (requiring hospitalisation or parenteral antibiotic; two probably related and three possibly related) and one other infection (possibly related). One of the serious infections (septicaemia gram negative bacilli) resulted in death.

Cost analysis

Use of resources and costs were similar when comparing groups with two exceptions. The cost of treatment drugs was significantly higher for the ciclosporin group, as would be anticipated. There was a significant increase in time in hospital in the prednisolone group. Of the six patients with greater than 10 days admission during the study, five received prednisolone (54, 48, 46, 38 and 16 days) and one received ciclosporin (14 days).

Table 9: Analysis of resource use and costs

Resources	Ciclosporin (N=47)	Prednisolone (N=40)	Diff.	95%CI	p

	Mean	SD	Mean	SD			
GP visits	2.91	7.56	1.53	2.29	-1.39	(-3.97 to 0.47)	0.24
GP home visits	0.02	0.15	0.43	1.52	0.40	(-0.04 to 0.85)	0.26
Practice Nurse contacts	5.85	13.03	6.30	14.26	0.45	(-5.11 to 6.33)	0.88
District Nurse visits	3.91	11.92	6.48	24.78	2.56	(-4.39 to 12.27)	0.60
Outpatient visits	8.30	14.32	5.15	9.03	-3.15	(-8.17 to 1.61)	0.22
Inpatient (days)	0.53	2.30	5.48	14.26	4.94	(0.34 to 9.55)	0.04
Cost (NHS, 2012)							
Cost (no drugs)	£1686	£2420	£2935	£5102	£1250	(£-330 to £3046)	0.171
Drug cost	£965	£442	£328	£198	£-638	(£-779 to £-498)	<0.001
Total cost	£2651	£2465	£3263	£5105	612	(£-971 to £2405)	0.487

Discussion

Given the lack of good quality published data relating to the management of PG, there is a clear need for trials that are robust in design and relevant to clinical practice. On this basis, STOP GAP has, for the first time we are aware of, compared two of the most commonly used treatments in an RCT setting. Patients were recruited from a range of centres around the UK in order to ensure that the sample was representative. The trial procedure was designed to reflect normal clinical practice as closely as possible; with dosing adjusted according to clinician opinion. The data collected in this trial included assessments by clinicians and patients, as well as independent analysis of digital photographs, thus providing both objective and subjective measures of treatment success.

Since starting the STOP GAP trial, an expert opinion consensus document considering safety, efficacy and cost placed prednisolone as preferred treatment and ciclosporin as second-

ranked therapy amongst the many suggested interventions[713]. Nonetheless, prior to the design of this trial, various studies had reported high proportions of patients with PG achieving complete responses with ciclosporin treatment [707-710] which lead STOP GAP to test the hypothesis that ciclosporin was superior to prednisolone for the treatment of PG. However, the study revealed no difference between the two treatments across a range of efficacy outcomes, and there were narrow confidence intervals around those lack of differences, suggesting that the study was large enough to exclude clinically important differences that might have been missed. Perhaps the most important finding was that contrary to the commonly-held impression that these drugs are very efficacious in PG, fewer than half of the ulcers were healed by either treatment after prolonged therapy The large number of patients who switched treatments or added in topical medications during the trial further reflects this generally poor treatment response.

In this study, approximately two thirds of patients reported adverse reactions in both treatment groups; 12% of whom experienced at least one serious event (two randomised to ciclosporin and seven randomised to prednisolone). This information is important given that less than half of the participants achieved complete healing. Though the overall rates of adverse reactions were almost identical in the two groups, the side effects observed differed in-line with the known side-effect profiles of these drugs. More serious adverse reactions including infections were reported in the prednisolone group than the ciclosporin group. It is worth considering that the median time to healing of almost 4 months indicates that patients would need to be exposed to the treatment-associated risks over long periods of time.

Analysis of resource use and cost data appears to support the clinical findings, in that these provide no strong rationale for ranking one treatment before another, rather informed decision making should reflect awareness of the side-effect profiles and patient preference.

Initial pilot work for the STOP GAP trial included discussions with patients as to the most important outcomes to be included in PG trials. As a result, the degree of exudation was added to the PG severity assessment scale proposed by Foss[717], and pain was recorded daily for the first 6 weeks. Although patients reported pain as being the most important symptom associated with PG, the pain scores reported here were relatively low (approximate mean of

1 on a 0–5 scale). It is possible that these scores were confounded by concurrent use of pain-killing medication.

Study limitations

The study was observer-blind rather than double-blind in design, owing to logistic and methodological difficulties in blinding treatment allocation within the resources available to the trial. However, every effort was made to capture the primary outcome in a blinded fashion, and all secondary analyses were supportive of this main analysis; suggesting that minimal bias was introduced by this approach. Power to explore the impact on quality-of-life was limited due to missing data from postal questionnaires and so a full cost-effectiveness analysis was not possible and data were presented descriptively in order to guide clinically decision making.

Given the lack of a placebo or no treatment third arm in this study, it is possible that neither drug is effective in treating PG. However, such a notion is not consistent with clinical experiences of some cases of rapid healing once systemic therapy is introduced. The costs, risks and logistics of developing appropriate placebos for two potent active systemic treatments with flexible dosing precluded the use of placebos or over encapsulation in our study.

The eventual sample size of 121 patients was slightly smaller than the 140 that had been planned. Funding for the study was available for a 5-year period, by the end of which the recruitment target had nearly been reached, and the decision was made to close the trial on the scheduled end date.

Generalizability

This was a pragmatic RCT that recruited in multiple secondary hospitals throughout the UK. As such, it is likely that the study has relatively good external validity and the patients recruited into this trial are reflective of the kinds of patients who commonly present with PG in the UK.

Clinical Conclusions

The results from this trial suggest that the outcome for patients with PG requiring systemic therapy is likely to be similar whether prednisolone or ciclosporin is chosen as the first

agent. However, there are differences in side effect profiles, which should be considered on a case-by-case basis to select the optimal treatment. For example, a person with a previous history of infections such as recurrent cellulitis or infections associated with joint prostheses might be more suitable for ciclosporin, and a person with previous hypertension or borderline renal impairment might be more suitable for prednisolone.

Based on the data from this study, it might be expected that with ciclosporin or prednisolone monotherapy, approximately one in six patients should be healed at 6 weeks. The median time to healing is closer to 4 months; and over half of patients may not achieve resolution even after 6 months of treatment. This knowledge may help clinicians to measure treatment response, to manage patients' expectations, and to provide a baseline comparison for future cohorts or observational studies. The fact that almost 30% of patients experienced a recurrence indicates that this is a long-term condition, for which follow-up of patients and possible preventive maintenance therapy may be required.

Patient information resources on PG can now be updated on the basis of these trial findings in order to provide better quality information to patients with this painful and debilitating condition.