Mason, J.M. and Thomas, K.S. and Ormerod, A.D. and Craig, F.E. and Mitchell, E. and Norrie, J. and Williams, H.C. (2017) Ciclosporin compared to prednisolone therapy for patients with pyoderma gangrenosum: cost-effectiveness analysis of the STOP GAP trial. British Journal of Dermatology . ISSN 1365-2133

BJD
British Journal of Dermatology

Ciclosporin compared with prednisolone therapy for patients with pyoderma gangrenosum: cost-effectiveness analysis of the STOP GAP trial

J.M. Mason,[1] K.S. Thomas,[2] A.D. Ormerod,[3] F.E. Craig,[4] E. Mitchell,[5] J. Norrie,[6] and H.C. Williams[2] on behalf of the U.K. Dermatology Clinical Trials Network's STOP GAP team

[1]*Warwick Medical School, University of Warwick, Coventry, CV4 7AL, U.K.*
[2]*Centre of Evidence Based Dermatology, University of Nottingham, NG7 2NR, U.K.*
[3]*Division of Applied Medicine, Aberdeen University, Aberdeen, AB24 2ZD, U.K.*
[4]*Department of Dermatology, NHS Forth Valley, Stirling, FK8 2AU, U.K.*
[5]*Nottingham Clinical Trials Unit, University of Nottingham, Nottingham, NG7 2UH, U.K.*
[6]*Edinburgh Clinical Trials Unit, Usher Institute, University of Edinburgh, EH16 4TU, U.K.*

Summary

Correspondence
James Mason.
E-mail: J.Mason@warwick.ac.uk

Accepted for publication
5 April 2017

Funding sources
This publication presents independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research Programme (RP-PG-0407-10177). The views expressed in this publication are those of the authors and not necessarily those of the U.K. National Health Service, the NIHR or the Department of Health. The study was developed with support from the U.K. Dermatology Clinical Trials Network (DCTN). The U.K. DCTN is grateful to the British Association of Dermatologists and the University of Nottingham for financial support of the Network.

Conflicts of interest
None declared.

DOI 10.1111/bjd.15561

Background Pyoderma gangrenosum (PG) is a painful, ulcerating skin disease with poor evidence for management. Prednisolone and ciclosporin are the most commonly used treatments, although not previously compared within a randomized controlled trial (RCT).
Objectives To compare the cost-effectiveness of ciclosporin and prednisolone-initiated treatment for patients with PG.
Methods Quality of life (QoL, EuroQoL five dimensions three level questionnaire, EQ-5D-3L) and resource data were collected as part of the STOP GAP trial: a multicentre, parallel-group, observer-blind RCT. Within-trial analysis used bivariate regression of costs and quality-adjusted life years (QALYs), with multiple imputation of missing data, informing a probabilistic assessment of incremental treatment cost-effectiveness from a health service perspective.
Results In the base case analysis, when compared with prednisolone, ciclosporin was cost-effective due to a reduction in costs [net cost: −£1160; 95% confidence interval (CI) −2991 to 672] and improvement in QoL (net QALYs: 0·055; 95% CI 0·018–0·093). However, this finding appears driven by a minority of patients with large lesions (≥ 20 cm^2) (net cost: −£5310; 95% CI −9729 to −891; net QALYs: 0·077; 95% CI 0·004–0·151). The incremental cost-effectiveness of ciclosporin for the majority of patients with smaller lesions was £23 374/QALY, although the estimate is imprecise: the probability of being cost-effective at a willingness-to-pay of £20 000/QALY was 43%.
Conclusions Consistent with the clinical findings of the STOP GAP trial, patients with small lesions should receive treatment guided by the side-effect profiles of the drugs and patient preference – neither strategy is clearly a preferred use of National Health Service resources. However, ciclosporin-initiated treatment may be more cost-effective for patients with large lesions.

What's already known about this topic?

- Pyoderma gangrenosum is characterized by severe, painful skin ulcers.
- Although prednisolone has been the main systemic treatment, ciclosporin has been used increasingly because of its perceived greater effectiveness and fewer side-effects.

- STOP GAP was a pragmatic randomized controlled trial comparing ciclosporin and prednisolone: clinical effectiveness was similar, but only 50% of ulcers had healed by 6 months on either drug and adverse events were common with both drugs.

What does this study add?

- For patients with small lesions (< 20 cm^2), neither treatment is clearly more cost-effective than the other.
- However, ciclosporin-initiated treatment may be the more cost-effective option in patients with large (≥ 20 cm^2) lesions.
- Decisions about treatment will continue to be informed primarily by patient preference, underlying comorbidities, and drug side-effect profiles (e.g. serious infections with prednisolone, hypertension and renal dysfunction with ciclosporin).

Pyoderma gangrenosum (PG) is a rare, inflammatory skin disease characterized by progressive and painful necrotizing ulcers. Typically, PG presents as a tender erythematous nodule or pustule, quickly progressing to a large, demarcated ulcer with purplish, undermined edges.[1] PG is associated with underlying systemic disease, and in particular with inflammatory bowel disease, arthritis and haematological malignancies.[2] Additionally it may develop following incidental or iatrogenic trauma.[3–5] Compared with mortality in the general population, PG is associated with a three fold increased risk of death;[6] its ulcers are characterized by debilitating pain and may require narcotic analgesia.[1,2,7]

PG is diagnosed clinically after excluding other diagnoses because there are no adequate diagnostic tests and histological findings are relatively nonspecific. No national or international guidelines address PG management, which currently includes a range of poorly evidenced topical and systemic treatment options including antibiotics, steroids, calcineurin inhibitors and immunosupressants.[8,9] Only one randomized controlled trial (RCT) of treatments has previously been reported in patients with PG: a study of 30 patients compared infliximab against placebo and showed benefit for infliximab at 2 weeks. However, infliximab is not a first-line treatment for this condition.[10]

Given the absence of high-quality evidence for the management of PG, the STOP GAP trial was designed to test whether treatment with ciclosporin was superior to prednisolone. In brief, STOP GAP was a multicentre, parallel-group, assessor-blind RCT, recruiting 112 adult patients, with outcomes assessed at baseline, 6 weeks and when the ulcer had healed (if within 6 months).[11,12] Groups were balanced at baseline. The primary end point of velocity of healing at 6 weeks was similar between groups [adjusted mean difference 0·003 cm^2 daily, 95% confidence interval (CI) −0·20 to 0·21; P = 0·97]; healing within 6 months was similar (ciclosporin 47·5%, prednisolone 47·2%; P = 0·84). Adverse reactions were similar (ciclosporin 67·8%, prednisolone 66·0%; P = 0·84), but serious adverse reactions may have been more common in the prednisolone group (ciclosporin 3%, prednisolone 13%; P = 0·082), in particular due to five serious infections that required hospitalization for parenteral antibiotics. Having found no difference for a range of objective and patient-reported outcomes, the trialists concluded that treatment decisions for individual patients should be guided by the different side-effect profiles of the two drugs and patient preference.

Economic analysis is intended to inform decision-makers about the value-for-money of treatment alternatives in a context where healthcare resources are limited and prioritization is informed (at least in part) by the efficient use of resources.[13] An economic analysis was designed integrally within the STOP GAP trial, following a prospective analysis plan, to provide robust evidence of cost-effectiveness to inform health service decision-making.

Patients and methods

A within-trial patient-level cost-effectiveness analysis was undertaken using data from the STOP GAP trial. The analysis was from the National Health Service (NHS) perspective; individual patient data collected within the STOP GAP trial included NHS treatment costs and health status, estimated as quality-adjusted life years (QALYs). Cost-effectiveness analysis captures the effect of treatment as changes in cost and QALYs. Because follow-up was limited to 24 weeks, no discounting of costs and benefits was applied. The analysis followed intention-to-treat principles, in which patients were included in the analysis according to the treatment allocated by randomization and irrespective of subsequent care.

Outcomes

Generic health-related quality of life (QoL) was assessed using the EuroQol (EQ) questionnaire: a patient-completed two-page questionnaire consisting of the EQ five dimensions three level questionnaire (EQ-5D-3L) descriptive system and the EQ visual analogue scale (EQ-VAS).[14,15] The EQ-5D-3L includes

five questions addressing mobility, self-care, usual activities, pain/discomfort and anxiety/depression), with each dimension assessed at three levels: no problems, some problems and extreme problems. EQ-5D scores were converted to health status scores using the U.K. time-trade-off value set recommended by the EuroQol group,[16] providing a single health-related index including 0 (death) and 1 (perfect health), where negative scores are possible for some health states. Patients who died during the study were subsequently scored 0 at later scheduled follow-up visits. The EQ-VAS reports self-rated health on a vertical VAS where 100 denotes 'best imaginable health state' and 0 denotes 'worst imaginable health state'. Additionally, the Dermatology Life Quality Index (DLQI) was recorded as a disease-specific measure: DLQI asks patients 10 questions about how their skin condition has affected their life over the past week providing an aggregate score of range 0–30.[17] QoL measures were captured at baseline, 8 weeks and up to 24 weeks (unless healing had occurred).

Using the trapezoidal rule, the 'area under the curve' (AUC) of health status scores was calculated, providing patient-level QALY estimates for the cost-effectiveness analysis.[18] Similarly, EQ-VAS and DLQI scores were integrated discretely over time. Because AUC estimates were predicted to correlate with baseline scores (and thus potential baseline imbalances), AUC estimates were adjusted for baseline scores.[19]

Resource use and cost

Resource assessments occurred at 8 and 24 weeks (or when healed if earlier), during mandatory clinical visits, augmented by telephone calls with patients when clinics were missed. Use by patients of study drugs was recorded in the trial drug log. PG-related health service contacts were recorded during clinic visits using patient diaries as an *aide memoire*.

Patient costs were initially estimated in U.K. pounds sterling (2012) as the sum of resources used weighted by their reference costs. Study drugs were prescribed at varying doses and durations. Using national Prescribing Cost Analysis (PCA) data,[20] average costs per unit weight of therapeutic were determined and applied to patient drug-use records: ciclosporin £0·0242 mg^{-1} and prednisolone £0·0237 mg^{-1}. Costs for inpatient stays (in days) and outpatient visits were estimated using Hospital Episodes Statistics (HES) and the National Schedule of Reference Costs (NSRC).[21,22] National HES data were explored for inpatient episodes with a primary diagnosis of L88 Pyoderma gangrenosum: the five most common admission codes associated with that diagnosis were included, accounting for 83% of admissions. (HES admission codes are similar to U.S. Diagnosis-Related Group codes).[23] Per diem costs for each code were estimated from the NSRC, and a volume-weighted average cost per admission for PG was estimated as the cost per day. Inpatient stays were costed at £387 per day and outpatient visits at £130 per visit. General practitioner (GP) clinic and home visits, and practice and district nurse visits were costed using unit costs provided by the Personal Social Services Research Unit at the University of Kent;[24] community care contacts: GP (clinic) £43, GP (home) £110, practice nurse £14 and district nurse £39. Patient costs were subsequently updated to 2015 U.K. pounds sterling using the Hospital and Community Health Services index.[25]

Analysis

Follow-up of patients with PG within trials is problematic and some incompleteness of data was anticipated. Consequently, a base case analysis was constructed where missing data were imputed using multiple imputation. The base case analysis included the imputed within-trial incremental cost and QALYs gained, adjusted for trial baseline covariates. Supportive sensitivity analyses included only patients with complete data, thus exploring the impact of imputation.

The base case analysis used multiple imputation, conducted according to good practice guidance.[26,27] Multiple imputation provides unbiased estimates of treatment effect if data are missing at random: this assumption was explored in the data, for example by using logistic regression for missingness of costs and QALYs against baseline variables.[28] A regression model was used to generate multiple imputed datasets (or 'draws') for individual treatment groups, where missing values were predicted drawing on predictive covariates: these included age, sex, target lesion size (< 20 cm^2; ≥ 20 cm^2), presence or absence of underlying systemic disease. Outcome measures (at each time point) and costs contributed as both predictors and imputed variables. Each draw provided a complete dataset, which reflected the distributions and correlations between variables. Predictive mean matching was used to enhance the plausibility and robustness of imputed values, as normality could not be assumed. The imputation model used fully conditional (Markov chain Monte Carlo, MCMC) methods (multiple imputation by chained equations), which are appropriate when missing and correlated data occur in more than one variable. Each draw was analysed independently using bivariate regression (see below) and the estimates obtained were pooled to generate mean and variance estimates of costs and QALYs using Rubin's rule – a method that captures within and between variances for imputed samples.[29] To minimize the information loss of finite imputation sampling, 50 draws were taken, resulting in a loss of efficiency relative to infinite sampling of less than 0·5% in all imputed values. The distribution of imputed and observed values was compared visually and statistically to establish that imputation did not introduce bias into subsequent estimation.

Bivariate regression using seemingly unrelated regression equations was used to model incremental changes in costs and QALYs. This method respects the correlation of costs and outcomes within the data, and allows adjustment for a set of covariates, which can be explored and which improve precision.[30] Baseline QoL scores were included within all models to allow for potential baseline imbalances.[19] Joint distributions of costs and outcomes were generated using the

(nonparametric) bootstrap method, with replicates used to populate a cost-effectiveness plane. Bootstrapping jointly resamples costs and outcomes from the original data with replacement (maintaining the sample correlation structure) to create a new bootstrap sample from which a change in costs and QALYs are estimated. Using bias-corrected nonparametric bootstrapping, 5000 bootstraps were taken per model or draw evaluated. Mean estimates are reported with 95% credible intervals.

The incremental cost-effectiveness ratio (ICER) was estimated as the difference between treatments in mean total costs divided by the difference in mean total QALYs. Value-for-money is determined by comparing the ICER with a threshold value, typically the National Institute for Health and Care Excellence threshold for U.K. studies, of £20K–30K/QALY.[31] This represents the willingness-to-pay (WTP) for an additional QALY, and lower values than the threshold could be considered cost-effective for use in the NHS. Base case assumptions are explored using a range of supportive sensitivity analyses.

The net monetary benefit (NMB) of changing treatment was reported as a recalculation of the ICER at a range of thresholds of WTP for an additional QALY. The NMB succinctly describes the resource gain (or loss) when investing in a new treatment when resources can be used elsewhere at the same threshold. NMB estimates were used to generate cost-effectiveness acceptability curves (CEACs). The CEAC compares the likelihood that treatments are cost-effective as the WTP threshold varies.[30]

The expected value of perfect information (EVPI) is the upper limit of the value to a healthcare system of further research to eliminate uncertainty.[32] Findings from cost-effectiveness analyses remain uncertain because of the imperfect information they use. If a wrong adoption decision (to make a treatment available) is made this will bring with it costs in terms of health benefit forgone: the NMB framework allows this expected cost of uncertainty to be determined and guide whether further research should be conducted to eliminate uncertainty.

Analyses and modelling were undertaken in Stata 14 SE (StataCorp LLC; College Station, TX, U.S.A.). Reporting follows the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement.[33]

Results

Completeness of data

All 112 patients included within the trial primary analysis of effectiveness were included in the economic analysis (Table 1). Patients with complete EQ-5D assessments for all periods numbered 67 in total (60%). One patient died during the study and was subsequently scored 0 on visits that followed for both cost and EQ-5D score and is included in the analysis. There was a pattern of decreasing completeness as follow-up proceeded. Resource data was complete for 67 patients (60%). When considering both utilities and resource use, complete information was available for 55 patients (49%). Completeness of data was similar when comparing treatment arms. Missing values were imputed to provide a base case analysis including all 112 patients.

Table 1 Completeness of data by follow-up visit

	Ciclosporin (N = 59) n (%)	Prednisolone (N = 53) n (%)	Total (N = 112) n (%)
Health status			
EQ-5D baseline	56 (95)	52 (98)	108 (96)
EQ-5D 8 weeks	46 (78)	41 (77)	87 (78)
EQ-5D 24 weeks	41 (69)	29 (55)	70 (63)
Complete cases	39 (66)	28 (53)	67 (60)
Resource use			
Drug use	59 (100)	53 (100)	112 (100)
Health service 8 weeks	47 (80)	38 (72)	85 (76)
Health service 24 weeks	41 (69)	34 (64)	75 (67)
Complete cases	38 (64)	29 (55)	67 (60)
Health status and resource use			
Complete cases	32 (54)	23 (43)	55 (49)

EQ-5D, EuroQol five dimensions questionnaire.

Complete case estimates

Mean EQ-5D scores, resource use and cost data are reported by treatment in Table 2. Over the 24-week follow-up period there were no significant differences in QALYs when comparing treatments. Differences in resource use comparing groups were not statistically significant at any time point, although there is a suggestion of greater inpatient usage by patients on prednisolone. Time as an inpatient was recorded for only 10 patients: total durations were 14 and 5 days for ciclosporin and 54, 48, 46, 38, 16, 7, 6 and 1 for prednisolone. Patterns of resource use were costed using national reference values (see Patients and methods: Resource use and cost). Although costs for patients receiving ciclosporin were less over 24 weeks, the decrease was not statistically significant: −£1046 (95% CI: −£3534 to £1341) (see Table 3).

Cost-effectiveness analysis

The joint distribution of incremental cost and outcome for the base case analysis is shown graphically in Figure 1 (see also Table 3). Patients allocated to ciclosporin (compared with prednisolone) experienced a modest average increase in QoL (0·055 QALYs; 95% CI 0·018–0·093) over 24 weeks. Health costs were lower for patients receiving ciclosporin but the difference was imprecise: −£1160; 95% CI: −2991 to 672); cost differences were predominantly driven by differences in hospitalization (Table 2). The joint distribution of cost and outcome is summarized within the NMB metric. Using a WTP criterion of less than £20 000 per QALY gained, the NMB associated with ciclosporin-initiated therapy was positive £2263 (95% CI: 216–4311). Thus, the base case analysis

Table 2 Health status, resource use and cost (complete cases)

	Ciclosporin (C) Mean (SD)	Prednisolone (Pr) Mean (SD)	(C)−(Pr)[a] Mean (95% CI)
Health status			
EQ-5D baseline	0·51 (0·35)	0·44 (0·38)	0·08 (−0·06 to 0·22)
EQ-5D 8 weeks	0·65 (0·30)	0·53 (0·39)	0·12 (−0·03 to 0·27)
EQ-5D 24 weeks	0·80 (0·22)	0·66 (0·38)	0·15 (−0·01 to 0·30)
EQ-5D AUC	0·33 (0·08)	0·29 (0·15)	0·04 (−0·02 to 0·10)
Resource use[b]			
Drug use (g)[c]	45·6 (23·3)	8·5 (5·1)	
NHS contacts (0–8 weeks)			
GP clinic visits	2·34 (7·32)	1·11 (1·66)	1·24 (−0·92 to 3·39)
GP home visits	0·02 (0·15)	0·21 (0·62)	−0·19 (−0·39 to 0·01)
Practice nurse visits	3·57 (8·54)	3·82 (7·60)	−0·24 (−3·68 to 3·19)
District nurse visits	1·53 (4·23)	2·42 (7·09)	−0·89 (−3·45 to 1·67)
Outpatient visits	4·55 (8·24)	3·34 (6·71)	1·21 (−1·97 to 4·39)
Inpatient days	0·53 (2·30)	3·86 (10·70)	−3·33 (−6·84 to 0·18)
NHS contacts (9–24 weeks)			
GP clinic visits	0·66 (1·68)	0·56 (1·21)	0·10 (−0·56 to 0·76)
GP home visits	0·00 (0·00)	0·26 (1·54)	−0·26 (−0·78 to 0·25)
Practice nurse visits	2·61 (7·65)	3·15 (8·28)	−0·54 (−4·17 to 3·10)
District nurse visits	2·73 (8·78)	5·12 (20·40)	−2·39 (−9·75 to 4·98)
Outpatient visits	4·29 (9·44)	2·15 (3·45)	2·15 (−0·97 to 5·26)
Inpatient days	0·00 (0·00)	2·21 (7·65)	−2·21 (−4·78 to 0·36)
Cost			
Drug cost (0–24 weeks)	1211 (618)	222 (132)	989 (828 to 1151)
Care cost (0–8 weeks)	1151 (1869)	2344 (4816)	−1193 (−2814 to 429)
Care cost (9–24 weeks)	841 (1585)	1587 (3652)	−746 (−2066 to 574)

AUC, area under the curve; CI, confidence interval; EQ-5D, EuroQol five dimensions questionnaire; GP, general practitioner; NHS, National Health Service. [a]Ordinary least squares regression-estimated means and 95% CIs. [b]Resource use has different missing values in the two periods: overall resource use is not a simple sum of these items. [c]Average (mean) weight of allocated study drug.

suggests NHS resources would be better directed to ciclosporin- than prednisolone-initiated therapy in terms of cost-effectiveness. This finding is echoed in the cost-effectiveness acceptability curve, which expresses the NMB finding as a probability (Fig. 2: all patients). The likelihood that ciclosporin-initiated treatment is cost-effective is 98·5% given a WTP criterion of less than £20 000/QALY gained (Table 3).

Sensitivity analyses

Comparing mean costs and QoL estimates using different modelling assumptions supported the base case finding (Table 3). The qualitative similarity of imputed and complete case estimates supports the validity of the imputation process and assumptions.

Subgroup analyses

There was no interaction between treatment effect and baseline covariates except in the case of index lesion size. The 40 patients (36%) recruited with large lesions ($\geq$ 20 cm^2) experienced a different pattern of costs from the patients with smaller lesions (see Fig. 3 and Table 3).

Ciclosporin-initiated treatment markedly lowered costs for patients presenting with large lesions (−£5310, 95% CI −9729 to −891), but not patients with smaller lesions (£1007, 95% CI −269 to 2283). These differences were driven by the pattern of hospitalization, which predominantly occurred in patients receiving prednisolone and may be linked to the occurrence of serious adverse events.[12]

For patients presenting with large lesions, ciclosporin-initiated treatment appears to be a cost-effective strategy (Figs 2 and 3: index lesion $\geq$ 20 cm^2). However, for patients presenting with smaller lesions, for ciclosporin-initiated treatment, cost-effectiveness (£23 374/QALY) is uncertain with the 95% confidence region including preference for either treatment; consequently neither strategy is clearly a preferred use of NHS resources for patients with smaller lesions (Figs 2 and 3: index lesion $<$ 20 cm^2).

Value of further research

An EVPI analysis was conducted to explore the value of reducing uncertainty about the cost-effectiveness of ciclosporin- or prednisolone-initiated therapy. EVPI analysis at the patient level was conducted treating the two trial strata for index lesion size as independent trials. There is considerable certainty about the findings for the trial as a whole as well as for patients with large lesions: the remaining value of obtaining perfect information is low (Fig. 4). However, there remains

Table 3 Cost-effectiveness, cost/QALY (£, 2015): ciclosporin compared with prednisolone

	Incremental cost (95% CI)	Incremental QALYs (95% CI)	ICER[a] (95% CI)	p[b]	p[c]	NMB[b] (£)	NMB[c] (£)
Base case							
Imputed attributable costs and QALYs, covariate adjusted	−1160 (−2991 to 672)	0·055 (0·018 to 0·093)	Dominant	0·985	0·993	2263 (216 to 4311)	2815 (569 to 5061)
Sensitivity analyses							
1. Imputed attributable costs and QALYs, baseline EQ-5D adjusted	−813 (−2666 to 1039)	0·046 (0·010 to 0·083)	Dominant	0·948	0·970	1739 (−360 to 3839)	2202 (−101 to 4506)
2. Complete case attributable costs and QALYs, covariate adjusted	−1046 (−3534 to 1341)	0·047 (0·001 to 0·090)	Dominant	0·931	0·957	1978 (−616 to 4776)	2444 (−395 to 5381)
Subgroup analysis							
Base case: index lesion < 20 cm^2	1007 (−269 to 2283)	0·043 (0·001 to 0·085)	23 374 (undefined)	0·427	0·619	−146 (−1713 to 1420)	284 (−1559 to 2127)
Base case: index lesion ≥ 20 cm^2	−5310 (−9729 to −891)	0·077 (0·004 to 0·151)	Dominant	0·997	0·998	6858 (2039 to 11 678)	7632 (2464 to 12 801)

CI, confidence interval; EQ-5D, EuroQol five dimensions questionnaire; ICER, incremental cost-effectiveness ratio; NMB, net monetary benefit; QALY, quality-adjusted life year. [a]ICER: dominance indicates average costs were less and average benefit greater for ciclosporin-initiated therapy. [b]Probability cost-effective or NMB if willing to pay £20 000 per QALY gained. [c]Probability cost-effective or NMB if willing to pay £30 000/QALY.

considerable uncertainty about the management of smaller lesions and the probabilistic value of removing this uncertainty is significant. In 2011, 580 patients were hospitalized in England with a primary diagnosis of PG.[22] Within the trial, 64% of patients presented with smaller lesions: if generalizable then the treatment of 370 patients a year might be affected by greater certainty about treatment. Assuming the findings affected care for 10 years the population affected might number 3700; taking the upper bound of £336/patient, the population EVPI is £1·24 million, similar to the average cost of a nationally recruiting multicentre clinical trial (undiscounted costs).[34] However, taking into consideration the particular difficulties of conducting trials in this patient group, there may not be scope to conduct further definitive trial in patients with smaller lesions, and efforts might be better placed in investigating new interventions or topical treatments.

Other end-points

The three QoL measures used in the STOP GAP trial are reported in Table 4. EQ-VAS, like EQ-5D, provides 6-month approximations of quality-adjusted survival for each treatment group. EQ-VAS is scored 1–100: equivalent QALY scores are obtained by dividing by 100, although EQ-VAS is not recommended for QALY estimation within trials, as values are preference-rated rather than societal. DLQI is scored 0–30: the average score over the 6-month follow-up period is reported, using the AUC between the three time points to calculate the average. Being a disease-specific QoL measure, the DLQI is potentially more sensitive to change than a generic measure.

Between-group differences for all three imputed QoL measures are shown in Table 4, including unadjusted, baseline score-adjusted and full covariate-adjusted estimates. For each measure, there is a trend favouring ciclosporin.

Discussion

The STOP GAP trial featured a pragmatic multicentre design reflecting real-world clinical practice; thus, cost and outcome profiles are likely to reflect routine care in NHS settings. Patient-level data from the STOP GAP trial provide the most robust evidence to date on whether ciclosporin or prednisolone is cost-effective as first-line treatment for patients with PG. The base case analysis (using multiple imputation) found ciclosporin-initiated treatment to be cost-effective compared with prednisolone, primarily due to a modest net cost savings and improvement in QoL. However, this finding was driven by the performance of the subgroup of patients with large lesions. In the majority of patients with smaller lesions (< 20 cm^2) the estimated cost-effectiveness was too imprecise to differentiate between treatments. These findings are consistent with the results of the clinical trial, which found no difference between treatments in speed of healing, 6-month healing rates or recurrence, but a (near-significant) difference in the EQ-5D based on complete cases. Further, the trial reported a (near-significant) difference in more serious

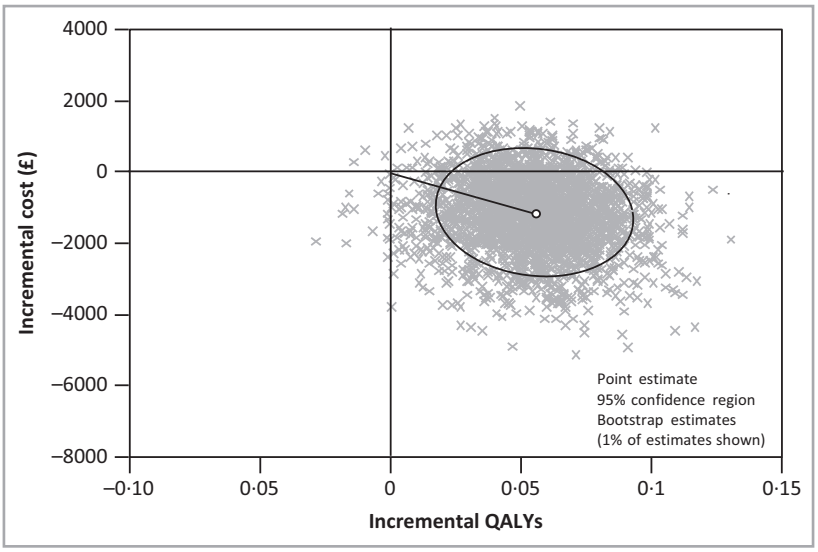

Fig 1. Cost-effectiveness plane: ciclosporin compared with prednisolone, base case analysis (cost per QALY, £, 2015). QALY, quality-adjusted life year.

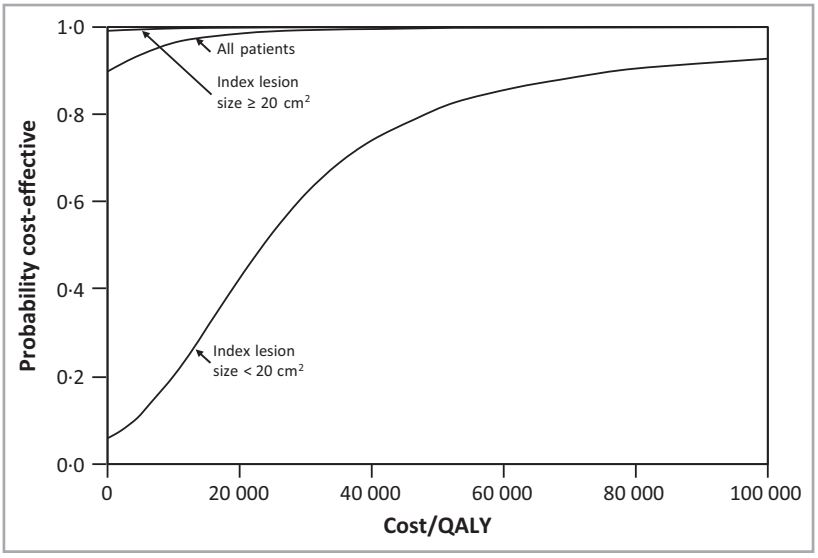

Fig 2. Cost-effectiveness acceptability curve: ciclosporin compared with prednisolone, base case and subgroup analyses.

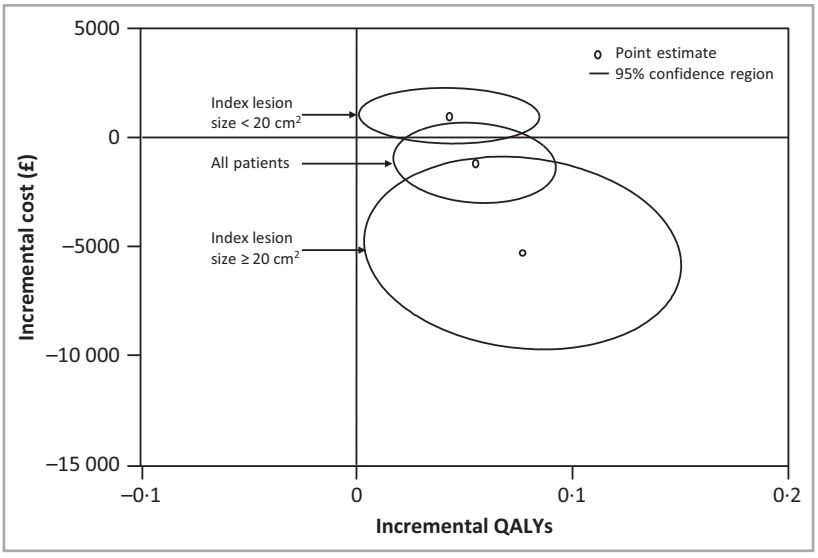

Fig 3. Cost-effectiveness plane: ciclosporin compared with prednisolone, base case and subgroup analyses (cost per QALY, £, 2015). QALY, quality-adjusted life year.

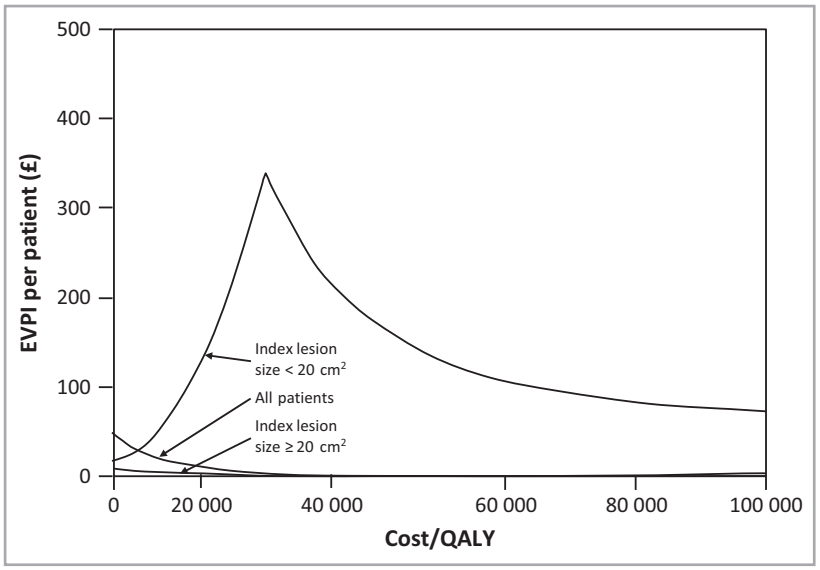

Fig 4. Expected value of perfect information: overall and subgroup analysis (£, 2015). QALY, quality-adjusted life year.

Table 4 EQ-5D, EQ-VAS and DLQI estimates: ciclosporin compared with prednisolone

AUC estimates	Mean difference	95% CI
EQ-5D imputed, unadjusted	0·061	0·016 to 0·105
EQ-5D imputed, baseline adjusted	0·046	0·010 to 0·083
EQ-5D imputed, covariate adjusted	0·055	0·018 to 0·093
Index lesion < 20 cm^2	0·043	0·001 to 0·085
Index lesion ≥ 20 cm^2	0·077	0·004 to 0·151
EQ-VAS imputed, unadjusted	2·214	−0·799 to 5·227
EQ-VAS imputed, baseline adjusted	2·051	−0·501 to 4·603
EQ-VAS imputed, covariate adjusted	2·556	−0·117 to 5·229
Index lesion < 20 cm^2	2·862	−0·431 to 6·155
Index lesion ≥ 20 cm^2	1·970	−2·315 to 6·255
DLQI imputed, unadjusted	−2·646	−4·796 to −0·497
DLQI imputed, baseline adjusted	−1·214	−2·685 to 0·258
DLQI imputed, covariate adjusted	−1·202	−2·719 to 0·316
Index lesion < 20 cm^2	−1·005	−2·795 to 0·785
Index lesion ≥ 20 cm^2	−1·566	−4·350 to 1·218

AUC, area under the curve; CI, confidence interval; DLQI, Dermatology Life Quality Index; EQ-5D, EuroQol five dimensions questionnaire; EQ-VAS, EQ visual analogue scale.

adverse reactions with prednisolone, particularly for infections, which might increase costs.

There are several caveats to these findings. PG is a rare disease and recruitment is problematic; although the largest trial of its kind, STOP GAP recruited only 112 patients from 39 U.K. hospitals over 3·5 years. Lesion size was a stratification variable within the trial randomization making the strata subgroups nested randomized controlled comparisons within the overall trial. Reflecting the subgroup patient numbers, differentiation of cost-effectiveness by lesion size would be strengthened by further evidence before prioritizing ciclosporin routinely for patients with large lesions.

Another weakness is the incompleteness of the data contributing to the economic analysis, a consequence of trying to maintain data quality over so many sites and time, and when the energy of trialists might focus on the clinical data. Exploring the consequences of imputation, the findings appear robust within a range of sensitivity analyses.

A final issue concerns the profile of costs and EQ-5D scores over time (Table 2). In the case of QoL, differences seem to be present and continuing beyond 24 weeks while costs have not clearly converged (accepting the different time periods involved). Thus, there might be a case to model extrapolated costs and outcomes beyond 24 weeks. In essence, modelling an extrapolated time horizon is appropriate when it (i) permits better characterization of the decision problem; (ii) allows evidence synthesis (e.g. from multiple trials); or, (iii) improves characterization of uncertainty.[35] While the within-trial analysis presented provides findings relevant to health service decision-makers, evidence is lacking on which to model plausible longer-term treatment and prognosis of patients with PG. Although the assumptions involved and quality of available trial data further limit the value of such modelling, any attempt would be likely to further emphasize the value of ciclosporin in preference to prednisolone in large lesions. The trial also captured relapses of symptoms beyond 24 weeks: these were infrequent and balanced between groups; thus, their inclusion would not influence the findings.

For patients presenting with smaller lesions the economic and clinical findings align in the sense that clinical outcomes are similar and the cost-effectiveness analysis is too imprecise to differentiate between these strategies. Uncertainty about the cost-effectiveness of ciclosporin- or prednisolone-initiated therapy for patients with small lesions is unlikely to be

resolved, at least within the NHS jurisdiction, given the challenges in conducting a further definitive trial, although EVPI suggests there might be value in doing so. It is likely in the health service setting that uncertainty about cost-effectiveness will be a secondary concern, with the clinical findings of similar effectiveness permitting continued use of either ciclosporin- or prednisolone-initiated therapy as the clinical context dictates. The subgroup analysis indicates ciclosporin may be preferred on cost-effectiveness grounds, particularly in patients with large lesions.

References

1 Brooklyn T, Dunnill G, Probert C. Diagnosis and treatment of pyoderma gangrenosum. *BMJ* 2006; **333**:181–4.
2 Binus AM, Qureshi AA, Li VW *et al.* Pyoderma gangrenosum: a retrospective review of patient characteristics, comorbidities and therapy in 103 patients. *Br J Dermatol* 2011; **165**:1244–50.
3 Su WP, Davis MD, Weenig RH *et al.* Pyoderma gangrenosum: clinicopathologic correlation and proposed diagnostic criteria. *Int J Dermatol* 2004; **43**:790–800.
4 Tremezaygues L, Schmaltz R, Vogt T *et al.* [Management of pyoderma gangrenosum. An update on clinical features, diagnosis and therapy]. *Hautarzt* 2010; **61**:345–53.
5 Zuo KJ, Fung E, Tredget EE *et al.* A systematic review of post-surgical pyoderma gangrenosum: identification of risk factors and proposed management strategy. *J Plast Reconstr Aesthet Surg* 2015; **68**:295–303.
6 Langan SM, Groves RW, Card TR *et al.* Incidence, mortality, and disease associations of pyoderma gangrenosum in the United Kingdom: a retrospective cohort study. *J Invest Dermatol* 2012; **132**:2166–70.
7 Miller J, Yentzer BA, Clark A *et al.* Pyoderma gangrenosum: a review and update on new therapies. *J Am Acad Dermatol* 2010; **62**:646–54.
8 British Association of Dermatologists. Patient information leaflet on pyoderma gangrenosum. (London, BAD: 2010, updated 2013). Available at: http://www.bad.org.uk/for-the-public/patient-information-leaflets/pyoderma-gangrenosum (last accessed 25 November 2016).
9 Reichrath J, Bens G, Bonowitz A *et al.* Treatment recommendations for pyoderma gangrenosum: an evidence-based review of the literature based on more than 350 patients. *J Am Acad Dermatol* 2005; **53**:273–83.
10 Brooklyn TN, Dunnill MG, Shetty A *et al.* Infliximab for the treatment of pyoderma gangrenosum: a randomised, double blind, placebo controlled trial. *Gut* 2006; **55**:505–9.
11 Craig FF, Thomas KS, Mitchell EJ *et al.*; UK Dermatology Clinical Trials Network's STOP GAP Trial Team. UK Dermatology Clinical Trials Network's STOP GAP trial (a multicentre trial of prednisolone versus ciclosporin for pyoderma gangrenosum): protocol for a randomised controlled trial. *Trials* 2012; **13**:51.
12 Ormerod AD, Thomas KS, Craig FE *et al.* UK Dermatology Clinical Trials Network's STOP GAP Team. Comparison of the two most commonly used treatments for pyoderma gangrenosum: results of the STOP GAP randomised controlled trial. *BMJ* 2015; **350**:h2958.
13 Drummond MF, Sculpher MJ, O'Brien B, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press, 2005.
14 Kind P, Dolan P, Gudex C, Williams A. Variations in population health status: results from a United Kingdom national questionnaire survey. *BMJ* 1998; **316**:736–41.
15 Dolan P, Gudex C, Kind P, Williams A. *A Social Tariff for EuroQol: Results from a UK general population survey*. Discussion paper 138. York, U.K.: Centre for Health Economics, University of York, 1995. Available at: http://www.york.ac.uk/media/che/documents/papers/discussionpapers/CHE%20Discussion%20Paper%20138.pdf (last accessed 17 April 2017).
16 Dolan P, Gudex C, Kind P, Williams A. The time trade-off method: results from a general population survey. *Health Econ* 1996; **5**:141–54.
17 Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI) – a simple practical measure for routine clinical use. *Clin Exp Dermatol* 1994; **19**:210–6.
18 Billingham L, Abrams KR, Jones DR. Methods for the analysis of quality-of-life and survival data in health technology assessment. *Health Technol Assess* 1999; **3**:1–152.
19 Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ* 2005; **14**:487–96.
20 NHS Prescription Services. *Prescription Cost Analysis (PCA) Data, England 2012*. Available at: http://content.digital.nhs.uk/catalogue/PUB10610 (accessed 25 November 2016).
21 Department of Health. *Reference Cost Collection: National schedule of reference costs – year 2011–12 – NHS trusts and NHS foundation trusts*. Available at: https://www.gov.uk/government/publications/nhs-reference-costs-financial-year-2011-to-2012 (accessed 25 November 2016).
22 NHS Digital. *Hospital Episode Statistics, Admitted Patient Care – England, 2011–12*. Available at: http://content.digital.nhs.uk/catalogue/PUB08288 (accessed 25 November 2016).
23 Krawzik K, Kenney A (eds). *DRG Desk Reference (ICD-10-CM) 2015*. Salt Lake City, UT: Optum 360; 2015.
24 Curtis L. *Unit Costs of Health and Social Care 2012*. Personal Social Services Research Unit. Canterbury, U.K.: University of Kent; 2012. Available at: http://www.pssru.ac.uk/project-pages/unit-costs/2012 (accessed 30 December 2016).
25 Curtis L, Burns A. *Unit Costs of Health and Social Care 2016*. Personal Social Services Research Unit. Canterbury, U.K.: University of Kent; 2016. Available at: http://www.pssru.ac.uk/project-pages/unit-costs/2016 (accessed 30 December 2016).
26 Sterne JA, White IR, Carlin JB *et al.* Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009; **338**:b2393.
27 White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. *BMJ* 2011; **342**:d40.
28 Faria R, Gomes M, Epstein D, White IR. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. *Pharmacoeconomics* 2014; **32**:1157–70.
29 White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011; **30**:377–99.
30 Willan AR, Briggs AH, Hoch JS. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. *Health Econ* 2004; **13**:461–75.
31 National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*. PMG9. London: NICE, 2013. Available at: https://www.nice.org.uk/process/pmg9 (accessed 25 November 2016).
32 Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute for Clinical Excellence (NICE). *Lancet* 2002; **360**:711–5.
33 Husereau D, Drummond M, Petrou S *et al.* CHEERS Task Force. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *BMJ* 2013; **346**:f1049.

34 Raftery J, Young A, Stanton L *et al*. Clinical trial metadata: defining and extracting metadata on the design, conduct, results and costs of 125 randomised clinical trials funded by the National Institute for Health Research Health Technology Assessment programme. *Health Technol Assess* 2015; **19**:1–138.

35 Sculpher MJ, Claxton K, Drummond M, McCabe C. Whither trial-based economic evaluation for health care decision making? *Health Econ* 2006; **15**:677–87.