HEALTH TECHNOLOGY ASSESSMENT

VOLUME 19 ISSUE 102 DECEMBER 2015
ISSN 1366-5278

The clinical effectiveness and cost-effectiveness of STeroids Or Pentoxifylline for Alcoholic Hepatitis (STOPAH): a 2 × 2 factorial randomised controlled trial

Mark Thursz, Ewan Forrest, Paul Roderick, Christopher Day, Andrew Austin, John O'Grady, Stephen Ryder, Michael Allison, Dermot Gleeson, Anne McCune, David Patch, Mark Wright, Steven Masson, Paul Richardson, Luke Vale, Jane Mellor, Louise Stanton, Megan Bowers, Ian Ratcliffe, Nichola Downs, Scott Kirkman, Tara Homer and Laura Ternent

National Institute for Health Research

DOI 10.3310/hta191020

The clinical effectiveness and cost-effectiveness of STeroids Or Pentoxifylline for Alcoholic Hepatitis (STOPAH): a 2 × 2 factorial randomised controlled trial

Mark Thursz,[1*] Ewan Forrest,[2] Paul Roderick,[3] Christopher Day,[4] Andrew Austin,[5] John O'Grady,[6] Stephen Ryder,[7] Michael Allison,[8] Dermot Gleeson,[9] Anne McCune,[10] David Patch,[11] Mark Wright,[12] Steven Masson,[4] Paul Richardson,[13] Luke Vale,[14] Jane Mellor,[15] Louise Stanton,[15] Megan Bowers,[15] Ian Ratcliffe,[15] Nichola Downs,[15] Scott Kirkman,[14] Tara Homer[14] and Laura Ternent[14]

[1]Department of Surgery and Cancer, Imperial College London, London, UK
[2]Department of Gastroenterology, Glasgow Royal Infirmary, Glasgow, UK
[3]Primary Care & Population Sciences, University of Southampton, Southampton, UK
[4]Institute of Cellular Medicine, The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK
[5]Department of Gastroenterology, Derby Royal Hospital, Derby, UK
[6]Institute of Liver Studies, King's College Hospital, London, UK
[7]Department of Hepatology, Nottingham University Hospitals NHS Trust and National Institute for Health Research Biomedical Research Unit, Queens Medical Centre, Nottingham, UK
[8]Department of Hepatology, Addenbrookes Hospital, Cambridge, UK
[9]Department of Hepatology, Sheffield Teaching Hospitals Foundation Trust, Sheffield, UK
[10]Department of Hepatology, Bristol Royal Infirmary, Bristol, UK
[11]Sheila Sherlock Liver Centre, Royal Free Hospital, London, UK
[12]Department of Hepatology, University Hospital Southampton NHS Foundation Trust, Southampton, UK
[13]Department of Hepatology, Royal Liverpool Hospital, Liverpool, UK
[14]Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK
[15]Southampton Clinical Trials Unit, University of Southampton, Southampton, UK

*Corresponding author

Declared competing interests of authors: Mark Thursz has received fees for advisory boards and speaker engagements from Gilead, BMS, Abbvi, MSD, Jenssen and Abbott Laboratories. Paul Roderick has received grant support from Pfizer and is a member of the Health Services and Delivery Research Board. Michael Allison has received fees for advisory board engagements from Norgine and Luke Vale is a member of the Clinical Trials Board.

Published December 2015
DOI: 10.3310/hta191020

This report should be referenced as follows:

Thursz M, Forrest E, Roderick P, Day C, Austin A, O'Grady J, *et al*. The clinical effectiveness and cost-effectiveness of STeroids Or Pentoxifylline for Alcoholic Hepatitis (STOPAH): a 2 × 2 factorial randomised controlled trial. *Health Technol Assess* 2015;**19**(102).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (SciSearch®) and *Current Contents®*/ Clinical Medicine.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 5.116

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 08/14/44. The contractual start date was in July 2010. The draft report began editorial review in October 2014 and was accepted for publication in August 2015. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Abstract

The clinical effectiveness and cost-effectiveness of STeroids Or Pentoxifylline for Alcoholic Hepatitis (STOPAH): a 2 × 2 factorial randomised controlled trial

Mark Thursz,[1*] Ewan Forrest,[2] Paul Roderick,[3] Christopher Day,[4] Andrew Austin,[5] John O'Grady,[6] Stephen Ryder,[7] Michael Allison,[8] Dermot Gleeson,[9] Anne McCune,[10] David Patch,[11] Mark Wright,[12] Steven Masson,[4] Paul Richardson,[13] Luke Vale,[14] Jane Mellor,[15] Louise Stanton,[15] Megan Bowers,[15] Ian Ratcliffe,[15] Nichola Downs,[15] Scott Kirkman,[14] Tara Homer[14] and Laura Ternent[14]

[1]Department of Surgery and Cancer, Imperial College London, London, UK
[2]Department of Gastroenterology, Glasgow Royal Infirmary, Glasgow, UK
[3]Primary Care & Population Sciences, University of Southampton, Southampton, UK
[4]Institute of Cellular Medicine, The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK
[5]Department of Gastroenterology, Derby Royal Hospital, Derby, UK
[6]Institute of Liver Studies, King's College Hospital, London, UK
[7]Department of Hepatology, Nottingham University Hospitals NHS Trust and National Institute for Health Research Biomedical Research Unit, Queens Medical Centre, Nottingham, UK
[8]Department of Hepatology, Addenbrookes Hospital, Cambridge, UK
[9]Department of Hepatology, Sheffield Teaching Hospitals Foundation Trust, Sheffield, UK
[10]Department of Hepatology, Bristol Royal Infirmary, Bristol, UK
[11]Sheila Sherlock Liver Centre, Royal Free Hospital, London, UK
[12]Department of Hepatology, University Hospital Southampton NHS Foundation Trust, Southampton, UK
[13]Department of Hepatology, Royal Liverpool Hospital, Liverpool, UK
[14]Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK
[15]Southampton Clinical Trials Unit, University of Southampton, Southampton, UK

*Corresponding author m.thursz@imperial.ac.uk

Background: Alcoholic hepatitis (AH) is a distinct presentation of alcoholic liver disease arising in patients who have been drinking to excess for prolonged periods, which is characterised by jaundice and liver failure. Severe disease is associated with high short-term mortality. Prednisolone and pentoxifylline (PTX) are recommended in guidelines for treatment of severe AH, but trials supporting their use have given heterogeneous results and controversy persists about their benefit.

Objectives: The aim of the clinical effectiveness and cost-effectiveness of STeroids Or Pentoxifylline for Alcoholic Hepatitis trial was to resolve the clinical dilemma on the use of prednisolone or PTX.

Design: The trial was a randomised, double-blind, 2 × 2 factorial, multicentre design.

Setting: Sixty-five gastroenterology and hepatology inpatient units across the UK.

Participants: Patients with a clinical diagnosis of AH who had a Maddrey's discriminant function value of ≥ 32 were randomised into four arms: A, placebo/placebo; B, placebo/prednisolone; C, PTX/placebo; and D, PTX/prednisolone. Of the 5234 patients screened for the trial, 1103 were randomised and after withdrawals, 1053 were available for primary end-point analysis.

Interventions: Those allocated to prednisolone were given 40 mg daily for 28 days and those allocated to PTX were given 400 mg three times per day for 28 days.

Outcomes: The primary outcome measure was mortality at 28 days. Secondary outcome measures included mortality or liver transplant at 90 days and at 1 year. Rates of recidivism among survivors and the impact of recidivism on mortality were assessed.

Results: At 28 days, in arm A, 45 of 269 (16.7%) patients died; in arm B, 38 of 266 (14.3%) died; in arm C, 50 of 258 (19.4%) died; and in arm D, 35 of 260 (13.5%) died. For PTX, the odds ratio for 28-day mortality was 1.07 [95% confidence interval (CI) 0.77 to 1.40; $p = 0.686$)] and for prednisolone the odds ratio was 0.72 (95% CI 0.52 to 1.01; $p = 0.056$). In the logistic regression analysis, accounting for indices of disease severity and prognosis, the odds ratio for 28-day mortality in the prednisolone-treated group was 0.61 (95% CI 0.41 to 0.91; $p = 0.015$). At 90 days and 1 year there were no significant differences in mortality rates between the treatment groups. Serious infections occurred in 13% of patients treated with prednisolone compared with 7% of controls ($p = 0.002$). At the 90-day follow-up, 45% of patients reported being completely abstinent, 9% reported drinking within safety limits and 33% had an unknown level of alcohol consumption. At 1 year, 37% of patients reported being completely abstinent, 10% reported drinking within safety limits and 39% had an unknown level of alcohol consumption. Only 22% of patients had attended alcohol rehabilitation treatment at 90 days and 1 year.

Conclusions: We conclude that prednisolone reduces the risk of mortality at 28 days, but this benefit is not sustained beyond 28 days. PTX had no impact on survival. Future research should focus on interventions to promote abstinence and on treatments that suppress the hepatic inflammation without increasing susceptibility to infection.

Trial registration: This trial is registered as EudraCT 2009-013897-42 and Current Controlled Trials ISRCTN88782125.

Funding: This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 19, No. 102. See the NIHR Journals Library website for further project information. The NIHR Clinical Research Network provided research nurse support and the Imperial College Biomedical Research Centre also provided funding.

Contents

List of tables xiii

List of figures xvii

List of boxes xix

List of abbreviations xxi

Plain English summary xxiii

Scientific summary xxv

Chapter 1 Aims 1

Chapter 2 Background 3
Importance of the health problem to the NHS 3
Summary of current evidence 3
Corticosteroids 3
Pentoxifylline 3
Steroids and/or pentoxifylline 4

Chapter 3 Trial design and methods 5
Study design 5
Ethical approval and research governance 5
Changes to original protocol 5
Study setting and sample 7
Inclusion criteria 7
Exclusion criteria 7
Study interventions 7
Study procedures 8
Recruitment 8
Informed consent 8
Randomisation and concealment 8
Unblinding service 8
Data collection and management 9
Baseline 9
Follow-up 9
Outcome measures 9
Primary outcome 9
Secondary outcomes 9
Sample size 10
Stopping guidelines 11
Statistical analysis 12
Analysis populations 12
Preliminary analyses 12
Primary analyses 12

Secondary analyses 13
Sensitivity analyses 13
Subgroup analyses 14
Patient and public involvement 14

Chapter 4 Trial results **15**
Recruitment and randomisation 15
Flow of participants through the trial 15
Unblinding of patients 15
Characteristics of the study sample 19
Allocation of treatments 20
Treatment adherence 22
Primary outcome 22
Secondary outcomes 23
Day 90, year 1 and overall survival 23
Prognostic scores 26
Other predictive factors of mortality 31
Serious adverse events and deaths 33
Alcohol behaviour 33
Sensitivity analyses 44
Subgroup analyses 44

Chapter 5 Health economic evaluation **47**
Introduction 47
Part 1: eliciting health state utility values for cirrhosis owing to alcohol abuse and cirrhosis owing to viral hepatitis – a standard gamble exercise 47
Methods 47
Results 50
Discussion 50
Conclusion 52
Part 2: economic evaluation 53
Methods overview 53
Within-trial analysis 60
Sensitivity analysis 60
Results 61
Model-based analysis 70
Modelling results 73
Summary 76

Chapter 6 Discussion and conclusions **77**
Summary of findings 77
Therapeutic effects 77
Adverse events 77
Prognostic scores 77
Health economics 78
Strengths and weaknesses of the study 78
Comparison with results of other studies 79
Implications for services and future research 80

Acknowledgements **83**

References **85**

Appendix 1 Prognostic score formula 91

Appendix 2 Substudies 93

Appendix 3 Sensitivity analysis of the primary outcome 95

Appendix 4 Unit cost: intervention, initial admission and adverse events costs 97

Appendix 5 Average adverse event investigations per treatment arm 103

List of tables

TABLE 1 Summary of changes to original STOPAH trial protocol approved by the REC 5

TABLE 2 Estimated 28-day mortality rates for sample size 11

TABLE 3 Summary of patients unblinded during the STOPAH trial 18

TABLE 4 Baseline characteristics 19

TABLE 5 Allocation of treatment arms by stratification factors 21

TABLE 6 Observed 28-day mortality 22

TABLE 7 Primary analysis results: logistic regression analysis for 28-day mortality 22

TABLE 8 Observed 90-day mortality or liver transplant 23

TABLE 9 Logistic regression analysis for 90-day mortality 23

TABLE 10 Observed 1-year mortality or liver transplant 24

TABLE 11 Logistic regression analysis for 1-year mortality or liver transplant 24

TABLE 12 Univariate logistic regression analysis for prognostic scores 26

TABLE 13 Receiver operating characteristics curve analysis 26

TABLE 14 Prednisolone vs. no prednisolone logistic regression model results for 28-day, 90-day and 1-year mortality by factorial design and risk group (ITT population at day 28, day 90 and 1 year) 28

TABLE 15 Receiver operating characteristics analysis for Lille score (prednisolone-treated group only) 29

TABLE 16 Cox proportional hazards model for mortality up to 1 year by Lille categories (< 0.45 and ≥ 0.45) (prednisolone-treated group only) 31

TABLE 17 Univariate logistic regression analysis 32

TABLE 18 Multivariate logistic regression (after backward elimination) 34

TABLE 19 Multivariate logistic regression analysis after removal of outlying observations after backward elimination 35

TABLE 20 Causes of death 36

TABLE 21 Summary of the SAEs reported 37

TABLE 22 Summary of the events reported on the SAE form (sorted by % of total events) 39

TABLE 23 Incidence of infection 42

TABLE 24 Alcohol consumption and counselling at day 90 and 1 year 43

TABLE 25 Univariate logistic regression analysis to assess the effect of alcohol consumption status at day 90 on 1-year mortality 44

TABLE 26 Subgroup analyses for 28-day, 90-day and 1-year mortality by factorial design 45

TABLE 27 Study sample characteristics 50

TABLE 28 Health state valuations 52

TABLE 29 A 2 × 2 factorial design for the STOPAH study 53

TABLE 30 Unit costs: follow-up 57

TABLE 31 Follow-up resource use at 90 days: liver related 61

TABLE 32 Follow-up resource use at 90 days: non-liver related 62

TABLE 33 Follow-up resource use at 1 year: liver related 63

TABLE 34 Follow-up resource use at 1 year: non-liver related 63

TABLE 35 Utility values at discharge, 90 days and 1 year for survivors only 64

TABLE 36 A 2 × 2 factorial design: within-the-table cost-effectiveness analysis 65

TABLE 37 Incremental cost per additional survivor at 28 days 65

TABLE 38 Cost-effectiveness analysis: prednisolone vs. no prednisolone 66

TABLE 39 Cost-effectiveness analysis: PTX vs. no PTX 68

TABLE 40 Cost-effectiveness sensitivity analysis with the most costly 10% of patients removed from each arm 69

TABLE 41 Parameters used in the economic model 72

TABLE 42 Analysis at 1 year comparing the four possible treatment arms 73

TABLE 43 Lifetime (10-year) Markov results: undiscounted 73

TABLE 44 Lifetime (10-year) Markov results: discounted 74

TABLE 45 Deterministic sensitivity analysis with 1-year Markov results 75

TABLE 46 Glasgow Alcoholic Hepatitis Score 91

TABLE 47 Observed 28-day mortality (sensitivity analysis) 95

TABLE 48 Logistic regression analysis for 28-day mortality (sensitivity analysis) 95

TABLE 49 Unit cost: intervention, initial admission and adverse events costs 97

TABLE 50 Average adverse event investigations per treatment arm 103

List of figures

FIGURE 1 Number of patients randomised by site, ordered by number of months open to recruitment 16

FIGURE 2 Consolidated Standards of Reporting Trials flow chart 17

FIGURE 3 Prednisolone vs. no prednisolone 24

FIGURE 4 Pentoxifylline vs. no PTX 25

FIGURE 5 Individual treatment arms 25

FIGURE 6 Receiver operating characteristics curves for prognostic scores 27

FIGURE 7 Receiver operating characteristics curves to compare the Lille score (prednisolone-treated group only) 29

FIGURE 8 Kaplan–Meier graph for over survival by Lille score (< 0.45 and > 0.45) (prednisolone-treated group only) 31

FIGURE 9 Example of a SG 48

FIGURE 10 Cost-effectiveness acceptability curve: placebo vs. prednisolone vs. PTX vs. prednisolone and PTX 66

FIGURE 11 Cost-effectiveness scatterplot: prednisolone vs. no prednisolone at 28 days 67

FIGURE 12 Cost-effectiveness acceptability curve: prednisolone vs. no prednisolone at 28 days 67

FIGURE 13 Cost-effectiveness scatterplot: PTX vs. no PTX at 28 days 68

FIGURE 14 Cost-effectiveness acceptability curve: PTX vs. no PTX at 28 days 69

FIGURE 15 Economic model 71

FIGURE 16 One-year cycle PSA results 74

FIGURE 17 Lifetime (10-year) cycle PSA results 75

List of boxes

BOX 1 Description of health state used in SG exercise 49

List of abbreviations

AH	alcoholic hepatitis
ALD	alcohol-related liver disease
ALP	alkaline phosphatase
ALT	alanine aminotransferase
AST	aspartate aminotransferase
BNF	*British National Formulary*
CI	confidence interval
CONSORT	Consolidated Standards of Reporting Trials
CRF	case report form
DF	Maddrey's discriminant function
EASL	European Association for Study of the Liver
EQ-5D	European Quality of Life-5 Dimensions
EQ-5D-3L	European Quality of Life-5 Dimensions-3 Levels
FBC	full blood count
GAHS	Glasgow Alcoholic Hepatitis Score
GI	gastrointestinal
GP	general practitioner
ICER	incremental cost-effectiveness ratio
ICU	intensive care unit
IMP	investigational medicinal product
INR	international normalised ratio
ITT	intention to treat
LFT	liver function test
MELD	model for end-stage liver disease
NAC	*N*-acetylcysteine
OS	overall survival
PSA	probabilistic sensitivity analysis
PT	prothrombin time
PTX	pentoxifylline
QALD	quality-adjusted life-day
QALY	quality-adjusted life-year
QoL	quality of life
REC	Research Ethics Committee
ROC	receiver operating characteristic
SAE	serious adverse event
SCTU	Southampton Clinical Trials Unit
SG	standard gamble
STOPAH	The clinical effectiveness and cost-effectiveness of STeroids Or Pentoxifylline for Alcoholic Hepatitis
TENALEA	Trans European Network for clinical trials services
TMG	Trial Management Group
WBC	white blood cell

Plain English summary

What was the problem?

Alcoholic hepatitis is an inflammatory condition of the liver caused by prolonged excessive alcohol consumption. Alcoholic hepatitis results in jaundice and liver failure; in the severe form of disease death may occur in 30% of patients within the first month. Previous clinical trials have failed to conclusively identify a treatment for this condition.

What did we do?

This trial tested whether or not prednisolone or pentoxifylline (which suppress inflammation) reduced the number of deaths in patients with severe alcoholic hepatitis. Over 1100 patients in 65 hospitals were included.

What did we find?

The number of deaths in the first month was reduced among patients taking prednisolone, but the effect was relatively weak and did not persist. Pentoxifylline did not appear to confer any benefit at any time. Patients treated with prednisolone had serious infections twice as frequently (13% vs. 7%) as in those who did not receive this drug. The trial showed that certain laboratory (prothrombin time and levels of bilirubin, creatinine and urea) and clinical (age and presence of encephalopathy) values helped predict a poor outcome of alcoholic hepatitis.

After 1 year, half the patients had died, 37% had stopped drinking alcohol and 20% attended an alcohol rehabilitation programme.

What does this mean?

Prednisolone has a modest effect and could be used to treat alcoholic hepatitis. In the longer term, more effort needs to be made to get patients into rehabilitation programmes. New treatments are required to suppress the inflammation in the liver without increasing the risk of infection.

Scientific summary

Background

Alcoholic hepatitis (AH) is a florid manifestation of alcohol-related liver disease with a 1-month mortality of 20–30% for those with severe disease [Maddrey's discriminant function (DF) ≥ 32]. It is characterised by the onset of jaundice in the context of excessive alcohol misuse and is a major cause of acute-on-chronic liver failure. Despite the serious nature of this illness there is controversy regarding the most effective medical therapy. Although there are numerous trials on this topic, there is still debate about the role of corticosteroids, with advocates citing significant improvement in the short- to medium-term mortality, while detractors raise concerns about the risks of sepsis and gastrointestinal (GI) haemorrhage. Pentoxifylline (PTX) has also been suggested as a treatment but meta-analyses to date have been inconclusive. Therefore, optimal medical treatment of AH is unclear and its management inconsistent.

Objectives

We aimed to evaluate whether or not prednisolone or PTX administered for 28 days improved short- and medium-term mortality in patients admitted with severe AH. We also aimed to assess their relative cost-effectiveness. The primary outcome measure was mortality at 28 days, with mortality at 90 days and 1 year being secondary outcomes. In addition, we aimed to assess the outcome relative to the Glasgow Alcoholic Hepatitis Score (GAHS), the model for end-stage liver disease (MELD) score and the Lille score, all of which have been advocated as methods of assessing disease severity and identifying patients for treatment.

Methods

Study design

The study design was a multicentre, double-blind, factorial (2 × 2) trial in which patients were randomised to one of four arms:

1. arm A – placebo/placebo
2. arm B – placebo/prednisolone
3. arm C – PTX/placebo
4. arm D – PTX/prednisolone.

Participants

Patients ≥ 18 years with a clinical diagnosis of AH on admission to hospital were considered for inclusion. Eligibility criteria were age ≥ 18 years; alcohol consumption > 80 g/day for males and 60 g/day for females to within 2 months of randomisation; serum bilirubin > 80 µmol/l; and DF ≥ 32. Key exclusion criteria were: duration of jaundice > 3 months; other causes of liver disease present; aspartate aminotransferase > 500 IU/ml or alanine aminotransferase > 300 IU/ml; previous entry to the study; previous use of prednisolone or PTX within 6 weeks of admission; renal failure (creatinine > 500 µmol/l or requiring renal replacement therapy); active GI bleeding; untreated sepsis; and patients requiring inotropic support. As the trial was conducted in 65 hospitals across the UK, and many of which do not have access to transjugular liver biopsy, it was decided not to make liver histology an entry criterion. Patients with GI bleeding, renal impairment or sepsis during the admission were allowed specific treatment for up to 7 days and randomised if the condition had been stabilised.

Main outcome measures

The primary end point was mortality at 28 days; this time point represents the end of the peak period of mortality for AH and is consistent with other trials in the field. Secondary end points looked at mortality or liver transplant at 90 days and 1 year, outcome relative to other prognostic scores (GAHS, MELD and Lille score), rates of recidivism, hospital readmission rates for liver or non-liver-related events, rates of GI haemorrhage and sepsis, and rates of new or recurrent renal failure (serum creatinine greater than 500 µmol/l or requiring renal support).

Study procedures

Potential patients for the trial were identified on admission to hospital with suspected AH. After clinical assessment, suitable candidates were given a patient information sheet and were given at least 24 hours to consider the study and ask questions, after which they (or their legal representatives) were asked to give written informed consent. Special arrangements were in place to ensure that the interests of patients with hepatic encephalopathy were protected. After enrolment, patients were registered via Trans European Network for clinical trials services, a web-based registration and randomisation system. If eligible for the study, patients were randomised to a study treatment arm, which was blinded to the site staff and the patient by means of a unique four-digit patient pack number.

Randomisation was performed using the following two stratification factors:

1. geographic region (28 in total)
2. risk group: either high or intermediate risk (high risk was defined as either sepsis or history of GI bleeding in the previous 7 days, or creatinine > 150 µmol/l, or any combination of the these; intermediate risk was defined as no sepsis and no history of GI bleeding in the previous 7 days, and creatinine ≤ 150 µmol/l).

All patients were given one capsule containing 400 mg of PTX (or identical placebo) three times per day, plus one capsule containing 40 mg of prednisolone (or identical placebo) once daily. Both medications were administered for 28 days.

Patients were evaluated on treatment days 7, 14, 21 and 28, and at each time point recordings were made of vital signs, World Health Organization performance status, concomitant medication and adverse events. Blood samples were taken for liver function tests, prothrombin time, full blood count, urea and creatinine. Patients were assessed for the presence of hepatic encephalopathy and the occurrence of GI bleed or sepsis in the past 7 days. If patients were discharged from hospital before the end of the 28-day treatment period, assessments were made at 28 days by telephone interview.

After discharge from hospital, patients were similarly evaluated at 90 days and at 1 year.

Statistical methods

A power calculation was performed to estimate sample sizes using the following parameters:

- power = 90%
- significance = 5%
- estimated mortality in placebo-treated group = 35%
- estimated mortality in prednisolone-treated group = 25%
- estimated mortality in PTX-treated group = 25%
- estimated mortality in prednisolone- and PTX- treated group = 17% (estimated assuming no interaction).

Based on a reduction in the 28-day mortality rate from 30% to 21%, a sample size of 513 per group of single agent versus no single agent was required. Thus, in total, the trial required 1026 patients. We allowed for a ≈ 10% withdrawal/lost to follow-up rate and therefore aimed to recruit 1200 patients to the study, with patients being evenly allocated to each treatment arm.

Analysis was on the basis of intention to treat. In order to determine the efficacy of prednisolone, the 28-day mortality rate in the prednisolone-treated group (arms B and D) was compared with the mortality rate in the control group (arms A and C). Similarly, PTX efficacy was assessed by comparing the 28-day mortality rate in the PTX-treated group (arms C and D) with the mortality rate in the control groups (arms A and B).

The impact of pre-treatment variables such as GI bleeding, sepsis or renal impairment on admission was estimated by adding these covariates to the logistic regression analysis. Mortality rates at 90 days and 1 year were compared using the same strategy.

Economic analysis

Within-trial cost-effectiveness and model-based cost–utility analyses were conducted. The economic evaluation was conducted from the perspective of the UK NHS and all costs were reported in 2014 UK pounds. The results of the cost-effectiveness analysis are reported as the incremental cost per additional survivor at 28 days and the results of the cost–utility analysis were reported as the incremental cost per quality-adjusted life-year (QALY) gained at 1 year and the patient's lifetime. QALYs were derived from responses to the European Quality of Life-5 Dimensions administered at baseline, 90 days and 1 year. Costs were based on the use of primary and secondary care services over the trial follow-up and cost using routine sources and study-specific estimates.

Results

Main trial results

Between January 2011 and February 2014, 5234 patients were screened for the trial and after applying eligibility criteria, 1103 patients were randomly allocated to the four treatment arms: 276 to placebo/placebo; 277 to placebo/prednisolone; 276 to PTX/placebo; and 274 to PTX/prednisolone. Patients were followed for 1 year or up until the time of their death with the exception of the final 223 patients recruited, when early cessation of follow-up meant that the trial was completed prior to them reaching this time point. The four arms were well matched with regard to their baseline characteristics and laboratory variables. At day 28, 16% of patients had died, 1% had been lost to follow-up and 2% had withdrawn from the study. At 90 days, 285 of 968 (29%) patients had died, 5% were lost to follow-up and 7% had withdrawn (including 4% owing to early cessation of follow-up). At 1 year, 421 of 747 (56%) patients had died or had a liver transplantation ($n = 3$), 8% were lost to follow-up and 24% of patients had withdrawn (including 20% owing to early cessation of follow-up).

At 28 days in the placebo/placebo arm, 45 of 269 (16.7%) patients had died, in the placebo/prednisolone arm 38 of 266 (14.3%) patients had died, in the PTX/placebo arm 50 of 258 (19.4%) patients had died and in the PTX/prednisolone arm 35 of 260 (13.5%) had died. There was no significant interaction between prednisolone and PTX. In the logistic regression analysis, the odds ratio for 28-day mortality in the PTX-treated group versus no-PTX-treated group was 1.07 [95% confidence interval (CI) 0.77 to 1.49; $p = 0.686$] and for the prednisolone-treated group versus no-prednisolone-treated group, the odds ratio was 0.72 (95% CI 0.52 to 1.01; $p = 0.056$). Although this result is of borderline significance, any effect of the steroid in the first month is rapidly lost: neither prednisolone nor PTX were found to influence mortality at 90 days and 1 year. At 90 days, 29.8% of patients treated with prednisolone had died, compared with 29.1% who did not receive this drug (odds ratio 1.02, 95% CI 0.77 to 1.35). At 90 days, 29.1% of patients treated with PTX had died, compared with 29.8% who did not receive this drug (odds ratio 0.97, 95% CI 0.73 to 1.28). At 1 year, 56.6% of patients treated with prednisolone had died, compared with 56.1% who did not receive this drug (odds ratio 1.01, 95% CI 0.76 to 1.35). At 1 year, 56.2% of patients treated with PTX had died, compared with 56.5% who did not receive this drug (odds ratio 0.99, 95% CI 0.74 to 1.33).

Baseline variables were assessed to establish whether or not prednisolone or PTX influenced (i.e. statistical significant at 5% level) 28-day mortality. In the univariate analysis, DF, GAHS, MELD, Lille score, age, encephalopathy, white blood cell (WBC) count, prothrombin ratio, serum bilirubin, urea and creatinine were all significant. In the multivariate analysis, age, encephalopathy, WBC count, urea, creatinine and prothrombin ratio remained significant. Using the multivariate logistic regression model to take into account these prognostic variables, we found that the odds ratio for 28-day mortality in the prednisolone-treated group versus no-prednisolone-treated group was 0.61 (95% CI 0.41 to 0.91; $p = 0.015$). However, the odds ratio for mortality in the prednisolone-treated group compared with the group not treated with prednisolone at 90 days was 1.00 (95% CI 0.73 to 1.36; $p = 0.976$) and at 1 year was 1.01 (95% CI 0.74 to 1.39; $p = 0.942$) remained statistically non-significant. This may indicate that minor variation in baseline factors influenced the result at 28 days. However, even if there is a benefit of steroid use at this early time point, it does not translate into long-term benefit.

In the univariate analysis, each of the four existing scoring systems provided prognostic information on patient survival at 28 days, 90 days and 1 year with p-values < 0.001. MELD had a slightly greater area under the receiver operating characteristic curve compared with the other scoring systems, although the GAHS had the higher odds ratio on logistic regression analysis at day 28. However, the prognostic value of each scoring system diminished with increased duration of follow-up. When the Lille score was studied as an indicator of corticosteroid response using cox proportional hazards regression modelling, Lille non-responders (≥ 0.45) had a higher overall mortality of 51.9% (70/135) compared with 24.6% (49/199) for responders (< 0.45) (hazard ratio 2.66, 95% CI 1.95 to 4.05; $p < 0.001$).

Serious adverse events (SAEs) were reported in 42% of patients with an equal distribution in each treatment arm and 29% of all SAEs resulted in deaths. Infection occurred in 74 of 547 (13.5%) patients in the prednisolone-treated group compared with 43 of 545 (7.9%) patients who did not receive prednisolone ($p = 0.003$). Acute kidney injury occurred in 9 of 546 (1.65%) patients who received PTX and 14 of 546 (2.56%) who did not receive this drug ($p = 0.4$).

Economic analysis

In terms of incremental cost per additional survivor at 28 days, prednisolone was the most likely treatment to be cost-effective. In the model-based analysis prednisolone was, on average, associated with an incremental cost per QALY that society may be willing to pay. No other treatment has an incremental cost per QALY of $<$ £30,000. However, the cautious assumptions made and limited longer-term data available means that there is considerable imprecision around estimates of cost-effectiveness, so that definitive conclusions cannot be drawn.

Conclusions

In this study we are able to show a reduction in the 28-day mortality in the prednisolone-treated group on logistic regression model analysis, but there was not clear evidence of benefit sustained beyond this point. In contrast there is no evidence of survival benefit for PTX either alone or in combination with steroids, at any measured time point.

The GAHS, MELD and Lille score performed similarly in assessing overall prognosis, but a high Lille score did not effectively identify a severe group of patients who may benefit from transplantation, as has been suggested previously.

Prednisolone use was associated with an increased rate of sepsis, which probably negated some of the benefits. There was no significant reduction in renal impairment with PTX use; however, there was a low rate of acute kidney injury in this trial, making a definitive conclusion about renal protection inappropriate.

Prednisolone was the most likely treatment to be considered cost-effective at 28 days and the results are suggestive of an incremental cost per QALY of < £30,000 at 1 year and over a lifetime, but remain tentative owing to limited longer-term data.

Recommendations for future research

1. Development of disease severity scores that allow identification of those most likely to benefit from corticosteroid use and to identify those for whom other treatment should be considered.
2. To address whether or not additional measures such as the addition of *N*-acetylcysteine or granulocyte colony stimulating factor to corticosteroids could reduce the rate of sepsis and improve outcome beyond 28 days.
3. To investigate methods to improve rates of abstinence after discharge from hospital.

Trial registration

This trial is registered as EudraCT 2009-013897-42 and ISRCTN88782125.

Funding

This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme. The NIHR Clinical Research Network provided research nurse support and the Imperial College Biomedical Research Centre also provided funding.

Chapter 1 Aims

The primary objective of this study was to determine whether or not pentoxifylline (PTX) or corticosteroids reduce the mortality associated with severe alcoholic hepatitis (AH) and thereby to resolve the ongoing dilemma about the use of these two agents in clinical practice. In order to avoid the controversies caused by underpowered studies in this field, we aimed to conduct a well-powered, definitive study.

Gastrointestinal (GI) bleeding, sepsis and renal impairment have previously been exclusion criteria in many AH treatment studies but patients with these complications have the highest mortality risk and therefore have potentially most to gain. They were therefore included in the trial.

Early treatment benefits may subsequently be lost owing to an increased incidence of sepsis in the medium term or recidivism in the longer term. Mortality at 28 days, 90 days and 1 year after treatment was therefore assessed.

Virtually all trials of therapeutic interventions in AH have used Maddrey's discriminant function (DF) to identify a group of patients with severe disease and an appreciable mortality risk. Recent research from the UK suggests that the Glasgow Alcoholic Hepatitis Score (GAHS) provides a more accurate prognosis.[1] In order to make this trial comparable to previous studies, we elected to keep the DF as an inclusion criterion, but we have compared the DF, GAHS, model for end-stage liver disease (MELD) and the Lille scores for their ability to predict mortality and/or response to therapy.

The effect of resumed alcohol abuse is thought to be one of the important predictors of mortality after the first month and we therefore aimed to collect data on recidivism and abstinence in order to assess the effect on patient outcomes.

Finally, we aimed to conduct within-trial and longer-term horizon economic evaluations of the two interventions.

Chapter 2 Background

Importance of the health problem to the NHS

Alcohol-related illness places an enormous burden on the NHS. It has been estimated by the Royal College of Physicians that the in-patient costs in 1998–9 arising from the consequences of alcohol misuse were as high as £2.9B.[2] Overall, alcohol-related deaths have more than doubled since 1979,[3] and in Scotland, they increased by 236% between 1980 and 2002.[4] Throughout the UK, deaths from cirrhosis rose dramatically between 1987 and 1991.[5] Alcohol-related liver disease (ALD) accounts for the majority of alcohol-related deaths in the UK.[6] While many patients presenting with ALD will have cirrhosis, as many as 60% will have evidence of alcohol-related hepatitis.[7] AH is the most florid manifestation of ALD, but is potentially reversible. However, the short-term mortality of AH is particularly high among those with indicators of severe disease. The 28-day mortality of patients who have a DF ≥ 32 is 20–30% and historically up to 40%.[8–10] The 28-day mortality of patients who have a GAHS of ≥ 9 is approximately 60%[1] (see *Appendix 1* for description of DF and GAHS). AH affects a relatively young population (average age 50 years; patients may present in their twenties to thirties). Despite the increasing prevalence and the severity of this disease there is no consistency in its management. Considerable controversy exists, especially regarding the use of corticosteroids.

Summary of current evidence

Corticosteroids

Since 1971 there have been 13 randomised studies and four meta-analyses investigating the role of corticosteroid therapy for this condition.[11,12] Despite this apparent wealth of evidence, controversy persists. There remains deep division with regard to the use of corticosteroids. Advocates of the treatment cite significant improvement in the short- to medium-term mortality, while detractors cite the risks of sepsis and GI haemorrhage with corticosteroid therapy. Many of the published studies have been plagued by widely varying inclusion and exclusion criteria. The largest placebo-controlled study[13] treated 90 patients and found no benefit with prednisolone compared with a similar placebo-treated group. This study was hampered by the inclusion of patients with both moderate and severe AH, as well as end-stage ALD. In the only study to require histological confirmation of AH in all patients, prednisolone was associated with a short-term improvement in mortality in patients, although this benefit was not apparent after 2 years.[14,15] On review of the published studies, none of these reached an adequate statistical power to make a statement with 80% confidence. The most recent meta-analysis of all of the available trials[16] demonstrated that, although there was a trend of benefit with steroids, the results were not statistically significant ($p = 0.2$). However, a reanalysis of the three largest studies indicted a significant benefit from corticosteroids.[17] In this meta-analysis, patients with a DF ≥ 32 treated with prednisolone had 28-day mortality of 15%, compared with mortality of 35% among placebo-treated patients ($p = 0.001$).

Pentoxifylline

Pentoxifylline has also been studied in the treatment of AH.[18–20] It is believed to act, in part, by inhibiting the synthesis of the proinflammatory cytokine tumour necrosis factor alpha.[20] There has been one randomised controlled trial[18] that showed significant benefit. In this study, 100 patients, all with a DF > 32, were enrolled. PTX was administered for 4 weeks, at a dose of 400 mg three times per day. In the PTX group, 12 of 49 patients (24.5%) died, compared with 24 of 52 (46.1%) in the placebo group during the index hospitalisation ($p = 0.037$). The principal benefit for the agent appeared to be a reduction in deaths attributed to hepatorenal syndrome. However, other studies found contrasting results and meta-analysis has failed to show any significant benefit of this drug.[19–21] Although published evidence for the use of PTX is inconclusive, the drug is widely used as an alternative to steroids, particularly in the USA, and further evaluation in a clinical trial is, therefore, clearly warranted.[22]

Steroids and/or pentoxifylline

Combinations of steroids and PTX have been evaluated in clinical trials, using steroids alone as the control arm; however, the combination does not appear to have any benefit over steroids alone.[23,24] Three small studies[25–27] have compared steroids directly with PTX but the results were inconsistent. One study[28] explored the use of PTX as a rescue therapy once steroids had failed to improve liver function tests (LFTs) but this strategy did not influence survival.

Chapter 3 Trial design and methods

Study design

The clinical effectiveness and cost-effectiveness of STeroids Or Pentoxifylline for Alcoholic Hepatitis (STOPAH) study was a pragmatic, multicentre, double-blind, factorial (2 × 2) trial to assess prednisolone and PTX in patients with severe AH. The factorial design meant that the two active treatments could be concurrently assessed when doubt exists over the efficacy of both medications.

The trial included an economic evaluation to assess which treatment is the most cost-effective, as well as the quality of life (QoL) and long-term prognosis in patients with AH (see *Chapter 5*). The main trial was also supplemented with a number of substudies (see *Appendix 2*). A description of the trial protocol has already been published.[29]

Ethical approval and research governance

Ethical approval for the study was given by Wales Research Ethics Committee (REC) 3 (formerly REC for Wales) on 27 April 2010 (reference number 09/MRE09/59). The trial was registered with the International Standard Randomised Controlled Trial Register under the reference number ISRCTN88782125.

Changes to original protocol

A summary of the changes made to the original protocol is given in *Table 1*.

TABLE 1 Summary of changes to original STOPAH trial protocol approved by the REC

Change to protocol	REC approval
'4 weeks' changed to '28 days'	19 July 2011
'Dosing instructions' section added	19 July 2011
Patient target added – 1200 randomised patients	19 July 2011
Liver transplant added as secondary end point at 90 days	19 July 2011
Past medication substudy text clarified	19 July 2011
Informed consent process clarified. Time for consideration of study changed to < 24 hours if appropriate	19 July 2011
Clarification that consent should be sought from 'incapacitated' patient, once able	19 July 2011
'TENALEA account creation and registration' sections added	19 July 2011
History of excess alcohol timeline added to ensure that patients have been drinking sufficiently heavily recently (to within 2 months of randomisation)	19 July 2011
Period of abstinence changed from '6 weeks' to '> 2 months'	19 July 2011
'Clinically apparent' jaundice added	19 July 2011
'6-month' timeline for previous use of study drugs changed to '6 weeks'	19 July 2011
Treatment of patients with renal impairment, sepsis or GI bleed clarified	19 July 2011
Prohibited drugs clarified, i.e. *N*-acetylcysteine and ketorolac, plus prescription of study drugs during the treatment phase	19 July 2011

continued

TABLE 1 Summary of changes to original STOPAH trial protocol approved by the REC (*continued*)

Change to protocol	REC approval
'Start of treatment and treatment breaks' section added	19 July 2011
'Renal impairment dose reduction' section added	19 July 2011
IMP manufacture, labelling and storage procedures clarified, and IMP shipment process expanded	19 July 2011
IMP 'temperature monitoring' section added, to explain storage requirements and use of WarmMark Temp Tags (Tela Temp Inc., Anaheim, CA, USA) for transit	19 July 2011
IMP 'documentation', 'dispensing procedures' and 'drug returns' sections updated	19 July 2011
'Patient adherence' section added	19 July 2011
Screening assessments updated/corrected	19 July 2011
'EDTA sample for DNA analysis' moved to baseline	19 July 2011
'Central pathology review' analysis section added	19 July 2011
Baseline assessments and timelines updated, including permitted 14-day window between screening and baseline visits	19 July 2011
'Inflammatory markers in serum analysis' section added	19 July 2011
'Analysis of genetic causes of AH' section added	19 July 2011
'Monocyte study' section added	19 July 2011
Treatment and follow-up assessments updated, including the permitted window around visits. Height removed from all assessments except screening	19 July 2011
WHO performance status added to all on treatment assessments	19 July 2011
Prior 48-hour window for blood samples at discharge added	19 July 2011
Details of day 28 telephone call added	19 July 2011
Only existing AEs to be followed up at day 90 and 1 year. No new AEs to be recorded after 4 weeks post IMP last dose	19 July 2011
Clarification of expected AEs and recording requirements for 4 weeks post IMP	19 July 2011
'Exclusions to AE recording requirements' section added	19 July 2011
Reporting of SAEs and SUSARs corrected, post-treatment SAE reporting and follow-up clarified, SUSAR causality assessment clarified, expedited SUSAR reporting clarified	19 July 2011
'Patient loss from study', 'patients lost', 'patients not lost' sections added	19 July 2011
'National registries' section added	19 July 2011
Other reasons for trial termination added	19 July 2011
'Quality of life in cirrhosis study' section clarified	19 July 2011
CRF completion and return to SCTU procedures clarified	19 July 2011
Change from 'minimum of 24 hours' to '< 24 hours for study consideration' for some patients	19 July 2011
Updated WHO performance status definitions added	19 July 2011
Parameters for standard gamble analysis added	19 July 2011

AE, adverse events; CRF, case report form; DNA, deoxyribonucleic acid; EDTA, ethylenediaminetetraacetic acid; IMP, investigational medicinal product; SAE, serious adverse event; SUSARS, suspected unexpected serious adverse event; TENALEA, Trans European Network for clinical trials services; SCTU, Southampton Clinical Trials Unit; WHO, World Health Organization.

Study setting and sample

Patients were identified, screened and recruited in the secondary care setting after admission with acute AH. A total of 66 hospitals across England, Wales and Scotland took part, although only 65 of these recruited patients to the study.

Inclusion criteria

- Aged ≥ 18 years.
- Clinical AH:
    - serum bilirubin level of > 80 µmol/l
    - history of excess alcohol (> 80 g/day for males and > 60 g/day for females) to within 2 months of randomisation.
- Less than 4 weeks since admission to hospital.
- DF of ≥ 32.
- Informed consent.

Exclusion criteria

- Abstinence of > 2 months prior to randomisation.
- Duration of clinically apparent jaundice > 3 months.
- Other causes of liver disease including:
    - evidence of chronic viral hepatitis (hepatitis B or C)
    - biliary obstruction
    - hepatocellular carcinoma.
- Evidence of current malignancy (except non-melanotic skin cancer).
- Previous entry into the study, or use of either prednisolone or PTX within 6 weeks of admission.
- aspartate aminotransferase (AST) level of > 500 IU or alanine aminotransferase (ALT) level of > 300 IU (not compatible with AH).
- Patients with a serum creatinine level of > 500 µmol/l or requiring renal support.
- Patients dependent on inotropic support (adrenaline or noradrenaline). Terlipressin is allowed.
- Active GI bleeding.
- Untreated sepsis.
- Patients with known hypersensitivity to PTX, other methylxanthines or any of the excipients.
- Patients with cerebral haemorrhage, extensive retinal haemorrhage, acute myocardial infarction (within the last 6 weeks) or severe cardiac arrhythmias (not including atrial fibrillation).
- Note that patients with evidence of sepsis, GI bleeding or renal failure may be treated for these conditions for up to 7 days and, if stable, the patients can then be rescreened for eligibility. Treatment can continue for > 7 days if they are stable. Patients who are *not* stable after 7 days of treatment will not be eligible for the study.

Study interventions

Patients were randomised to one of four arms:

- arm A – placebo/placebo
- arm B – placebo/prednisolone
- arm C – PTX/placebo
- arm D – PTX/prednisolone.

The Pharmacy Manufacturing Unit at the Royal Free Hospital in London manufactured the investigational medicinal product (IMP) by overencapsulating the active medication in gelatin capsules. Placebo preparations, which precisely matched the active medication in appearance, contained only microcrystalline cellulose. The IMP was provided in capsule form in two bottles (bottle A and bottle B). Bottle A contained PTX 400 mg or matched placebo and bottle B contained prednisolone 40 mg or matched placebo. Ideally, patients started their IMP within 48 hours of randomisation. It was administered for 28 days (bottle A one capsule daily and bottle B three times daily) and treatment breaks of up to 7 days were acceptable if required.

Concomitant medications may have been administered as medically indicated, including alcohol withdrawal therapy as required. All patients received supportive nutritional therapy. Nutritional supplements were offered but if the patient was unable to take these, enteral nutrition via a nasogastric tube was offered. The aim was to provide 35–40 kcal/kg/day non-protein energy with 1.5 g/kg/day of protein.

Study procedures

Recruitment

Site staff assessed patients admitted into secondary care with an acute episode of AH for potential eligibility for the trial.

Informed consent

Potentially eligible patients (or their legal representatives) were informed about the trial by one of the study team, and then given a patient information sheet to review. Patients were given a minimum of 24 hours in which to consider the study and ask questions, after which they (or their legal representatives) were asked to give written informed consent to participate, on the trial informed consent form.

For relevant patients, consent given by a legal representative was in place until the patient recovered capacity, at which point the patient was informed about the trial and asked to decide whether or not they wanted to continue in the trial. Consent to continue was then sought from the patients themselves.

Randomisation and concealment

Site staff registered patients onto the trial via Trans European Network for clinical trials services (TENALEA), a web-based registration and randomisation system, after written informed consent was obtained from the patient (or their legal representative). Subsequently, screening assessments took place to ascertain eligibility. Non-eligible patients were deemed screening failures, while eligible patients were then randomised by site staff, via the TENALEA, to one of four trial treatment arms. Treatment allocation was blinded to site staff and the patient by providing each patient with a unique four-digit patient pack number. The treatment arm was also concealed to investigators and researchers. Only the study statisticians were unblinded and this was for analysis purposes only.

Randomisation was stratified and performed using the following two stratification factors:

1. geographic region (28 in total)
2. risk group – either ‘high’ or ‘intermediate’ risk (‘high’ risk was defined as either sepsis or history of GI bleeding in the previous 7 days or creatinine level of > 150 µmol/l, or any combination of the these; ‘intermediate’ risk was defined as no sepsis and no history of GI bleeding in the previous 7 days and creatinine ≤ 150 µmol/l).

Unblinding service

The Emergency Scientific and Medical Services at Guy’s and St Thomas’ NHS Foundation Trust provided a 24-hour unblinding service for the trial. The Emergency Scientific and Medical Services held a copy of the study randomisation list and were able to perform emergency unblinding if the enquirer considered the medical management of the patient to be dependent on knowing their treatment allocation.

Data collection and management

Sites entered trial-specific data, as specified in the protocol, onto paper case report forms (CRFs). Completed paper CRFs were sent to the Southampton Clinical Trials Unit (SCTU) that was responsible for the data management of the trial. Data were transcribed from paper CRFs into an InForm database (InForm version 5.0, ORACLE, Redwood Shoves, CA, USA) at the SCTU. A range of data validation checks was carried out within both InForm and SAS version 9.3 (SAS Institute Inc., Cary, NC, USA) to minimise incorrect or missing data.

Source data verification was undertaken during site monitoring visits, in accordance with the SCTU's trial monitoring plan. For sites that recruited ≥ 15 patients a planned monitoring visit was scheduled. In total, 29 planned monitoring visits and two triggered monitoring visits took place.

Baseline

At baseline, a number of characteristics were collected to determine the patient's prognosis at the beginning of the trial. The outcomes collected at this time point included the following: vital signs, World Health Organization performance status, concomitant medication, LFTs, prothrombin time (PT), full blood count (FBC), urea, creatinine, GI bleed or sepsis since screening and time since admission to hospital. Prognostic scores at baseline (DF, MELD score, GAHS and Lille score) were derived using baseline data with the exception of the Lille score, which also uses data from day 7. Demographic data were captured at screening.

Follow-up

Patients who were being treated as in-patients were followed-up every 7 days from the start of treatment until the end of treatment (day 28). If the patient was discharged from hospital during the treatment phase, then the 28-day follow-up visit was conducted by phone, when possible, in order to gather treatment information. All patients were scheduled for a follow-up hospital visit at 90 days and 1 year from start of treatment, at which time details of the patients' alcohol consumption and attendance at alcohol counselling were collected. There were also repeated measurements of outcomes recorded at baseline (see *Baseline*) as well as if patients had experienced an occurrence of GI bleed or sepsis since their last visit. Finally, for the 1-year follow-up visit only, demographic data were collected on the patients' marital status, employment status and housing status.

All patients were asked to consent to information about their health status being held and maintained by the Health and Social Care Information Centre and the NHS Central Register to enable long-term follow-up. In January 2013, the National Institute for Health Research Health Technology Assessment (funding body) awarded a no-cost extension to recruitment (from the end of December 2012 to the end of February 2014), but with follow-up reduced to a minimum of 28 days (to the end of March 2014), so that primary outcome data could be collected for all randomised patients.

Outcome measures

Primary outcome

The primary outcome was 28-day mortality. This was chosen as this time point represents the end of the peak period of mortality for AH. Also, this is consistent with other AH trials, which allows a comparison to be made. Patient mortality at day 28 was treated as a binary outcome (1 = dead, 0 = alive).

Secondary outcomes

Secondary outcomes at day 90 and 1 year

Mortality or liver transplant at day 90 and 1 year were considered as secondary outcome measures. These were calculated in a similar fashion to the primary outcome; the main difference being that these also included the incidence of liver transplantation.

Overall survival (OS) was also evaluated, for which an event was defined as any death occurring during 1-year post-treatment start. The OS time was defined as the following:

- if the date of death was < 1-year post-treatment start, then the OS time was the difference between the treatment start date and death date
- if the date of death was > 1-year follow-up, then patients were censored with an OS time of 365 days
- if no date of death was received, then patients were censored with an OS time of 365 days, if they were known to be alive at 1 year, or the difference between treatment start and date last known alive.

Outcomes of recidivism in patient's alcohol intake were collected at day 90 and 1 year. This included whether or not patients had attended any alcohol counselling sessions. Development of renal failure, sepsis and GI bleed at day 90 and 1 year were also assessed.

Prognostic scores and predictors of mortality

The outcomes of 28-day, 90-day and 1-year mortality were considered in relation to the GAHS at baseline. Other scores considered were DF, MELD score and the Lille score (see *Appendix 1*).

Other baseline measures were evaluated to see whether or not they were predictors of mortality, including: pyrexia (temperature > 37.2 °C), hypotension (systolic blood pressure < 100 mmHg), tachycardia (pulse > 100 beats per minute), alcohol intake (units/week), albumin (g/l), alkaline phosphatase (ALP) (U/l), bilirubin (µmol/l), hepatic encephalopathy (presence vs. absence), PT ratio or international normalised ratio (INR), age (years), white blood cell (WBC) count (10^9 WBC/l), urea (mmol/l) and creatinine (µg/dl).

Safety

Serious adverse events (SAEs) were measured during the treatment period using the National Cancer Institute Common Toxicity Criteria for Adverse Events version 4.0.

Health-related quality of life

Hospital readmission rates for liver- or non-liver related events (including access to other health services) and incremental NHS costs and QoL data [using the European Quality of Life-5 Dimensions (EQ-5D) score], were collected at day 90 and 1-year follow-up visits.

Duration of hospitalisation, type of hospital, time of initial presentation and time of admission to treating hospital (if a tertiary referral) were evaluated.

Sample size

The sample size was performed in nQuery Advisor version 2.0 (Statistical Solutions, Boston, MA, USA) using a two-group continuity-corrected chi-squared test and the following parameters:

- power = 90% (to allow for secondary outcomes)
- two-sided significance level of 5%.

Estimated 28-day mortality rate in each treatment arm is given in *Table 2*.

Based on a reduction in the 28-day mortality rate at the margins from 30% to 21%, a sample size of 513 per group of single agent versus no single agent was required. Thus, in total, the trial required 1026 patients. We allowed for an approximate 10% dropout rate up to day 28 and therefore aimed to randomise 1200 patients to the study, with an equal treatment group allocation.

TABLE 2 Estimated 28-day mortality rates for sample size

	PTX		
	No	Yes	Total
Prednisolone			
No	35% (placebo/placebo)	25% (PTX/placebo)	30%
Yes	25% (placebo/prednisolone)	17%[a] (prednisolone/PTX)	21%
Total	30%	21%	–

a Estimated assuming no interaction, that is multiplicative independent effects.

The sample size for this trial was not powered to assess for any observed treatment interaction and, in fact, assumed that there was no interaction between the two treatments (i.e. that receiving prednisolone in addition to PTX does not change the effect of PTX and vice versa). Assessing the size of any interaction was not of primary interest, as it was assumed to be small or non-existent and would require a fourfold increase in the sample size, so it was deemed appropriate not to power for assessing an interaction.

Stopping guidelines

Prior to this trial, the benefits and harms of PTX in this patient population were unknown, as only limited data were available, and prednisolone had a significant, although uncertain, evidence base. In addition, there was huge uncertainty about whether or not the combination of PTX and prednisolone would be harmful or beneficial, and whether or not there would be a treatment interaction. It was therefore important to conduct interim analyses with predefined criteria to assess both for harm and benefit.

Prespecified stopping guidelines were developed based on the Peto–Haybittle rule, as recommended by Pocock.[30] The guidelines recommended that certain treatment arms were stopped if a two-sided (i.e. for harm or benefit) p-value of < 0.001 was found in group comparisons of day-28 mortality using logistic regression analysis. Pocock[30] states this p-value ties in well with the concept of proof beyond reasonable doubt. This p-value was used for both stopping guidelines of harm and benefit, as the treatments were not new and were already in use in practice. Also, should evidence of harm or benefit arise, evidence needed to be sufficiently convincing to ensure that others would believe it and change their practice accordingly. No adjustment in the significance of the p-value for the final analysis was made, owing to the p-value of 0.001 being considered sufficiently small.

The Data Monitoring and Ethics Committee met to review cumulative safety and efficacy data at three time points during the trial. The following interim stopping guideline assessments were made:

- Primary end-point data from 200 patients: only stopping guidelines for harm were assessed. No treatment arms were stopped for harm at this interim assessment.
- Primary end-point data from 400 patients: stopping guidelines for harm and benefit were assessed. No treatment arms were stopped for harm or benefit at this interim assessment.
- Primary end-point data from 800 patients: stopping guidelines for harm and benefit were assessed. No treatment arms were stopped for harm or benefit at this interim assessment.

Statistical analysis

All trial analyses and reporting were carried out following the Consolidated Standards of Reporting Trials (CONSORT) guidelines. Statistical analysis was carried out in SAS version 9.2 and/or 9.3 following a predefined statistical analysis plan, which was approved by the Data Monitoring and Ethic Committee. Some subsequent analyses were carried out that were not predefined; these will be highlighted in the relevant sections of this report (see *Chapter 3, Secondary Outcomes*).

Analyses were carried out with respect to the factorial design for which:

- the prednisolone-treated group (arms B and D) was compared with the prednisolone control group (arms A and C)
- the PTX-treated group (arms C and D) was compared with the PTX control group (arms B and D).

If a significant interaction (at the 5% significance level) was found then the comparison of interest was individual treatment arm. All descriptive tables were presented by treatment arm.

All analyses were adjusted for risk group and factorial design unless otherwise stated, and were assessed at the two-sided 5% significance level.

Analysis populations

All summaries and analyses were carried out on the intention-to-treat (ITT) population unless otherwise specified. The ITT population included all patients who were randomised regardless of treatment adherence. For analyses at a certain time point, the ITT population was modified to include only patients who were known to be alive at the specified time point or had died prior to the time point, that is, for primary analysis, this included all randomised patients for whom we knew that they were alive at day 28 or who died prior to or on day 28.

A per-protocol population was defined as:

- treatment adherence $\geq 75\%$
    - patients who reached day 28 needed to have had taken at least 75% of treatment (> 21 days) to be included in the per-protocol population
    - patients who withdrew or died before day 28 needed to have had a minimum of 7 days of treatment and taken at least 75% of treatment up until they withdrew/died
- patients who had no major protocol violations.

The primary end-point analysis was repeated for the per-protocol population.

Preliminary analyses

Summary statistics were produced for a selection of key clinical and demographic characteristics at baseline to assess baseline comparability between the four treatment arms. Any significant differences were reported.

Primary analyses

The primary outcome of 28-day mortality was analysed using logistic regression. Although the trial was not powered to detect a treatment interaction, this was tested using a secondary logistic regression model in order to test the underlying model assumptions surrounding factorial design.

Secondary analyses

The primary outcome analysis method was repeated for the 90-day and 1-year mortality or liver transplant outcomes. For OS, comparisons were made using Cox proportional hazards regression and Kaplan–Meier curves.

A number of analyses were carried out on the DF, GAHS, MELD and Lille prognostic scores. First, to assess the predictive power of the scores on 28-day, 90-day and 1-year mortality, a receiver operating characteristic (ROC) curve analysis was carried out. Graphical representation of the ROC curves was produced along with area under the curve statistics. Further ROC curve analysis was carried out on the Lille score comparing the two prednisolone-treated groups (not predefined). Univariate logistic regression analysis was also carried out to look at the relationship between the scores and outcome.

Univariate logistic regression analysis was carried out to assess the relationship between baseline prognostic factors and 28-day, 90-day and 1-year mortality. The prognostic factors included in this analysis were as follows: pyrexia, hypotension, tachycardia, alcohol intake, albumin, ALP, bilirubin, encephalopathy, PT ratio/INR, age, WBCs, urea and creatinine. Odds ratios with a 95% confidence interval (CI) and corresponding *p*-value were produced for each model fitted. Any factors that were significant in the univariate analysis were added to a multivariate model, which also included the treatment indicators for prednisolone and PTX. Backward elimination was then applied to produce adjusted odds ratios for the effects of prednisolone and PTX. Please note that the use of backward elimination was not predefined in the statistical analysis plan.

Descriptive summaries of the following were produced: causes of death, SAEs, alcohol behaviour and alcohol counselling attendance.

Finally, the effect of drinking habits on 1-year mortality was explored by fitting a logistic regression model with abstinence as the reference event. This was compared with the other three alcohol consumption statuses (reduced drinking to below government defined safety limits, reduced drinking but still above safety limits and not reduced). Please note that this analysis was not predefined.

For all regression models applied, the assumptions underlying these models were checked using appropriate methods; the Hosmer–Lemeshow goodness-of-fit test for assessing logistic regression modelling and the proportional hazards assumption for Cox proportional hazards modelling. Covariates used in logistic regression modelling were assessed for outliers. Subsequently, outliers were excluded during a sensitivity analysis to allow for any differences in results to be examined.

Sensitivity analyses

A sensitivity analysis of the primary end-point outcome was produced by excluding any patients who were randomised and followed-up but were later found to be ineligible through central monitoring of baseline characteristics. Patients were excluded if they had any of the following characteristics at baseline:

- < 18 years old
- bilirubin level of < 80 µmol/l
- maximum alcohol consumption in past 2 months < 70 units/week for a male
- maximum alcohol consumption in past 2 months < 52.5 units/week for a female
- time between initial admission and screening date > 4 weeks
- DF of < 32
- AST level of > 500 IU/l
- ALT level of > 300 U/l
- creatinine level of > 500 µmol/l
- unresolved sepsis
- unresolved GI bleed.

Please note that this sensitivity analysis was not predefined.

Subgroup analyses

Subgroup analyses were performed to further explore whether or not particular subsets of the patient population were at more benefit to treatment and prolongation of survival. The subgroups considered were the prognostic scores (DF, GAHS, MELD and Lille score) and risk at baseline. These analyses looked at 28-day, 90-day and 1-year mortality using logistic regression analysis methods as described in *Statistical analysis*. Please note that these subgroup analyses were not predefined.

Patient and public involvement

The Trial Management Group (TMG) was delighted to recruit Mr Colin Stanfield as a patient representative. He attended the majority of TMG meetings during the set up and running of the STOPAH trial. His input from a patient perspective was valuable in terms of the design of the patient information sheet and consent form. He also provided guidance on a number of occasions about patient behaviour and acceptability of interventions, and data collection.

Chapter 4 Trial results

Recruitment and randomisation

The SCTU site staff performed initiation visits between October 2010 and August 2012. In total, 66 trial sites were opened but subsequently one site was closed after failing to recruit any patients. The details of recruitment by site are given in *Figure 1*.

A total of 5234 patients were screened for the trial, and after applying inclusion and exclusion criteria, 1103 were randomised and evenly allocated to the four treatment arms. Analysis of screening logs indicates that the trial recruited 85% of eligible patients; disease severity being too mild was the most common reason for patients not being recruited. There were a total of 50 dropouts prior to the primary end-point time point at 28 days. A total of 168 patients died between 28 days and the next time point at 90 days, and 221 patients dropped out of the study between 90 days and 1 year. We were unable to follow-up 228 patients recruited during the last 9 months of the trial as the study was closed. Details of the patient flow are provided in the CONSORT diagram (*Figure 2*).

As detailed in the sample size and power calculation in *Chapter 3*, *Sample size*, the trial aimed to recruit 1200 patients in order to be confident of having primary end-point data on 1026 patients. Although the final number of patients recruited fell short of the target number, the trial exceeded the number needed for the trial power to be maintained in the primary end-point analysis.

Flow of participants through the trial

See *Figure 2* for CONSORT flow diagram.

Unblinding of patients

Requests for unblinding were received for a small number of patients. In all cases there were reasonable clinical justifications for the unblinding (detailed in *Table 3*), which were judged to be acceptable by the trial chief investigator (ChI). In view of the hard trial end points and the small number of unblinding events these are not thought to have any impact on the trial result.

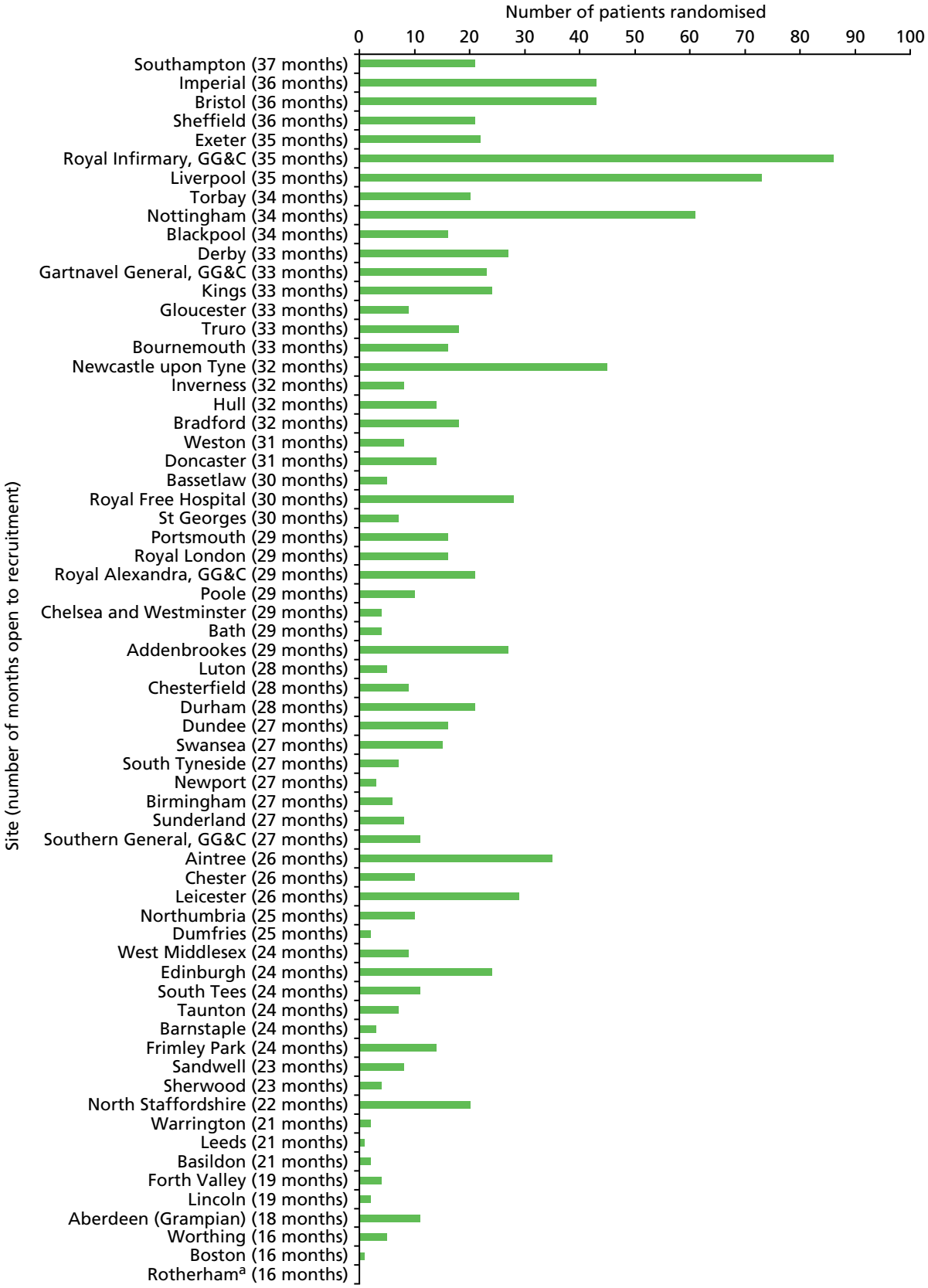

FIGURE 1 Number of patients randomised by site, ordered by number of months open to recruitment. a, All sites closed to recruitment on 28 February 2014 with the exception of Rotherham, which closed on 19 October 2012, after 16 months open, owing to poor recruitment activity. GG&C, Greater Glasgow and Clyde.

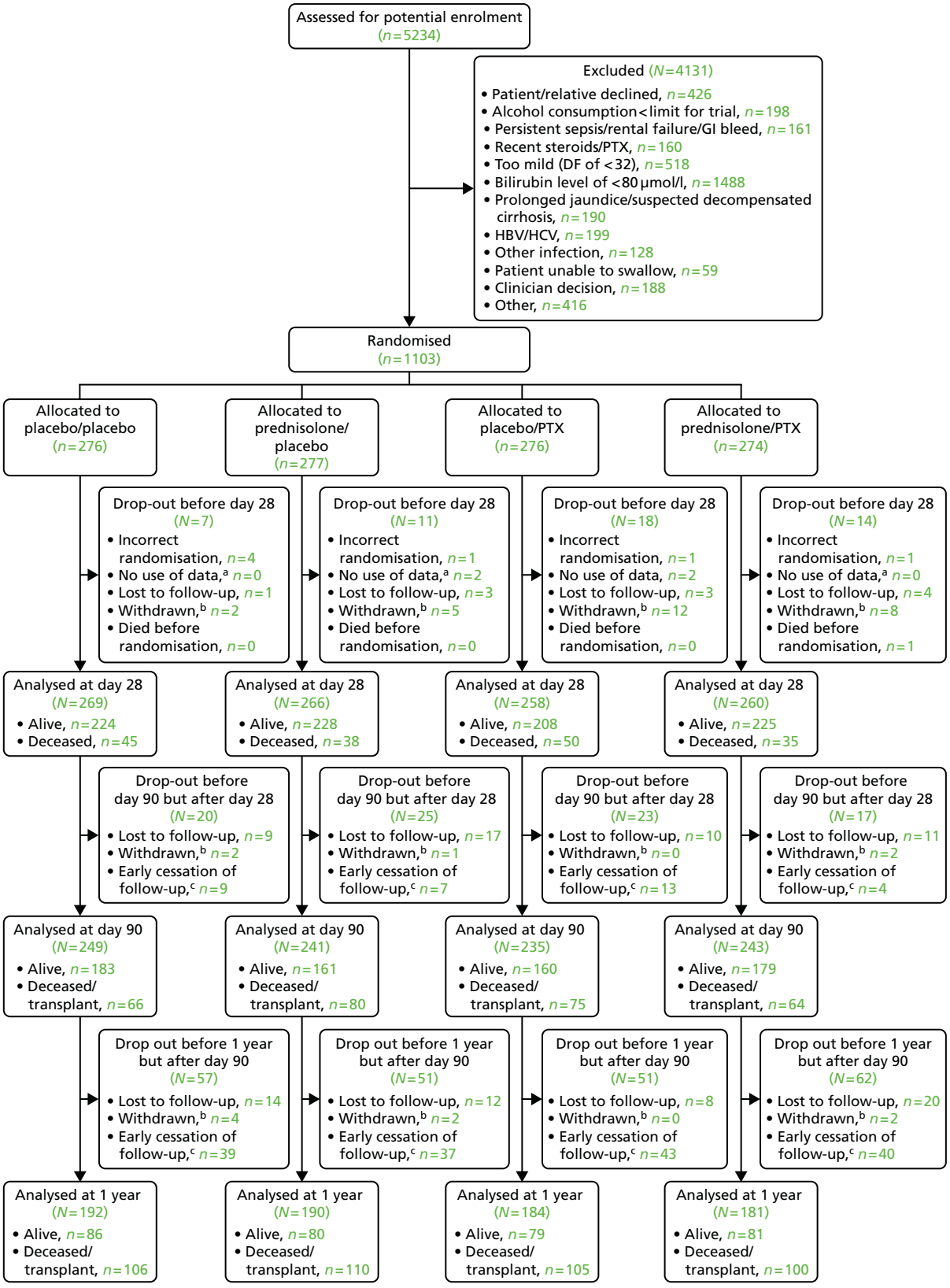

FIGURE 2 Consolidated Standards of Reporting Trials flow chart. HBV, hepatitis B virus; HCV, hepatitis C virus. a, No usage of data allowed – patient withdrew consent and would not allow any data collection to be used; b, withdrew – patient withdrew consent but allowed use of data collection up to point of withdrawal; and c, early cessation of follow-up – recruitment extended to end of February 2014, but follow-up for all patients ceased end March 2013, so that primary end-point data could be collected for all patients.

TABLE 3 Summary of patients unblinded during the STOPAH trial

Treatment allocation	Reason for unblinding
Placebo/prednisolone	*Research nurse requested unblinding of patient who had had surgery following bowel perforation previous night. Treating medical/surgical team wanted to know if patient had received steroids*
PTX/placebo	*Patient's clinical condition had deteriorated. Responsible clinician has discussed this with ChI who has authorised unblinding*
PTX/placebo	*Due to forthcoming inquest into patient's death, ChI requested unblinding via Hull RI research nurse*
Placebo/placebo	*Patient has been in ICU for 5–6 days following complications during an investigative procedure. Admitted to ICU 6 days after starting trial drug(s) and these ceased on admission. Patients condition 'stable but critical'. Patient put on hydrocortisone when admitted to ICU and needs to stop [prednisolone]. Caller insists decision to stop hydrocortisone dependent on knowledge of study treatment so they can decide to stop immediately or taper drug off*
PTX/prednisolone	*Patient admitted to West Middlesex Hospital with sepsis and pneumonia. Condition developing and worsening over last 2 days*
Placebo/prednisolone	*Patient is very poorly on ITU and the his/her medical management is dependent on knowing if he/she is receiving active or placebo drug*
PTX/placebo	*Unblinding request from ChI (no further details available)*
Placebo/prednisolone	*The patient died and the treatment allocation information would be useful for the post-mortem*
PTX/placebo	*ChI agreed to unblinding as patient developed severe psychosis*
Placebo/prednisolone	*Patient had persistent jaundice since November – caller would like to know what treatment the patient was receiving when taking trial medication. The patient finished treatment on 20/12/2012*
PTX/placebo	*Patient due to stop trial tomorrow (15/06/2013) but has been deteriorating. Consultant treating patient would like to start steroids, but only if patient hasn't already been on steroids in the study*
Placebo/prednisolone	*First call at 23.50 – patient on GI ward with suspected perforation. Dr wants to administer IV hydrocortisone and would like to know treatment that the patient is on. Second call at 00.12 – confirming that the Dr would like to unblind as the patient has an actual perforation and sepsis. No further details on patient's condition as caller is the on-call pharmacist*
Placebo/prednisolone	*Caller has been asked by the PI to find out what arm the patient was on. They finished the trial over 3 weeks ago and have been hospitalised with deranged LFTs. The caller has spoken to the ChI and they have confirmed that this subject can be unblinded*
Placebo/placebo	*Caller has a patient with severe VT and torsade and requires tertiary cardiac HDU care. The patient was enrolled on to the trial on 31/07/2013 and completed a month's supply of IMP. The patient was transported from their local hospital for tertiary cardiac HDU care from the cardiologist. Local PI has received confirmation to unblind from ChI*
PTX/prednisolone	*Caller requested unblinding of a patient that has died, stating that they needed to know the treatment allocation for the coroner's report. They think the patient died of necrotising pancreatitis and peritonitis*

ChI, chief investigator; ICU, intensive care unit; ITU, intensive therapy unit; IV, intravenous; HDU, high-dependency unit; RI, Royal Infirmary; PI, principal investigator; VT, ventricular tachycardia.

Characteristics of the study sample

The baseline characteristics of the patients are typical of those seen in other studies in this population of patients. In particular, disease severity scores (DF, MELD, GAHS and Lille) were consistent with those seen in other publications.[23,31] Details of the clinical and laboratory values are provided in *Table 4*. There were no significant differences in laboratory values or prognostic scores between the four treatment arms except for ALP and Lille score, which were significant at the 1% level, and PT, which was significant at the 5% level.

TABLE 4 Baseline characteristics

Baseline characteristics	Placebo/placebo (*N* = 272)	Prednisolone/placebo (*N* = 274)	Placebo/PTX (*N* = 273)	Prednisolone/PTX (*N* = 273)
n *(%)*				
Male	162 (60)	177 (65)	164 (60)	182 (67)
White	259 (95)	262 (96)	264 (97)	261 (96)
WHO performance status				
0 – asymptomatic	14 (5)	24 (9)	21 (8)	26 (10)
1 – symptomatic but completely ambulatory	75 (28)	84 (31)	90 (33)	77 (28)
2 – symptomatic < 50% in bed	91 (33)	77 (28)	69 (25)	77 (28)
3 – symptomatic > 50% in bed	66 (24)	69 (25)	74 (27)	67 (25)
4 – bedbound	22 (8)	17 (6)	15 (5)	17 (6)
Encephalopathy				
None	191 (70)	205 (75)	211 (77)	190 (70)
Grade I	46 (17)	38 (14)	33 (12)	48 (18)
Grade II	19 (7)	19 (7)	16 (6)	12 (4)
Grade III	5 (2)	1 (< 0.5)	5 (2)	7 (3)
Grade IV	0	0	0	0
Sepsis on admission	31 (11)	27 (10)	23 (8)	29 (11)
Renal failure on admission	1 (< 0.5)	0	1 (< 0.5)	0
GI bleeding on admission	16 (6)	21 (8)	15 (5)	15 (5)
Pyrexia	42 (15)	28 (10)	32 (12)	34 (12)
Hypotension	49 (18)	46 (17)	45 (16)	52 (19)
Tachycardia	47 (17)	46 (17)	56 (21)	52 (19)

continued

TABLE 4 Baseline characteristics (*continued*)

Baseline characteristics	Placebo/placebo (*N* = 272)	Prednisolone/placebo (*N* = 274)	Placebo/PTX (*N* = 273)	Prednisolone/PTX (*N* = 273)
Mean (SD)				
Age (years)	48.8 (10.28)	49.3 (10.57)	47.9 (10.23)	48.6 (9.81)
Alcohol consumption (g/day)				
Female	153.7 (98.54)	141.7 (75.42)	145.7 (93.11)	157.0 (143.77)
Male	195.4 (126.53)	209.9 (117.67)	192.4 (129.81)	201.7 (127.28)
Time from admission to treatment (days)	6.1 (3.82)	6.5 (3.85)	6.7 (4.16)	6.5 (4.40)
Laboratory values				
Bilirubin (μmol/l)	305.9 (157.94)	297.5 (155.19)	292.6 (144.62)	306.1 (163.09)
Albumin (g/l)	25.6 (6.26)	25.2 (6.22)	25.1 (5.37)	25.3 (6.01)
AST (IU/l)	143.7 (69.53)	133.6 (64.78)	134.3 (73.24)	143.4 (77.19)
ALP (U/l)[a]	184.7 (86.36)	207.7 (113.06)	182.4 (85.09)	196.1 (98.52)
Creatinine (μmol/l)	73.4 (38.33)	79.8 (46.31)	78.5 (49.08)	81.5 (51.65)
WBC ($\times 10^9$/l)	10.1 (5.56)	10.6 (8.13)	9.9 (5.42)	9.8 (4.94)
Neutrophils ($\times 10^9$/l)	7.6 (5.20)	7.7 (5.24)	7.4 (4.92)	7.3 (4.51)
PT (seconds)[a]	21.1 (5.27)	20.8 (5.26)	22.1 (6.79)	21.1 (5.17)
Prognostic scores				
DF	61.9 (25.66)	60.7 (25.34)	65.6 (31.64)	62.4 (25.62)
GAHS	8.2 (1.18)	8.3 (1.27)	8.3 (1.19)	8.3 (1.32)
MELD	20.7 (5.54)	21.2 (6.19)	21.4 (6.32)	21.5 (6.84)
Lille score[a,b]	0.5 (0.32)	0.4 (0.33)	0.5 (0.34)	0.4 (0.34)

SD, standard deviation; WHO, World Health Organization.
a There were no significant differences between treatment arms for the baseline characteristics except for ALP and Lille score, which were significant at the 1% level, and PT, which was significant at the 5% level (looking at non-missing data only).
b The Lille score uses data at baseline and day 7.

Allocation of treatments

Allocation to treatment arms is detailed in *Table 5*. Overall, 22% of patients were categorised in the high-risk classification at randomisation, although a small number of these were inappropriately classified and were correctly defined for the primary end-point analysis.

TABLE 5 Allocation of treatment arms by stratification factors

Stratification factors	Placebo/placebo (*N* = 272)	Prednisolone/placebo (*N* = 274)	Placebo/PTX (*N* = 273)	Prednisolone/PTX (*N* = 273)
Risk used in randomisation,*[a] n *(%)				
Intermediate	212 (78)	214 (78)	216 (79)	214 (78)
High[b]	60 (22)	60 (22)	57 (21)	59 (22)
Actual risk at baseline,*[c] n *(%)				
Intermediate	216 (79)	219 (80)	222 (81)	213 (78)
High[b]	56 (21)	55 (20)	51 (19)	60 (22)
IMP centre,* n *(%)				
Aberdeen	3 (1)	3 (1)	3 (1)	2 (1)
Addenbrookes	6 (2)	6 (2)	7 (3)	6 (2)
Basildon	0	0	1 (< 0.5)	0
Blackpool	4 (1)	4 (1)	4 (1)	4 (1)
Bradford	4 (1)	5 (2)	5 (2)	4 (1)
Bristol	18 (7)	18 (7)	18 (7)	18 (7)
Derby	17 (6)	18 (7)	17 (6)	18 (7)
Dumfries	1 (< 0.5)	0	1 (< 0.5)	0
Dundee	3 (1)	5 (2)	4 (1)	4 (1)
Edinburgh	6 (2)	6 (2)	6 (2)	6 (2)
Forth Valley	2 (1)	1 (< 0.5)	1 (< 0.5)	0
Frimley Park	4 (1)	4 (1)	3 (1)	3 (1)
Hull	4 (1)	3 (1)	3 (1)	4 (1)
Imperial	23 (8)	23 (8)	22 (8)	23 (8)
Inverness	2 (1)	1 (< 0.5)	3 (1)	2 (1)
King's	9 (3)	8 (3)	8 (3)	8 (3)
Leeds	0	1 (< 0.5)	0	0
Liverpool	30 (11)	29 (11)	30 (11)	29 (11)
Luton	1 (< 0.5)	1 (< 0.5)	1 (< 0.5)	2 (1)
Newcastle upon Tyne	26 (10)	25 (9)	24 (9)	26 (10)
Nottingham	23 (8)	23 (8)	25 (9)	25 (9)
Plymouth	20 (7)	21 (8)	21 (8)	21 (8)
Glasgow Royal Infirmary	34 (13)	36 (13)	34 (12)	35 (13)
Sheffield	9 (3)	10 (4)	11 (4)	10 (4)
Southampton	16 (6)	16 (6)	16 (6)	15 (5)
St Georges	2 (1)	2 (1)	1 (< 0.5)	2 (1)
Swansea	4 (1)	4 (1)	3 (1)	4 (1)
Worthing	1 (< 0.5)	1 (< 0.5)	1 (< 0.5)	2 (1)

a Risk classification as used in randomisation system.
b High risk is defined as either sepsis or history of GI bleeding in the previous 7 days or creatinine level of > 150 µmol/l, or any combination of the these; intermediate risk is defined as no sepsis and no history of GI bleeding in the previous 7 days and creatinine level of ≤ 150 µmol/l.
c Risk classification at baseline.

Treatment adherence

The predefined per-protocol analysis depended on measurement of treatment adherence as defined in *Chapter 3, Analysis population*. Per-protocol analysis required patients to have taken at least 75% of prescribed medication and that they had no protocol violations. Data on treatment adherence were poorly documented, so it was difficult to define a sufficiently sized 'per-protocol' population for analysis.

Primary outcome

Table 6 documents the primary outcome at 28 days. At day 28, 45 of 269 (16.7%) patients in the placebo/placebo arm had died, 50 of 258 (19.4%) patients in the PTX/placebo arm had died, 38 of 266 (14.3%) patients in the prednisolone/placebo arm had died and 35 of 260 (13.5%) patients in the PTX/prednisolone arm had died. Therefore, in the prednisolone-treated group, 13.9% of patients died compared with 18.0% in the group who did not receive prednisolone (odds ratio 0.72, 95% CI 0.52 to 1.01; $p = 0.056$). In the PTX-treated group, 16.5% of patients died compared with 15.5% in the group who did not receive PTX (odds ratio 1.07, 95% CI 0.77 to 1.49; $p = 0.686$). There was no interaction between the effects of prednisolone and PTX ($p = 0.407$) (*Table 7*).

TABLE 6 Observed 28-day mortality[a]

	PTX		
	No	Yes	Total
Prednisolone			
No	16.7% (45/269) (placebo/placebo)	19.4% (50/258) (placebo/PTX)	18.0% (95/527)
Yes	14.3% (38/266) (prednisolone/placebo)	13.5% (35/260) (prednisolone/PTX)	13.9% (73/526)
Total	15.5% (83/535)	16.5% (85/518)	16.0% (168/1053)

a Includes all randomised patients who have at least 28 days of data or died prior to or on day 28.

TABLE 7 Primary analysis results: logistic regression analysis for 28-day mortality[a]

	Prednisolone	No prednisolone	PTX	No PTX
28-day mortality (*n*/*N*)	13.9% (73/526)	18.0% (95/527)	16.4% (85/518)	15.5% (83/535)
Adjusted odds ratio[b] (95% CI)	0.72 (0.52 to 1.01)		1.07 (0.77 to 1.49)	
p-value	0.056		0.686	

a Includes all randomised patients who have at least 28 days of data or died prior to or on day 28.
b Adjusted for risk (high or intermediate) at baseline and factorial design (all comparisons are for the intervention compared with its respective control).

Notes
Odds ratios < 1 represent a favourable outcome for the corresponding intervention.
The interaction between the interventions was investigated as a secondary analysis and was found to be non-significant [interaction coefficient = –0.284, 95% CI –0.956 to 0.387; *p*-value = 0.407)].

Secondary outcomes

Day 90, year 1 and overall survival

Table 8 documents the secondary outcomes at day 90. At day 90, there had been 139 deaths out of 478 participants (29.1%) deaths in the PTX group compared with 146 deaths out of 490 participants (29.8%) deaths in the group not treated with PTX (odds ratio 0.97, 95% CI 0.73 to 1.28; $p = 0.807$). There had been 144 of 484 (29.8%) deaths in the prednisolone group compared with 141 of 484 (29.1%) deaths in the group not treated with prednisolone (odds ratio 1.02, 95% CI 0.77 to 1.35; $p = 0.875$) (*Table 9*).

At 1 year there had been 205 of 365 (56.2%) deaths in the PTX group, compared with 216 of 382 (56.5%) deaths in the group not treated with PTX (odds ratio 0.99, 95% CI 0.74 to 1.33; $p = 0.972$). There had been 210 of 371 (29.8%) deaths in the prednisolone group, compared with 211 of 376 (56.1%) deaths in the group not treated with prednisolone (odds ratio 1.01, 95% CI 0.76 to 1.35; $p = 0.937$) (*Tables 10* and *11*).

Neither prednisolone nor PTX appear to influence all-cause survival after 28 days (*Figures 3–5*). Analysis of the causes of mortality as recorded by the investigators indicate that most patients died from liver-related causes, and in many cases this included a contribution from sepsis. The 28-day mortality in this trial was appreciably less than that seen in other studies in this patient group. However, the overall mortality at 1 year is 56%, illustrating the extent to which ALD contributes to the alarming rise in liver disease mortality in the UK.

TABLE 8 Observed 90-day mortality or liver transplant[a]

	PTX		
	No	Yes	Total
Prednisolone			
No	26.5% (66/249) (placebo/placebo)	31.9% (75/235) (placebo/PTX)	29.1% (141/484)
Yes	33.2% (80/241) (prednisolone/placebo)	26.3% (64/243) (prednisolone/PTX)	29.8% (144/484)
Total	29.8% (146/490)	29.1% (139/478)	29.4% (285/968)

a Includes all randomised patients who have at least 28 days of data or died prior to or on day 28.

TABLE 9 Logistic regression analysis for 90-day mortality[a]

	Prednisolone	No prednisolone	PTX	No PTX
90-day mortality/liver transplant	29.8% (144/484)	29.1% (141/484)	29.1% (139/478)	29.8% (146/490)
Odds ratio (95% CI)[b]	1.02 (0.77 to 1.35)		0.97 (0.73 to 1.28)	
p-value	0.875		0.807	

a Includes all randomised patients who have at least 90 days of data or died prior to or on day 90.
b Adjusted for true risk at baseline and factorial design.
Notes
Odds ratios < 1 represent a favourable outcome for the corresponding intervention.
The interaction between the interventions was investigated as a secondary analysis and was found to be significant at the 5% significance level (interaction coefficient –0.622, 95% CI –1.181 to –0.064; $p = 0.029$).
Logistic regression model: 90-day mortality = intercept + prednisolone indicator + PTX indicator + risk.
No patients had a liver transplant by day 90.

TABLE 10 Observed 1-year mortality or liver transplant[a]

	PTX		
	No	Yes	Total
Prednisolone			
No	55.2% (106/192) (placebo/placebo)	57.1% (105/184) (placebo/PTX)	56.1% (211/376)
Yes	57.9% (110/190) (prednisolone/placebo)	55.2% (100/181) (prednisolone/PTX)	56.6% (210/371)
Total	56.5% (216/382)	56.2% (205/365)	56.4% (421/747)

a Includes all randomised patients who have 1 year of data, died prior to or at 1 year or had a liver transplant prior to or at 1 year.

Notes

Three patients had a liver transplant by 1 year: one in the prednisolone/PTX arm at day 216 and two in placebo/PTX arm at days 270 and 360 (patients situated at the following sites: Southampton, Nottingham and Glasgow Southern General).

TABLE 11 Logistic regression analysis for 1-year mortality or liver transplant[a]

	Prednisolone	No prednisolone	PTX	No PTX
1-year mortality/liver transplant, % (*n*/*N*)	56.6% (210/371)	56.1% (211/376)	56.2% (205/365)	56.5% (216/382)
Adjusted odds ratio[b] (95% CI)	1.01 (0.76 to 1.35)		0.99 (0.74 to 1.33)	
p-value	0.937		0.972	

a Includes all randomised patients who have 1 year of data, died prior to or at 1 year or had a liver transplant prior to or at 1 year.

b Adjusted for risk (high or intermediate) at baseline and factorial design (all comparisons are for the intervention compared with its respective control).

Notes

Odds ratios < 1 represent a favourable outcome for the corresponding intervention.

The interaction between the interventions was investigated as a secondary analysis and was found to be non-significant at the 5% significance level (interaction coefficient = –0.205, 95% CI –0.787 to 0.377; $p = 0.490$).

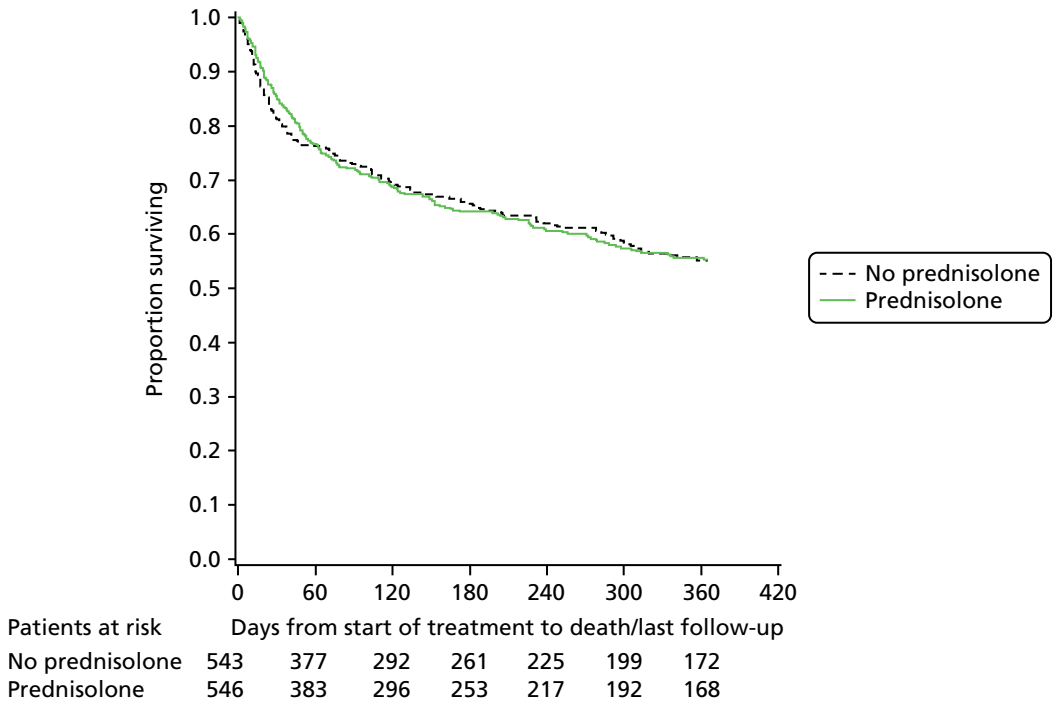

FIGURE 3 Prednisolone vs. no prednisolone.

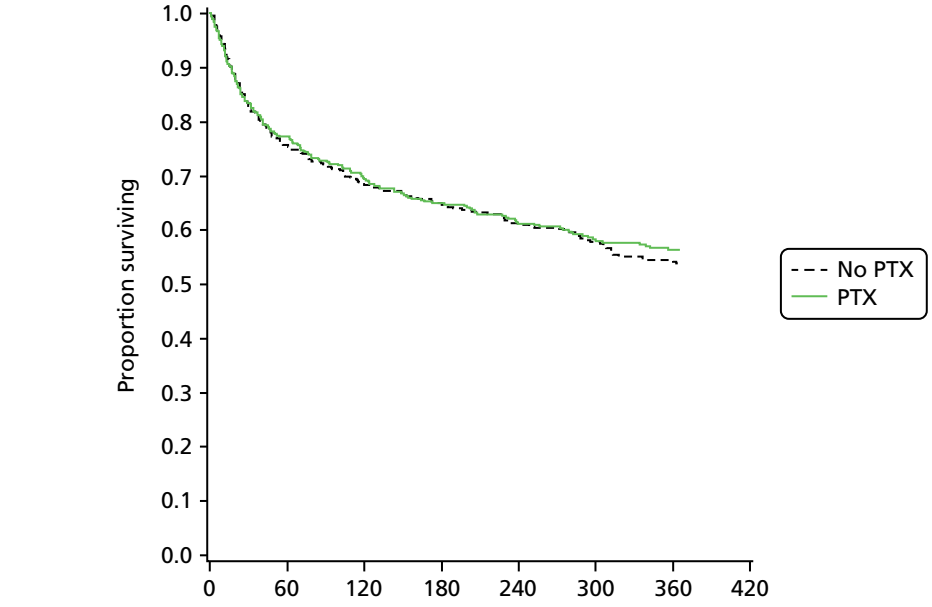

FIGURE 4 Pentoxifylline vs. no PTX.

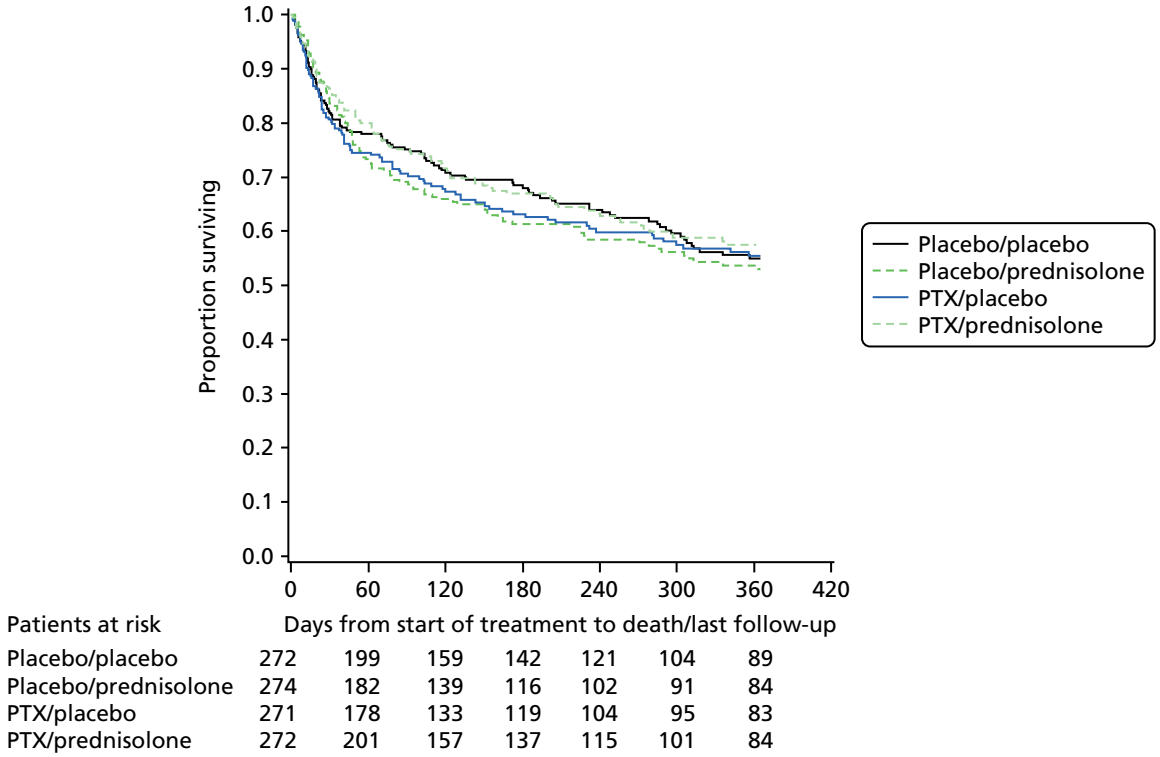

FIGURE 5 Individual treatment arms.

Prognostic scores

Four scoring systems have been described for use in providing prognostic information in patients with AH. DF, GAHS and MELD are derived from clinical and laboratory parameters at baseline whereas the Lille score also incorporated the response to 1 week of prednisolone therapy, determined by the change in bilirubin level at 7 days. Although each of the scores were highly significantly associated with mortality (*Table 12*) at each time point, the diagnostic performance was not considered to be adequate at a threshold for the area under the ROC curve of 0.75 (*Table 13* and *Figure 6*). Inevitably the prognostic value of the scores diminished with increasing duration of follow-up.

TABLE 12 Univariate logistic regression analysis for prognostic scores

Prognostic score	28-day mortality[a]				90-day mortality[b]				1-year mortality[c]			
	n	Odds ratio	95% CI	*p*-value	*n*	Odds ratio	95% CI	*p*-value	*n*	Odds ratio	95% CI	*p*-value
DF	1049	1.021	1.016 to 1.027	< 0.001	965	1.022	1.017 to 1.028	< 0.001	744	1.022	1.015 to 1.028	< 0.001
GAHS	930	2.128	1.789 to 2.532	< 0.001	855	1.956	1.698 to 2.253	< 0.001	650	1.406	1.239 to 1.595	< 0.001
MELD	1008	1.147	1.116 to 1.179	< 0.001	925	1.138	1.109 to 1.168	< 0.001	715	1.105	1.076 to 1.136	< 0.001
Lille[d]	697	1.027	1.019 to 1.034	< 0.001	637	1.024	1.019 to 1.030	< 0.001	498	1.017	1.012 to 1.023	< 0.001

a Includes all randomised patients who have at least 28 days of data or died prior to or on day 28 and have data for the relevant prognostic score.
b Includes all randomised patients who have at least 90 days of data or died prior to or on day 90 and have data for the relevant prognostic score.
c Includes all randomised patients who have 1 year of data or died prior to or on 1 year and have data for the relevant prognostic score.
d Lille fitted as a continuous variable using a simple transformation.

TABLE 13 Receiver operating characteristics curve analysis

Prognostic score	28-day mortality[a]			90-day mortality[b]			1-year mortality[c]		
	n	AUC	95% CI	*n*	AUC	95% CI	*n*	AUC	95% CI
DF	1049	0.688	0.646 to 0.730	965	0.683	0.647 to 0.720	744	0.653	0.614 to 0.693
GAHS	930	0.733	0.689 to 0.778	855	0.711	0.673 to 0.748	650	0.619	0.577 to 0.660
MELD	1008	0.741	0.699 to 0.783	925	0.712	0.676 to 0.749	715	0.664	0.624 to 0.704
Lille	697	0.735	0.682 to 0.789	637	0.718	0.675 to 0.761	498	0.663	0.616 to 0.711

AUC, area under the curve.
a Includes all randomised patients who have at least 28 days of data or died prior to or on day 28 and have data for the relevant prognostic score.
b Includes all randomised patients who have at least 90 days of data or died prior to or on day 90 and have data for the relevant prognostic score.
c Includes all randomised patients who have 1 year of data or died prior to or on 1 year and have data for the relevant prognostic score.

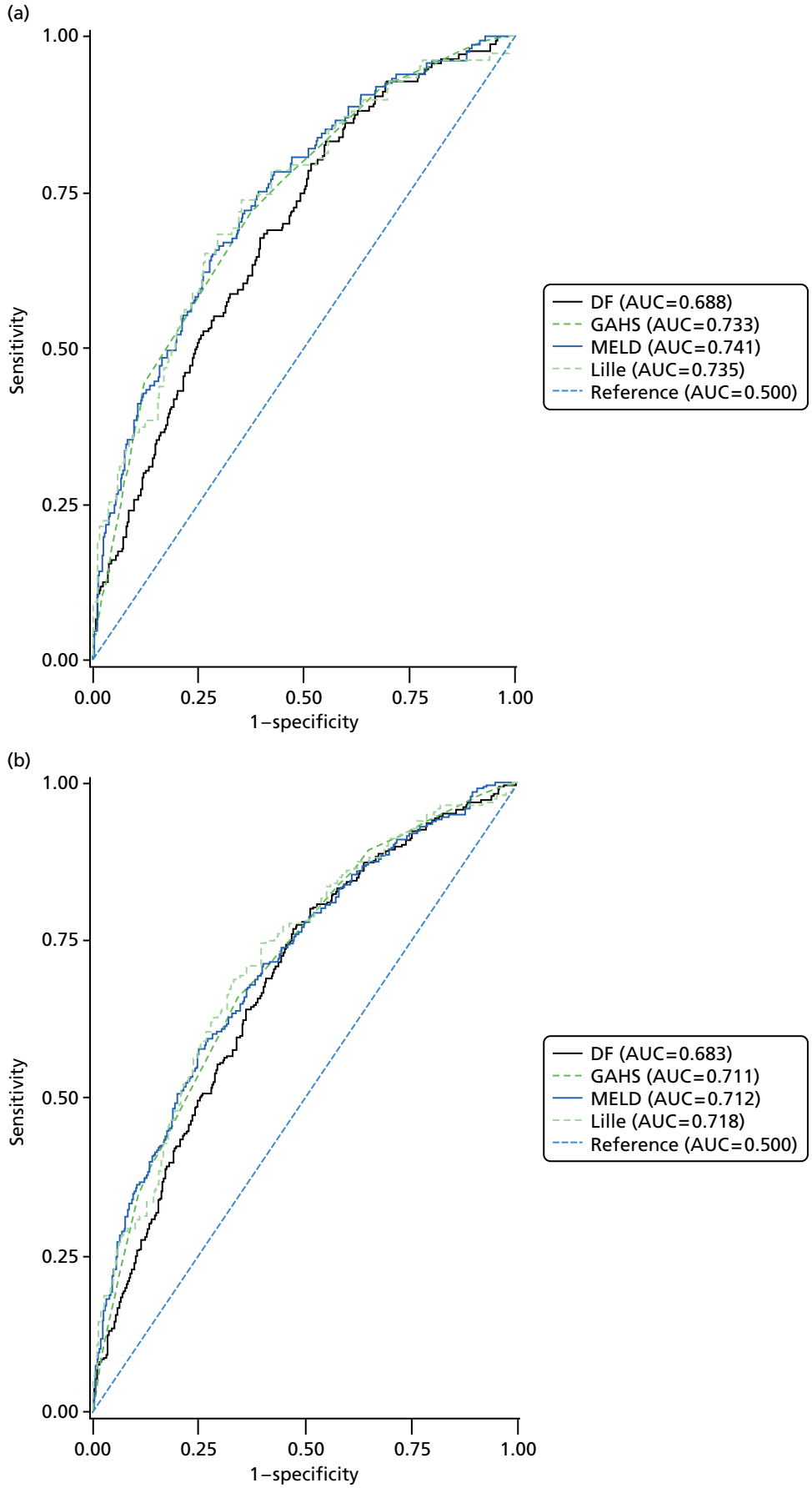

FIGURE 6 Receiver operating characteristics curves for prognostic scores. (a) 28-day mortality; (b) 90-day mortality; and (c) 1-year mortality. AUC, area under the curve. (*continued*)

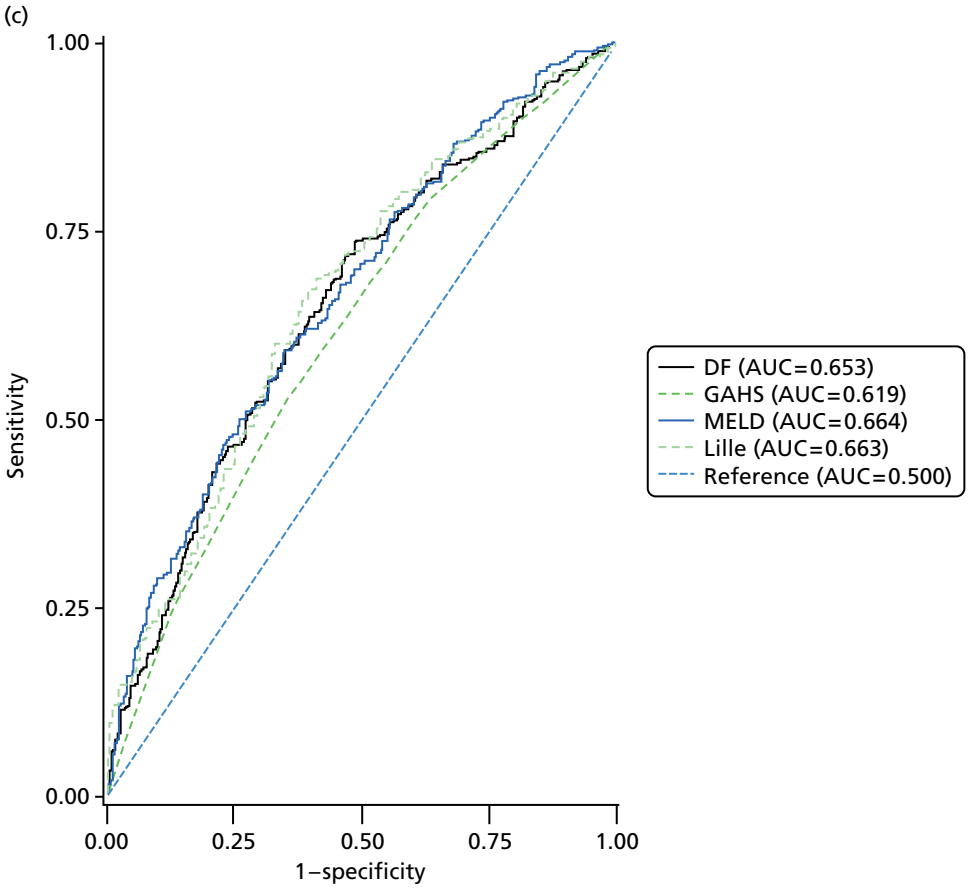

FIGURE 6 Receiver operating characteristics curves for prognostic scores. (a) 28-day mortality; (b) 90-day mortality; and (c) 1-year mortality. AUC, area under the curve.

The impact of prognostic scores on the response to prednisolone was explored by stratifying patients by the magnitude of the prognostic score value or by their risk category (*Table 14*). These data suggest that the risk group (high vs. intermediate) has no impact on the effect of prednisolone on mortality. Based on the odds ratios, patients with a low DF and low GAHS appear to derive more benefit from prednisolone. However, in contrast, patients may benefit from prednisolone when they have a higher MELD score. These results appear contradictory and will need to be explored in further studies.

The Lille score was originally derived to predict mortality based on the response to prednisolone. When the analysis of Lille performance was restricted to only those patients who received prednisolone, the diagnostic performance improved (*Table 15* and *Figure 7*). However, in the original description of the Lille score, a score ≥ 0.45 identified a population of patients who had an expected mortality of 75% by 6 months, whereas in our study a Lille score ≥ 0.45 was associated with around 50% survival at 6 months

TABLE 14 Prednisolone vs. no prednisolone logistic regression model results for 28-day, 90-day and 1-year mortality by factorial design and risk group (ITT population at day 28, day 90 and 1 year)

	Day 28			Day 90			1 year		
Risk	*n*	Odds ratio (95% CI)	*p*-value	*n*	Odds ratio (95% CI)	*p*-value	*n*	Odds ratio (95% CI)	*p*-value
High	218	0.72 (0.39 to 1.33)	0.291	203	0.86 (0.49 to 1.52)	0.603	151	0.82 (0.42 to 1.61)	0.561
Intermediate	835	0.72 (0.48 to 1.07)	0.105	765	1.08 (0.79 to 1.49)	0.631	596	1.07 (0.78 to 1.48)	0.665

TABLE 15 Receiver operating characteristics analysis for Lille score (prednisolone-treated group only)

Lille score by treatment group	28-day mortality[a]			90-day mortality[b]			1-year mortality[c]		
	n	AUC	95% CI	*n*	AUC	95% CI	*n*	AUC	95% CI
Prednisolone/placebo	168	0.765	0.661 to 0.870	152	0.726	0.640 to 0.812	116	0.650	0.549 to 0.751
Prednisolone/PTX	158	0.735	0.596 to 0.873	145	0.740	0.643 to 0.837	107	0.656	0.552 to 0.761

AUC, area under the curve.

a Includes all randomised patients in the prednisolone/placebo or prednisolone/PTX groups who have at least 28 days of data or died prior to or on day 28 and have data for the Lille score.

b Includes all randomised patients in the prednisolone/placebo or prednisolone/PTX groups who have at least 90 days of data or died prior to or on day 90 and have data for the Lille score.

c Includes all randomised patients in the prednisolone/placebo or prednisolone/PTX groups who have at least 1 year of data or died prior to or on 1 year and have data for the Lille score.

Notes

Four patients (prednisolone/placebo = 2, prednisolone/PTX = 2) have negative values of DF derived from PT and bilirubin at baseline. The PT values were queried but were not resolved prior to database lock; therefore, they are excluded from the table as all of the prognostic factors use PT in their derivation.

Interpretation: perfect tests yield an AUC of 1.0. As a rule of thumb, a test with an AUC of > 0.9 has high accuracy, whereas a score of 0.7–0.9 indicates moderate accuracy, a score of 0.5–0.7 indicates low accuracy and a score of 0.5 indicates a chance result.

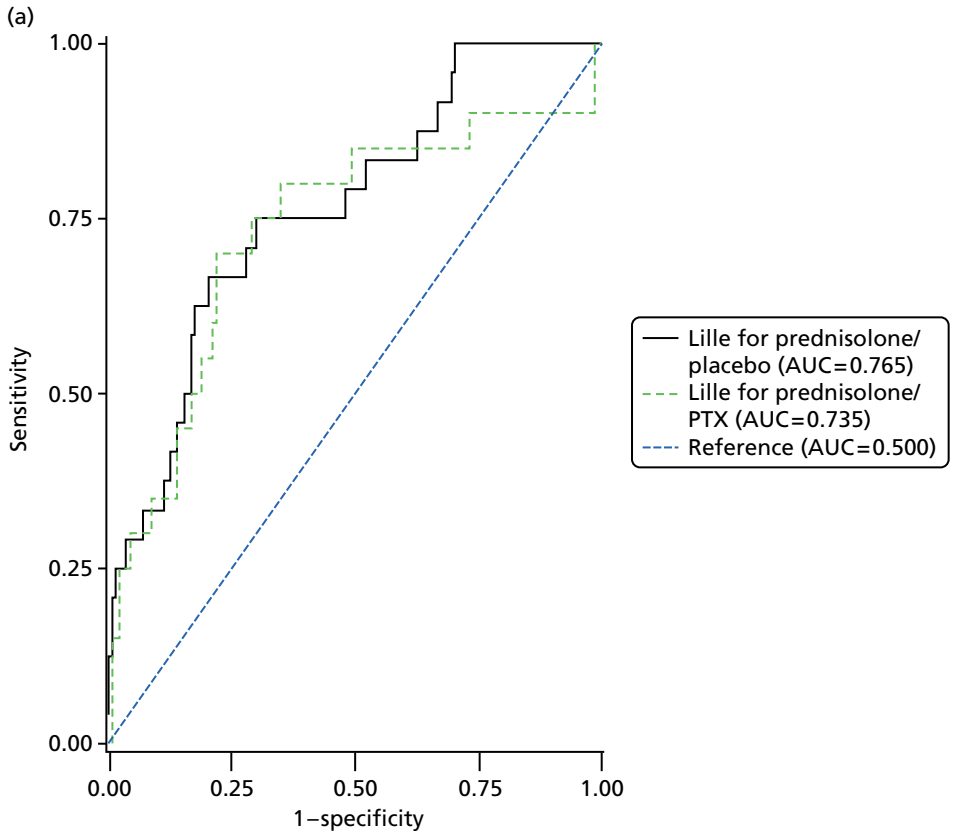

FIGURE 7 Receiver operating characteristics curves to compare the Lille score (prednisolone-treated group only). (a) 28-day mortality; (b) 90-day mortality; and (c) 1-year mortality. AUC, area under the curve. (*continued*)

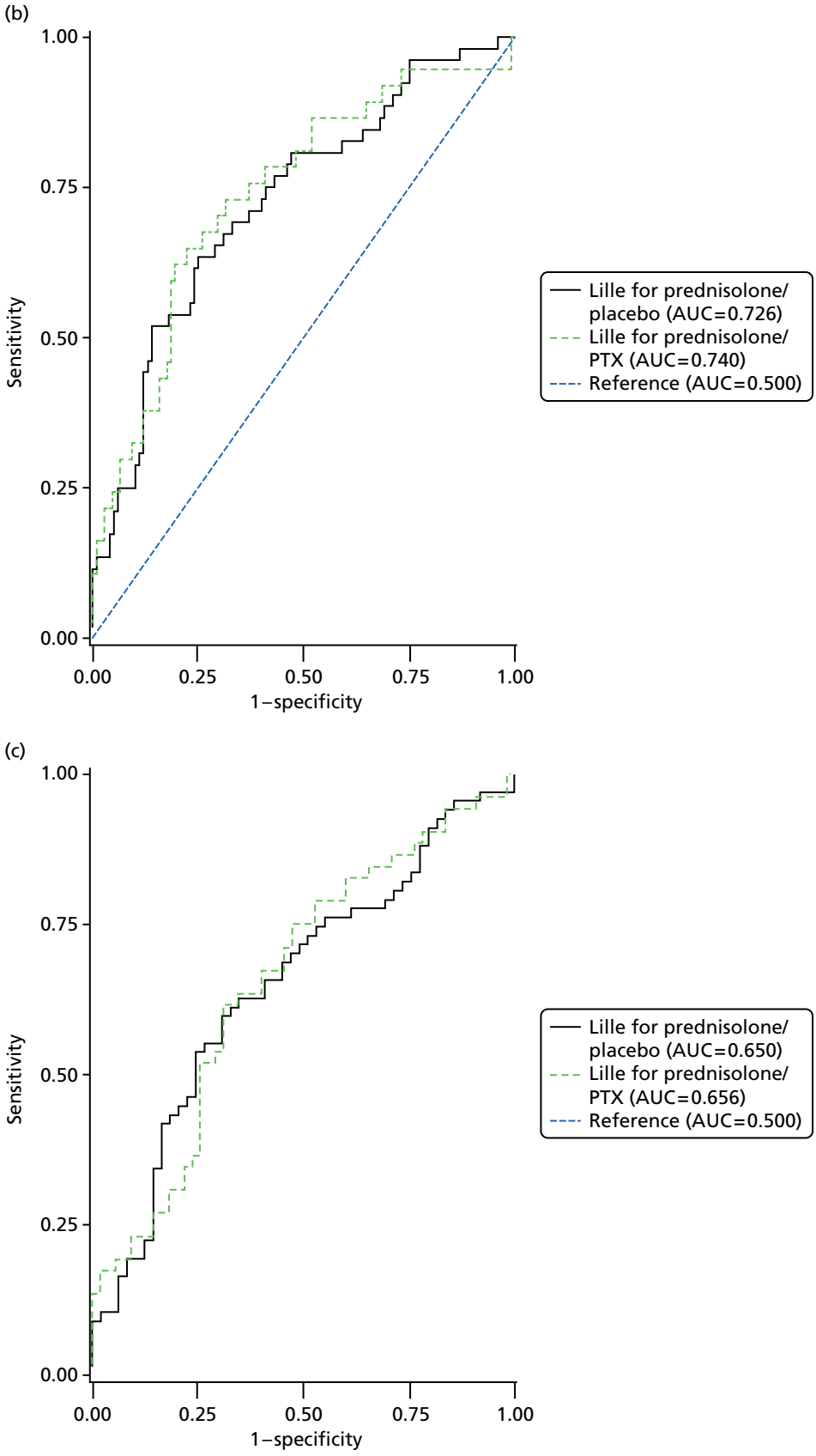

FIGURE 7 Receiver operating characteristics curves to compare the Lille score (prednisolone-treated group only). (a) 28-day mortality; (b) 90-day mortality; and (c) 1-year mortality. AUC, area under the curve.

(*Table 16* and *Figure 8*). At this survival rate the Lille score cannot be used to identify a population who may be suitable for liver transplantation as envisaged in the French pilot scheme.

Other predictive factors of mortality

Factors that influenced mortality at 28 days, 90 days and 1 year were analysed in univariate and multivariate analyses using baseline clinical and laboratory characteristics. *Table 17* summarises the findings from the univariate analysis. As expected, the variables significantly associated with mortality were encephalopathy, bilirubin, PT ratio, age, WBC count, creatinine and urea.

TABLE 16 Cox proportional hazards model for mortality up to 1 year by Lille categories (< 0.45 and ≥ 0.45) (prednisolone-treated group only)[a]

	Prednisolone group, Lille category ≥ 0.45	Prednisolone group, Lille category < 0.45
1-year mortality, (*n*/*N*)	51.9% (70/135)	24.6% (49/199)
Adjusted hazard ratio[b] (95% CI)	2.80 (1.94 to 4.03)	
p-value	< 0.001	

a Includes all randomised prednisolone-treated patients who have 1 year of data or died prior to or at 1 year and have data on the Lille score.
b Adjusted for risk (high or intermediate) at baseline.

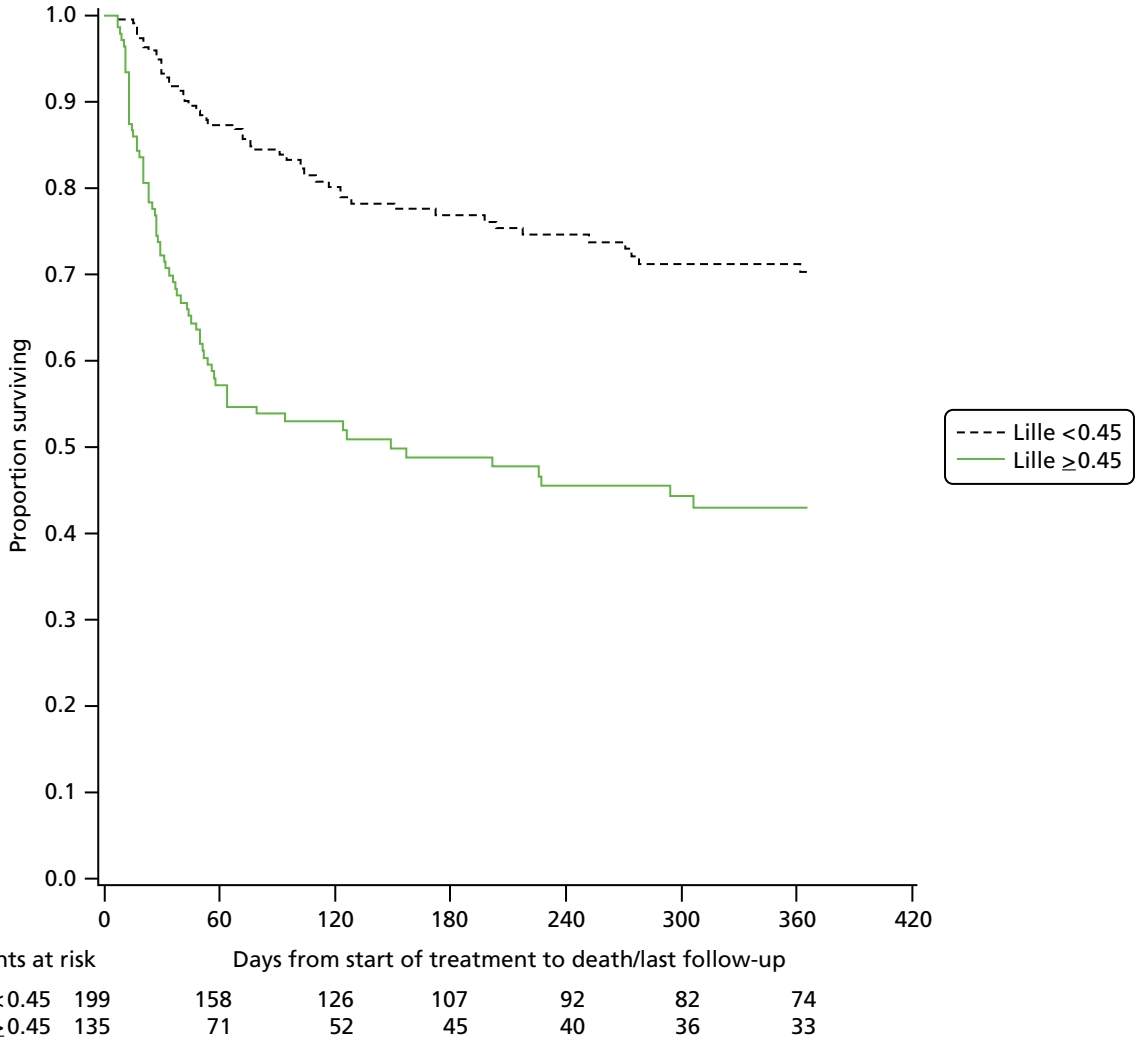

FIGURE 8 Kaplan–Meier graph for over survival by Lille score (< 0.45 and > 0.45) (prednisolone-treated group only).

TABLE 17 Univariate logistic regression analysis[a]

	28-day mortality[b]				90-day mortality[c]				1-year mortality[d]			
Variable (baseline value)	*n*	Odds ratio	95% CI	*p*-value	*n*	Odds ratio	95% CI	*p*-value	*n*	Odds ratio	95% CI	*p*-value
Pyrexia (temperature > 37.2 °C)	1049	0.658	0.374 to 1.159	0.147	965	0.664	0.422 to 1.045	0.077	745	0.780	0.503 to 1.210	0.267
Hypotension (SBP < 100mm/Hg)	1048	1.200	0.789 to 1.826	0.394	965	0.981	0.680 to 1.414	0.917	745	1.134	0.768 to 1.673	0.527
Tachycardia (pulse > 100 beats per minute)	1048	1.085	0.715 to 1.646	0.702	965	1.024	0.719 to 1.459	0.894	745	0.941	0.648 to 1.366	0.748
Alcohol intake (g/day)	1032	0.999	0.998 to 1.001	0.371	949	0.999	0.998 to 1.000	0.055	733	0.999	0.998 to 1.000	0.066
Albumin (g/l)	1044	0.988	0.961 to 1.016	0.405	960	0.986	0.963 to 1.009	0.235	743	0.993	0.969 to 1.017	0.557
ALP (U/l)	1031	0.998	0.996 to 1.000	0.074	948	0.999	0.997 to 1.000	0.119	733	0.999	0.997 to 1.000	0.178
Bilirubin (µmol/l)	1053	1.004	1.003 to 1.005	< 0.001	968	1.003	1.002 to 1.004	< 0.001	747	1.002[e]	1.001 to 1.003	0.001
Hepatic encephalopathy (presence vs. absence)	1009	3.700	2.587 to 5.292	< 0.001	929	2.352	1.720 to 3.217	< 0.001	717	2.191	1.539 to 3.118	< 0.001
PT ratio or INR[e]	1046	1.379[e]	1.101 to 1.727	0.005	962	1.909[e]	1.456 to 2.504	< 0.001	741	1.797	1.317 to 2.453	< 0.001
Age (years)	1053	1.049	1.032 to 1.067	< 0.001	968	1.050	1.035 to 1.065	< 0.001	747	1.036	1.021 to 1.051	< 0.001
WBCs (10^9/l)	1050	1.055	1.027 to 1.084	< 0.001	965	1.045	1.019 to 1.071	0.001	744	1.017	0.991 to 1.043	0.200
Urea (mmol/l)	1041	1.140[e]	1.099 to 1.182	< 0.001	958	1.142[e]	1.099 to 1.186	< 0.001	740	1.124	1.072 to 1.178	< 0.001
Creatinine (mg/dl)	1014	3.074[e]	2.315 to 4.082	< 0.001	930	3.070	2.266 to 4.161	< 0.001	720	2.695[e]	1.843 to 3.941	< 0.001

SBP, systolic blood pressure.
a Logistic regression model: 28-day (or 90-day/1-year) mortality = intercept plus prognostic factor.
b Includes all randomised patients who have at least 28 days of data or died prior to or on day 28 and have data for the corresponding variable.
c Includes all randomised patients who have at least 90 days of data or died prior to or on day 90 and have data for the corresponding variable.
d Includes all randomised patients who have at least 1 year of data or died prior to or on 1 year and have data for the corresponding variable.
e Model produced a significant result for presence of lack of fit at the 5% significance level.

Multivariate analysis was conducted by first including variables from the univariate analysis that were significant at the 5% level (*Table 18*). Variable selection was performed using backward elimination at the 5% significance level. The logistic regression analysis now demonstrated a clearly significant influence of prednisolone on 28-day mortality with an odds ratio of 0.609 (95% CI 0.409 to 0.909; $p = 0.015$). The odds ratio for 28-day mortality is adjusted for the following baseline factors: PT ratio or INR, bilirubin, age, WBC count, urea, creatinine and encephalopathy.

Further multivariate analysis was performed after the removal of patients who had outlying observations. A patient was classed as an outlier if they had a standardised residual with an absolute value of > 2 for the fitted variables. The results of this analysis (*Table 19*) show an important increase in the treatment effect of prednisolone (odds ratio 0.427, 95% CI 0.253 to 0.721); however, the treatment effect is confined to 28-day mortality with no impact on later time points.

Serious adverse events and deaths

As expected in this patient population, there were numerous adverse events and fatalities. As shown in *Table 20* the vast majority of deaths were liver related and include all the common complications of acute-on-chronic liver failure.

Serious adverse events are summarised in *Table 21*. Over 75% of SAEs were at Common Toxicity Criteria For Adverse Events grade 3 or above. SAEs resulted in mortality in 31% of cases. Details of the SAEs are given in *Table 22* and the incidence of infection is shown in *Table 23*.

Infections were frequent causes of SAEs and were reported more frequently in the prednisolone-treated group (see *Table 21*). Although infections led to mortality in approximately 35% of cases, the risk of mortality owing to an infection-related SAE was no greater in the prednisolone-treated group than in the controls, suggesting that corticosteroids increase the risk of infection but do not necessarily make the consequences of infection any worse (see *Table 22*). The most common site of infection was the lung and it is interesting to note that prednisolone had a more significant impact on pulmonary infections than any other site.

Alcohol behaviour

It is well recognised that return to heavy alcohol use is one of the most important determinants of medium- to long-term prognosis in patients admitted with AH or decompensated alcoholic cirrhosis.[32] We therefore sought to document the rates of recidivism in the STOPAH cohort. Accurate measures of alcohol consumption are difficult to ascertain, so we asked patients to report on if their alcohol consumption was:

1. completely abstinent
2. reduced to within government-defined safety limits
3. reduced but above safety limits (14 units/week for females, 21 units/week for males)
4. not reduced.

The alcohol use questions were asked at day 90 and 1-year follow-up.

Overall 45% of patients reported abstinence at 90 days and 36% reported abstinence at 1 year (*Table 24*). It is disappointing to record that only 22% of patients had attended one or more alcohol counselling sessions by the 90-day follow-up, which appears to have influenced alcohol recidivist behaviour. Inevitably there is a high proportion of missing data from these CRFs, reflecting a loss to follow-up, which in many cases is probably associated with recidivism.

TABLE 18 Multivariate logistic regression (after backward elimination)

	28-day mortality[a]				90-day mortality[b]				1-year mortality[c]			
Variable (baseline value)	*n*[d]	Odds ratio	95% CI	*p*-value	*n*[d]	Odds ratio	95% CI	*p*-value	*n*[d]	Odds ratio	95% CI	*p*-value
Prednisolone vs. no prednisolone	954	0.609	0.409 to 0.909	0.015	915	0.995	0.728 to 1.361	0.976	686	1.012	0.737 to 1.390	0.942
PTX vs. no PTX	954	1.105	0.743 to 1.642	0.623	915	0.909	0.664 to 1.244	0.550	686	0.948	0.690 to 1.303	0.742
PT ratio or INR[e]	954	1.381	1.129 to 1.691	0.002	915	1.936	1.428 to 2.626	< 0.001	686	1.643	1.171 to 2.305	0.004
Bilirubin (μmol/l)	954	1.002	1.001 to 1.003	0.003	915	1.002	1.001 to 1.003	< 0.001	–	–	–	–
Age (years)	954	1.050	1.029 to 1.071	< 0.001	915	1.050	1.033 to 1.067	< 0.001	686	1.030	1.014 to 1.047	< 0.001
WBCs (10^9/l)	954	1.030	1.002 to 1.060	0.037	–	–	–	–	–	–	–	–
Urea (mmol/l)	954	1.065	1.015 to 1.118	0.011	915	1.088	1.047 to 1.130	< 0.001	–	–	–	–
Creatinine (mg/dl)	954	1.564	1.048 to 2.332	0.028	–	–	–	–	686	2.127	1.452 to 3.117	< 0.001
Hepatic encephalopathy (presence vs. absence)	954	3.073	2.050 to 4.605	< 0.001	915	1.828	1.297 to 2.576	0.001	686	1.813	1.242 to 2.646	0.002

a Logistic regression model: 28-day mortality = intercept + prednisolone indicator + PTX indicator + PT ratio or INR + bilirubin + age + WBCs + urea + creatinine + hepatic encephalopathy.
b Logistic regression model: 90-day mortality = intercept + prednisolone indicator + PTX indicator + PT ratio or INR + bilirubin + age + urea + creatinine + hepatic encephalopathy. One iteration of backward elimination: model 2 removed WBCs as $p = 0.150$ in model 1, model 3 removed creatinine as $p = 0.062$.
c Logistic regression model: 1-year mortality = intercept + prednisolone indicator + PTX indicator + PT ratio or INR + age + creatinine + hepatic encephalopathy. Three iterations of backward elimination: model 2 removed bilirubin as $p = 0.398$ in model 1, model 3 removed urea as $p = 0.118$ in model 2).
d Number included in each model is the number of patients that has non-missing values for all of the variables included in the model.
e PT ratio or INR were fitted to all models. Therefore, the following variables were not included owing to the fact that they would be correlated with the PT: PT (patient), PT (control) and renal insufficiency (PT ratio > 1.3).

TABLE 19 Multivariate logistic regression analysis after removal of outlying observations after backward elimination

Variable (baseline value)	28-day mortality[a]					90-day mortality[b]					1-year mortality[c]				
	n removed[d]	*n*[e]	Odds ratio	95% CI	*p*-value	*n* removed[d]	*n*[e]	Odds ratio	95% CI	*p*-value	*n* removed[d]	*n*[e]	Odds ratio	95% CI	*p*-value
Prednisolone vs. no prednisolone	42	912	0.427	0.253 to 0.721	0.001	17	897	0.988	0.707 to 1.381	0.944	4	702	1.017	0.737 to 1.402	0.920
PTX vs. no PTX	42	912	1.282	0.770 to 2.136	0.340	17	897	0.960	0.686 to 1.343	0.811	4	702	0.939	0.681 to 1.296	0.703
PT ratio or INR[f]	42	912	1.903	1.417 to 2.558	< 0.001	17	897	3.627	2.566 to 5.126	< 0.001	4	702	3.397	2.283 to 5.054	< 0.001
Bilirubin (μmol/l)	42	912	1.003	1.001 to 1.005	< 0.001	17	897	1.002	1.001 to 1.003	< 0.001	4	–	–	–	–
Age (years)	42	912	1.094	1.063 to 1.125	< 0.001	17	897	1.068	1.049 to 1.087	< 0.001	4	702	1.032	1.015 to 1.049	< 0.001
WBCs (10^9/l)	42	912	1.069	1.022 to 1.118	0.004	17	897	1.039	1.005 to 1.074	0.024	4	–	–	–	–
Urea (mmol/l)	42	912	1.072	1.008 to 1.139	0.027	17	897	1.089	1.046 to 1.135	< 0.001	4	702	1.106	1.049 to 1.167	< 0.001
Creatinine (mg/dl)	42	912	2.070	1.276 to 3.360	0.003	–	–	–	–	–	4	–	–	–	–
Hepatic encephalopathy (presence vs. absence)	42	912	6.093	3.641 to 10.195	< 0.001	17	897	1.987	1.381 to 2.859	< 0.001	4	702	1.775	1.212 to 2.598	0.003

a Logistic regression model: 28-day mortality = intercept + prednisolone indicator + PTX indicator + PT ratio or INR + bilirubin + age + WBCs + urea + creatinine + hepatic encephalopathy. Two iterations of backward elimination: model 2 removed alcohol intake as $p = 0.850$ in model 1, model 3 removed ALP as $p = 0.183$ in model 2.

b Logistic regression model: 90-day mortality = intercept + prednisolone indicator + PTX indicator + PT ratio or INR + bilirubin + age + WBCs + urea + creatinine + alcohol intake + hepatic encephalopathy. Two iterations of backward elimination: model 2 removed alcohol intake as $p = 0.321$ in model 1, model 3 removed creatinine as $p = 0.093$ in model 2.

c Logistic regression model: 1-year mortality = intercept + prednisolone indicator + PTX indicator + PT ratio or INR + age + urea + hepatic encephalopathy. Three iterations of backward elimination: model 2 removed WBCs as $p = 0.578$ in model 1, model 3 removed bilirubin as $p = 0.415$ in model 2, model 4 removed creatinine as $p = 0.092$ in model 3.

d The number of observations removed because they were classed as outliers for the model fitted. An observation was an outlier if it had a standardised deviance residual of greater than 2 in absolute terms.

e The number included in each model is the number of patients that has non-missing values for all of the variables included in the model.

f PT ratio or INR were fitted to all models. Therefore, the following variables were not included owing to the fact that they would be correlated with the PTs: PT (patient), PT (control) and renal insufficiency (PT ratio > 1.3).

Notes

Four patients (prednisolone/placebo = 2, prednisolone/PTX = 2) have negative values of DF derived from PT and bilirubin at baseline. The PT values were queried but were not resolved prior to database lock; therefore, they are excluded from the table as all of the prognostic factors use PT in their derivation.

Bilirubin at day 7 was not included as this would correlate with bilirubin at baseline. Also, there were few patients who had a value for bilirubin at day 7 and this would decrease the number of patients included in the model, which reduces validity of results produced.

TABLE 20 Causes of death

Cause of death	Placebo/placebo ($n = 272$), n (%)	Prednisolone/placebo ($n = 274$), n (%)	Placebo/PTX ($n = 273$), n (%)	Prednisolone/PTX ($n = 273$), n (%)
Total number of deaths (during follow-up period)[a]	106 (39)	110 (40)	103 (38)	99 (36)
Liver related	92 (87)	99 (90)	93 (90)	81 (82)
Cardiac arrest (in addition to liver-related causes)	0	0	0	1 (1)
Encephalopathy	3 (3)	3 (3)	3 (3)	1 (1)
Gastric haemorrhage	1 (1)	0	1 (1)	0
Hepatic failure	39 (37)	36 (33)	40 (39)	38 (38)
Hepatobiliary disorders (other, alcohol abuse)	1 (1)	0	0	0
Infection-related causes	26 (25)	26 (24)	24 (23)	20 (20)
Multiorgan failure	10 (10)	16 (15)	9 (9)	5 (5)
Oesophageal varices haemorrhage	6 (6)	2 (2)	8 (8)	4 (4)
Pulmonary oedema	0	0	0	1 (1)
Renal failure-related causes	3 (3)	9 (8)	2 (2)	6 (6)
Upper GI haemorrhage	3 (3)	7 (6)	6 (6)	5 (5)
Non-liver related	2 (2)	1 (1)	0	4 (4)
Cardiac arrest (with no liver-related causes)	0	0	0	1 (1)
Infection-related causes	0	1 (1)	0	2 (2)
Intracranial haemorrhage	1 (1)	0	0	0
Nervous system disorders (other, hypoxic brain injury)	1 (1)	0	0	0
Respiratory failure	0	0	0	1 (1)
Both liver and non-liver related	7 (7)	10 (9)	8 (8)	6 (6)
Cardiac arrest (both liver-related and non-liver-related causes)	0	1 (1)	0	0
Duodenal haemorrhage	0	0	1 (1)	0
GI disorders (other, ruptured umbilical hernia)	0	0	0	1 (1)
Hepatic failure	1 (1)	5 (5)	3 (3)	1 (1)
Infection-related causes	1 (1)	1 (1)	1 (1)	1 (1)
Injury, poisoning and procedural complications (other, intoxicated)	0	1 (1)	0	0
Intracranial haemorrhage	0	1 (1)	0	1 (1)
Multiorgan failure	0	0	0	1 (1)
Nervous system disorders (other, traumatic brain injury)	2 (2)	0	1 (1)	0

TABLE 20 Causes of death (*continued*)

Cause of death	Placebo/placebo (*n* = 272), *n* (%)	Prednisolone/placebo (*n* = 274), *n* (%)	Placebo/PTX (*n* = 273), *n* (%)	Prednisolone/PTX (*n* = 273), *n* (%)
Pulmonary oedema	0	0	1 (1)	0
Renal failure-related causes	0	0	1 (1)	1 (1)
Respiratory failure	1 (1)	0	0	0
Stroke	2 (2)	0	0	0
Thromboembolic event	0	1 (1)	0	0
Unknown	4 (4)	0	2 (2)	7 (7)
General disorders and administration site conditions (other, bleeding)	1 (1)	0	0	0
Natural causes	0	0	0	1 (1)
Not available	3 (3)	0	2 (2)	5 (5)
Upper GI haemorrhage	0	0	0	1 (1)
Missing	1 (1)	0	0	1 (1)

TABLE 21 Summary of the SAEs reported

SAEs reported	Placebo/placebo (*n* = 272), *n* (%)	Prednisolone/placebo (*n* = 274), *n* (%)	Placebo/PTX (*n* = 273), *n* (%)	Prednisolone/PTX (*n* = 273), *n* (%)
Number of patients experiencing at least one SAE[a]	106 (39)	128 (47)	111 (41)	116 (42)
Total number of SAEs	136	184	145	159
Number of SAEs/SARs/SUSARs per patient[a]				
1	83 (31)	90 (33)	84 (31)	87 (32)
2	17 (6)	25 (9)	22 (8)	19 (7)
3	5 (2)	11 (4)	3 (1)	6 (2)
4	1 (< 0.5)	0	2 (1)	4 (1)
≥ 5	0	2 (1)	0	0
Principal investigator assessment (with reference to prednisolone/PTX)[b]				
SUSAR/SAE	0	3 (2)	0	1 (1)
SAE/SAR	1 (1)	2 (1)	5 (3)	0
SAR/SAE	5 (4)	24 (13)	13 (9)	19 (12)
SAR/SAR	4 (3)	11 (6)	4 (3)	3 (2)
SAE (no IMP taken)	3 (2)	6 (3)	3 (2)	13 (8)
SAE/SAE	123 (90)	138 (75)	120 (83)	123 (77)

continued

TABLE 21 Summary of the SAEs reported (*continued*)

SAEs reported	Placebo/placebo (*n* = 272), *n* (%)	Prednisolone/placebo (*n* = 274), *n* (%)	Placebo/PTX (*n* = 273), *n* (%)	Prednisolone/PTX (*n* = 273), *n* (%)
CTCAE grade[b]				
1 – mild	4 (3)	9 (5)	7 (5)	4 (3)
2 – moderate	28 (21)	41 (22)	30 (21)	42 (26)
3 – severe	37 (27)	49 (27)	36 (25)	41 (26)
4 – life-threatening	25 (18)	28 (15)	25 (17)	31 (19)
5 – death-related to SAE	42 (31)	57 (31)	47 (32)	41 (26)
Why the event was serious				
1 – resulted in death	45 (33)	54 (29)	47 (32)	37 (23)
2 – life-threatening	21 (15)	20 (11)	19 (13)	31 (19)
3 – required hospitalisation or prolongation of existing hospitalisation	68 (50)	110 (60)	75 (52)	91 (57)
4 – persistent or significant disability/incapacity	2 (1)	0	2 (1)	0
5 – congenital anomaly/birth defect	0	0	1 (1)	0
6 – medically significant event	0	0	1 (1)	0
Action taken following prednisolone[b]				
0 – none	–	131 (71)	–	107 (67)
1 – dose reduction	–	0	–	0
2 – treatment delayed	–	14 (8)	–	13 (8)
3 – treatment reduced and delayed	N/A	0	N/A	0
4 – treatment stopped	–	33 (18)	–	26 (16)
No IMP given	–	6 (3)	–	13 (8)
Action taken following PXT[b]				
0 – none	–	–	77 (53)	107 (67)
1 – dose reduction	–	–	2 (1)	1 (1)
2 – treatment delayed	–	–	10 (7)	12 (8)
3 – treatment reduced and delayed	N/A	N/A	1 (1)	0
4 – treatment stopped	–	–	52 (36)	26 (16)
No IMP given	–	–	3 (2)	13 (8)

CTCAE, Common Toxicity Criteria For Adverse Events; N/A, not applicable; SAR, serious adverse reaction; SUSAR, suspected unexpected serious adverse reaction.
a Percentages are based on the number of patients in the ITT population.
b Percentages are based on the number of SAEs.

TABLE 22 Summary of the events reported on the SAE form (sorted by % of total events)

System organ class and CTCAE event	Placebo/placebo (*n* = 272), *n* (%)	Prednisolone/placebo (*n* = 274), *n* (%)	Placebo/PTX (*n* = 273), *n* (%)	Prednisolone/PTX (*n* = 273), *n* (%)
Total number of SAEs	136	184	145	159
GI disorders	31 (23)	49 (27)	56 (39)	48 (30)
Ascites	7 (5)	13 (7)	14 (10)	13 (8)
Upper GI haemorrhage	5 (4)	12 (7)	10 (7)	18 (11)
Oesophageal varices haemorrhage	7 (5)	5 (3)	7 (5)	5 (3)
Abdominal pain	6 (4)	5 (3)	4 (3)	1 (1)
Pancreatitis	1 (1)	1 (1)	2 (1)	3 (2)
Vomiting	1 (1)	1 (1)	3 (2)	2 (1)
Abdominal distension	0	3 (2)	3 (2)	0
Diarrhoea	0	0	5 (3)	1 (1)
GI disorders: other	0	1 (1)	3 (2)	0
Oesophageal haemorrhage	1 (1)	2 (1)	0	0
Lower GI haemorrhage	2 (1)	1 (1)	0	0
Colonic perforation	0	2 (1)	0	0
Gastric haemorrhage	0	1 (1)	1 (1)	0
GI pain	0	0	1 (1)	1 (1)
Nausea	1 (1)	0	0	1 (1)
Rectal haemorrhage	0	1 (1)	0	1 (1)
Colitis	0	0	1 (1)	0
Colonic obstruction	0	0	0	1 (1)
Constipation	0	0	1 (1)	0
Duodenal perforation	0	1 (1)	0	0
Enterocolitis	0	0	0	1 (1)
Ileus	0	0	1 (1)	0
Infections and infestations	27 (20)	44 (24)	16 (11)	30 (19)
Lung infection[a]	11 (8)	20 (11)	6 (4)	18 (11)
Sepsis	8 (6)	11 (6)	6 (4)	3 (2)
Infections and infestations: other	6 (4)	9 (5)	3 (2)	3 (2)
Skin infection	1 (1)	1 (1)	1 (1)	2 (1)
Hepatic infection	0	2 (1)	0	1 (1)
Urinary tract infection	0	1 (1)	0	1 (1)
Abdominal infection	0	0	0	1 (1)
Endocarditis infective	1 (1)	0	0	0
Joint infection	0	0	0	1 (1)

continued

TABLE 22 Summary of the events reported on the SAE form (sorted by % of total events) (*continued*)

System organ class and CTCAE event	Placebo/placebo (*n* = 272), *n* (%)	Prednisolone/placebo (*n* = 274), *n* (%)	Placebo/PTX (*n* = 273), *n* (%)	Prednisolone/PTX (*n* = 273), *n* (%)
Hepatobiliary disorders	27 (20)	27 (15)	24 (17)	23 (14)
Hepatic failure	26 (19)	27 (15)	23 (16)	22 (14)
Hepatobiliary disorders: other	1 (1)	0	1 (1)	1 (1)
Renal and urinary disorders	10 (7)	10 (5)	13 (9)	8 (5)
Acute kidney injury	9 (7)	5 (3)	4 (3)	5 (3)
Renal and urinary disorders: other	1 (1)	5 (3)	9 (6)	3 (2)
Nervous system disorders	12 (9)	12 (7)	6 (4)	9 (6)
Encephalopathy	8 (6)	4 (2)	4 (3)	4 (3)
Seizure	1 (1)	3 (2)	2 (1)	2 (1)
Syncope	0	4 (2)	0	1 (1)
Stroke	2 (1)	0	0	1 (1)
Depressed level of consciousness	1 (1)	0	0	0
Intracranial haemorrhage	0	0	0	1 (1)
Nervous system disorders: other	0	1 (1)	0	0
Respiratory, thoracic and mediastinal disorders	7 (5)	12 (7)	9 (6)	11 (7)
Respiratory failure	1 (1)	5 (3)	1 (1)	5 (3)
Dyspnoea	2 (1)	4 (2)	2 (1)	1 (1)
Adult respiratory distress syndrome	0	0	5 (3)	1 (1)
Pulmonary oedema	1 (1)	1 (1)	0	2 (1)
Pleural effusion	0	1 (1)	0	1 (1)
Respiratory, thoracic and mediastinal disorders: other	0	1 (1)	0	1 (1)
Aspiration	1 (1)	0	0	0
Bronchospasm	1 (1)	0	0	0
Epistaxis	1 (1)	0	0	0
Hypoxia	0	0	1 (1)	0
General disorders and administration site conditions	8 (6)	6 (3)	7 (5)	9 (6)
Multiorgan failure	6 (4)	3 (2)	6 (4)	7 (4)
Malaise	1 (1)	2 (1)	0	1 (1)
General disorders and administration site conditions: other	1 (1)	0	1 (1)	0
Non-cardiac chest pain	0	1 (1)	0	0
Oedema limbs	0	0	0	1 (1)

TABLE 22 Summary of the events reported on the SAE form (sorted by % of total events) (*continued*)

System organ class and CTCAE event	Placebo/placebo (*n* = 272), *n* (%)	Prednisolone/placebo (*n* = 274), *n* (%)	Placebo/PTX (*n* = 273), *n* (%)	Prednisolone/PTX (*n* = 273), *n* (%)
Injury, poisoning and procedural complications	5 (4)	7 (4)	3 (2)	8 (5)
Fall	2 (1)	1 (1)	2 (1)	4 (3)
Injury, poisoning and procedural complications: other	0	4 (2)	1 (1)	2 (1)
Fracture	1 (1)	2 (1)	0	1 (1)
Arterial injury	1 (1)	0	0	0
Hip fracture	1 (1)	0	0	0
Seroma	0	0	0	1 (1)
Psychiatric disorders	4 (3)	2 (1)	1 (1)	4 (3)
Confusion	1 (1)	1 (1)	0	3 (2)
Hallucinations	2 (1)	0	0	0
Psychosis	0	0	1 (1)	1 (1)
Psychiatric disorders: other	0	1 (1)	0	0
Suicide attempt	1 (1)	0	0	0
Cardiac disorders	3 (2)	3 (2)	1 (1)	3 (2)
Cardiac arrest	1 (1)	1 (1)	1 (1)	2 (1)
Chest pain – cardiac	1 (1)	1 (1)	0	0
Atrial fibrillation	0	0	0	1 (1)
Myocardial infarction	0	1 (1)	0	0
Ventricular tachycardia	1 (1)	0	0	0
Metabolism and nutrition disorders	0	5 (3)	1 (1)	1 (1)
Hyperglycaemia	0	4 (2)	0	1 (1)
Dehydration	0	0	1 (1)	0
Hypoglycaemia	0	1 (1)	0	0
Investigations	0	3 (2)	1 (1)	2 (1)
Blood bilirubin increased	0	1 (1)	1 (1)	1 (1)
Investigations: other	0	2 (1)	0	1 (1)
Vascular disorders	1 (1)	1 (1)	2 (1)	2 (1)
Hematoma	1 (1)	1 (1)	2 (1)	1 (1)
Thromboembolic event	0	0	0	1 (1)
Blood and lymphatic system disorders	0	1 (1)	3 (2)	0
Anaemia	0	0	2 (1)	0
Haemolysis	0	0	1 (1)	0
Leucocytosis	0	1 (1)	0	0

continued

TABLE 22 Summary of the events reported on the SAE form (sorted by % of total events) (*continued*)

System organ class and CTCAE event	Placebo/placebo (*n* = 272), *n* (%)	Prednisolone/placebo (*n* = 274), *n* (%)	Placebo/PTX (*n* = 273), *n* (%)	Prednisolone/PTX (*n* = 273), *n* (%)
Musculoskeletal and connective tissue disorders	1 (1)	1 (1)	0	1 (1)
Back pain	0	1 (1)	0	0
Muscle weakness lower limb	0	0	0	1 (1)
Myositis	1 (1)	0	0	0
Skin and subcutaneous tissue disorders	0	1 (1)	1 (1)	0
Skin ulceration	0	1 (1)	0	0
Toxic epidermal necrosis	0	0	1 (1)	0
Social circumstances	0	0	1 (1)	0
Social circumstances: other	0	0	1 (1)	0

CTCAE, Common Toxicity Criteria For Adverse Events.
a A test for significance to compare the proportion of lung infection SAEs experienced between prednisolone and no prednisolone was computed [38 (69%) vs. 17 (31%) events]. The difference was found to be significant at the 5% significance level ($p = 0.005$), for which the null hypothesis is that the two proportions are equal.
Percentages are out of the total number of SAEs.

TABLE 23 Incidence of infection

Type of SAE and time point	Prednisolone (*n* = 551), *n* (%)	No prednisolone (*n* = 552), *n* (%)
Number of patients with an infection-related SAE	71 (13)	38 (7)
Before start of treatment	3 (1)	2 (< 0.5)
Day 1–28	50 (9)	31 (6)
Day 29–56	16 (3)	3 (1)
Multiple time points	2 (< 0.5)	2 (< 0.5)
Number of patients who died from an infection-related SAE[a]	25 (35)	12 (32)
Before start of treatment	1 (1)	2 (5)
Day 1–28	16 (23)	10 (26)
Day 29–56	8 (11)	0

a Denominator is the number of patients who had an infection-related SAE.
An infection-related SAE was any SAE with a system organ class of 'infections and infestations', which could happen from the date of informed consent up to 4 weeks after treatment completion (i.e. day 56).

TABLE 24 Alcohol consumption and counselling at day 90 and 1 year

Alcohol consumption and counselling at day 90[a]	Placebo/placebo (*n* = 183), *n* (%)	Prednisolone/placebo (*n* = 161), *n* (%)	Placebo/PTX (*n* = 160), *n* (%)	Prednisolone/PTX (*n* = 179), *n* (%)
Alcohol consumption at day 90 compared with last assessment				
Abstinent	88 (48)	74 (46)	65 (41)	80 (45)
Reduced drinking to below safety limits	18 (10)	19 (12)	12 (8)	10 (6)
Reduced drinking but above safety limits	9 (5)	10 (6)	17 (11)	12 (7)
Not reduced (i.e. still drinking as much as or more than when presented)	11 (6)	16 (10)	10 (6)	10 (6)
Missing	57 (31)	42 (26)	56 (35)	67 (37)
Has the patient attended one or more alcohol counselling session?				
No	76 (42)	80 (50)	69 (43)	70 (39)
Yes	45 (25)	36 (23)	29 (18)	38 (21)
Not still attending[b]	19 (42)	10 (28)	6 (21)	7 (18)
Still attending[b]	26 (58)	26 (72)	23 (79)	31 (82)
Missing	62 (34)	45 (28)	62 (39)	71 (40)
Alcohol consumption and counselling at 1 year[c]	**Placebo/placebo (*n* = 86), *n* (%)**	**Prednisolone/placebo (*n* = 80), *n* (%)**	**Placebo/PTX (*n* = 79), *n* (%)**	**Prednisolone/PTX (*n* = 81), *n* (%)**
Alcohol consumption at 1 year compared with last assessment				
Abstinent	43 (50)	29 (36)	24 (30)	23 (28)
Reduced drinking to below safety limits	5 (6)	12 (15)	9 (11)	8 (10)
Reduced drinking but above safety limits	4 (5)	8 (10)	5 (6)	4 (5)
Not reduced (i.e. still drinking as much as or more than when presented)	4 (5)	5 (6)	8 (10)	7 (9)
Missing	30 (35)	26 (33)	33 (42)	39 (48)
Has the patient attended one or more alcohol counselling session?				
No	28 (33)	31 (39)	35 (44)	26 (32)
Yes	22 (26)	21 (26)	11 (14)	13 (16)
Not still attending[d]	13 (59)	13 (62)	5 (45)	9 (69)
Still attending[d]	9 (41)	8 (38)	5 (45)	4 (31)
Missing on CRF (still attending question)	0	0	1 (9)	0
Missing	36 (42)	28 (35)	33 (42)	42 (52)

a Includes all randomised patients who have at least 90 days of data.
b Percentages are based on the number of patients who have attended one or more alcohol counselling session at day 90.
c Includes all randomised patients who have at least 1 year of data.
d Percentages are based on the number of patients who have attended one or more alcohol counselling session at 1 year.

The impact of abstinence (compared with other drinking patterns) at 90 days or mortality at 1 year was explored (see *Table 25*). These data indicate that even modest drinking, within government guidelines, is associated with increased mortality (odds ratio 2.17, 95% CI 1.07 to 4.39; $p = 0.03$) whereas resumption of previous levels of alcohol use results in approximately a threefold increased mortality compared with abstinence. These findings reinforce the need to promote and support abstinence in this patient group.

Sensitivity analyses

A sensitivity analysis of the primary end-point outcome was produced on 939 patients, excluding any patients who were randomised and followed up but were later found to be ineligible through central monitoring of baseline characteristics ($n = 114$). This was not substantially different from the primary analysis results. Detailed tables for the sensitivity analysis can be found in *Appendix 3*.

Subgroup analyses

We used the disease-specific scoring systems to subclassify patients into higher or lower severity categories to explore whether or not disease severity had an impact on response to treatment (*Tables 25* and *26*). We also analysed whether or not the risk classification used to stratify the trial randomisation influenced response to prednisolone.

Risk classification made no impact on the response to therapy, indicating that in the future patients admitted with complications of AH such as renal failure, GI bleeding and sepsis should be included in clinical trials without stratification once the complication has been stabilised.

Based on the odds ratio, it would appear that patients classified as having more severe disease by DF or MELD were more likely to benefit from steroids than patients with milder disease. However, these analyses did not reach statistical significance. In contrast, patients with a higher GAHS appear to derive less benefit from steroids than those with a lower score. This contrasts with the original report on GAHS that found a higher response in patients with a GAHS of > 9.[1,33]

TABLE 25 Univariate logistic regression analysis to assess the effect of alcohol consumption status at day 90 on 1-year mortality

	1-year mortality			
Alcohol consumption at day 90	*n*[a]	Odds ratio	95% CI	*p*-value
Not reduced (i.e. still drinking as much as or more than when presented) vs. abstinent	478	2.99	1.47 to 6.05	< 0.001
Reduced drinking but above safety limits vs. abstinent	478	2.28	1.07 to 4.86	0.032
Reduced drinking to below safety limits vs. abstinent	478	2.17	1.07 to 4.39	0.031

a Includes all randomised patients who have at least 1 year of data, or died prior to or on 1 year and have data for alcohol consumption status at day 90.
Logistic regression model: 1 year mortality = intercept + alcohol consumption status at day 90.

TABLE 26 Subgroup analyses for 28-day, 90-day and 1-year mortality by factorial design

	Prednisolone vs. no prednisolone									PTX vs. no PTX								
	Day 28			Day 90			1 year			Day 28			Day 90			1 year		
Disease severity classification	*n*	Odds ratio (95% CI)	*p*-value	*n*	Odds ratio (95% CI)	*p*-value	*n*	Odds ratio (95% CI)	*p*-value	*n*	Odds ratio (95% CI)	*p*-value	*n*	Odds ratio (95% CI)	*p*-value	*n*	Odds ratio (95% CI)	*p*-value
Risk																		
High	218	0.72 (0.39 to 1.33)	0.291	203	0.86 (0.49 to 1.52)	0.603	151	0.82 (0.42 to 1.61)	0.561	218	1.33 (0.72 to 2.46)	0.364	203	0.94 (0.53 to 1.65)	0.818	151	1.09 (0.55 to 2.15)	0.803
Intermediate	835	0.72 (0.48 to 1.07)	0.105	765	1.08 (0.79 to 1.49)	0.631	596	1.07 (0.78 to 1.48)	0.665	835	0.98 (0.66 to 1.46)	0.914	765	0.98 (0.71 to 1.35)	0.896	596	0.94 (0.68 to 1.29)	0.698
DF[a,b]																		
< 55	515	0.56 (0.31 to 1.04)	0.066	473	1.04 (0.65 to 1.67)	0.875	335	1.20 (0.78 to 1.85)	0.414	515	0.93 (0.51 to 1.70)	0.825	473	0.75 (0.46 to 1.20)	0.232	335	0.71 (0.46 to 1.10)	0.125
≥ 55	534	0.82 (0.54 to 1.23)	0.337	492	1.06 (0.74 to 1.51)	0.767	409	0.91 (0.60 to 1.37)	0.647	534	1.06 (0.71 to 1.60)	0.767	492	1.02 (0.71 to 1.46)	0.922	409	1.13 (0.75 to 1.71)	0.551
MELD																		
< 18	351	1.16 (0.44 to 3.01)	0.765	320	1.34 (0.71 to 2.52)	0.371	232	1.47 (0.86 to 2.49)	0.155	351	1.83 (0.69 to 4.83)	0.224	320	0.65 (0.34 to 1.25)	0.197	232	0.78 (0.46 to 1.33)	0.363
≥ 18	660	0.69 (0.47 to 1.00)	0.052	607	1.04 (0.75 to 1.45)	0.816	485	0.89 (0.62 to 1.29)	0.541	660	0.98 (0.68 to 1.42)	0.927	607	1.02 (0.73 to 1.41)	0.922	485	1.00 (0.69 to 1.44)	0.986
GAHS																		
< 9	535	0.47 (0.23 to 0.97)	0.040	495	0.93 (0.58 to 1.51)	0.783	367	1.07 (0.71 to 1.62)	0.743	535	1.61 (1.21 to 2.39)	0.583	495	0.81 (0.50 to 1.31)	0.380	367	0.91 (0.61 to 1.38)	0.668
≥ 9	398	0.79 (0.50 to 1.26)	0.324	362	1.15 (0.76 to 1.75)	0.507	285	1.02 (0.63 to 1.65)	0.942	398	0.99 (0.63 to 1.58)	0.982	365	1.05 (0.69 to 1.60)	0.817	385	1.01 (0.63 to 1.64)	0.959

a Logistic regression model: 28-day (or 90-day/1-year) mortality = intercept + prednisolone indicator + PTX indicator, by subgroup.
b Cut-off of 55 used for DF as this was close to the median value at baseline.
Number included in each model is the number of patients in the ITT populations at each of the three time points that have non-missing data for the subgroup variables at baseline.

Chapter 5 Health economic evaluation

Introduction

This chapter reports the two elements of the health economics work conducted as part of this study: a standard gamble (SG) preference elicitation exercise and an economic evaluation conducted as part of the STOPAH trial.

Part 1: eliciting health state utility values for cirrhosis owing to alcohol abuse and cirrhosis owing to viral hepatitis – a standard gamble exercise

A necessary part of any resource allocation decision is to measure the outcomes that arise from treatment, which in the case of many treatments for liver disease, are improvements in QoL in addition to improvements in life expectancy. One way of presenting the outcomes of a treatment is in the form of quality-adjusted life-years (QALYs).

There are a number of methods available to measure health-related QoL. Generic questionnaire-based tools such as the EQ-5D[34] or Short Form questionnaire-36 items (or 12-item version),[35] and disease-specific questionnaires, such as the chronic liver disease questionnaire and liver disease QoL questionnaire,[36,37] can be used. The limitation of some of these is that their methods of 'scoring' changes in patients QoL are ill-suited to informing decisions about how best to allocate our scarce health-care resources, that is they do not produce utility values.

There are a number of direct valuation techniques that can be used to assign values or utilities to different health states. These include the rating scale, SG and time trade-off methods.[38] Of these, the SG is considered to be the gold standard as it is based on expected utility theory, which is the dominant paradigm in economics on how the preferences of an individual behave.[39]

A systematic review of health-state utility values in liver disease found that few estimates exist for liver disease in the literature.[40] This systematic review spanned 1966–2006 and found only 30 studies that measured utility values in liver disease, none of which related to alcoholic liver disease. An updated review conducted in September 2014 found no significant further relevant research.

Given this paucity of existing evidence, the objectives of this study were to directly elicit QoL in patients with cirrhosis owing to alcohol abuse and cirrhosis owing viral hepatitis. A further objective was to relate direct measures of health-state valuations (using SG), to indirect measures of patients' health valuations as measured by the European Quality of Life-5 Dimensions-3 Levels (EQ-5D-3L).

Methods

Quality of life was measured in a sample of patients with liver cirrhosis caused by alcohol abuse or viral hepatitis, using the SG technique. Utility values were obtained for patients in three health states that correspond to different stages of cirrhosis. All utility values are measured on a scale from 0 to 1, for which 0 represents death and 1 represents full health. Each participant was also asked to value their own health that day using both the SG and EQ-5D-3L. Standard clinical and laboratory variables were also collected from patient medical records and were used to calculate the Child–Pugh and MELD scores for each participant.

Standard gamble methodology

The SG is the gold standard technique for eliciting preferences for both chronic and temporary health states.[38] *Figure 9* provides an example of a SG when respondents are presented with two alternatives.

In *Figure 9*, alternative A is a health state with certainty, and alternative B is a health state with uncertainty, that is, there is a gamble associated with this choice. Alternative B comprises two possible health states: the patient returns immediately to full health or dies immediately. Each of these states has probabilities attached to them. The respondent makes the choice between remaining in the chronic health state or taking the gamble with a probability that they will return to full health (p) and a probability that they would die (1 – p). The probabilities within alternative B continue to change in an iterative fashion until the participant becomes indifferent to alternatives A and B.

Development of health-state descriptions

Three health states were developed reflecting the Child–Pugh classification system for cirrhosis.[41,42] The Child–Pugh classification system is based on three categories A, B and C. A is the mildest form of cirrhosis and C is the most severe and is commonly used in studies of liver cirrhosis.

The health-state descriptions were based on previous research.[43] Mason *et al*.[43] based the health state descriptions on expert opinion, the domains of chronic liver disease questionnaire[36] and reviews of the literature (liver disease and cirrhosis of the liver). The health state descriptions for Child–Pugh classifications A, B and C can be seen in *Box 1*.

Sample

We aimed to recruit 100 participants with alcoholic cirrhosis and 100 participants with cirrhosis owing to viral hepatitis to be able to represent of all levels of disease severity. Five centres took part in the study (St Mary's Hospital, London; King's College Hospital, London; Queens Medical Centre, Nottingham; Freeman Hospital, Newcastle upon Tyne; Glasgow Royal Infirmary, Glasgow). Participants aged ≥ 18 years who were able to give informed consent were recruited to hepatology wards and outpatient clinics by a research nurse. Those individuals who had difficulty in understanding written or verbal information in English, or had deterioration in liver function (change in Child–Pugh score of > 2) in the last 2 weeks were excluded from the study. Patients were given a minimum of 24 hours to consider the study and ask questions, after which they (or their legal representatives) were asked to give written informed consent to participate, on the trial informed consent form. Full ethical approval for the study was granted (REC reference: 09/MRE09/59).

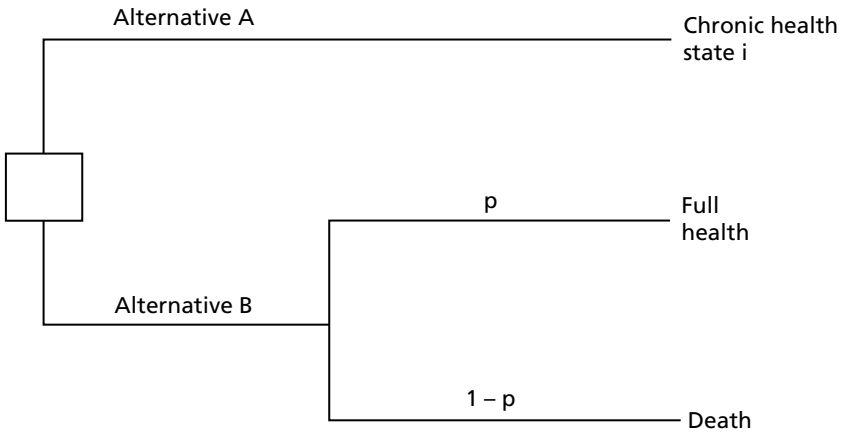

FIGURE 9 Example of a SG. p, probability.

BOX 1 Description of health state used in SG exercise

State X = Child–Pugh classification A

- You will not develop jaundice.
- You will feel tired some of the time.
- You will not develop a swollen stomach.
- You will not be troubled by itching.
- A little bit of the time you will be worried about your condition getting worse.

State Y = Child–Pugh classification B

- Occasionally you will have jaundice.
- You will feel tired a lot of the time.
- Occasionally you will have swollen stomach.
- Some of the time you will feel depressed about your condition.
- You may become confused.
- You may occasionally be admitted to hospital.

State Z = Child–Pugh classification C

- You will often have jaundice.
- You will feel tired all of the time.
- It is likely you will develop a swollen stomach.
- You may be troubled by itching.
- Some of the time you will be depressed about your condition.
- You may become confused.
- You are likely to spend several weeks a year in hospital.
- You may vomit blood.
- You may develop liver cancer.

State O

- Your own health today.

Interview process

Research nurses ($n = 10$) were fully trained in the use of the SG methodology. The interviews were conducted using a visual aid know as a chance board.[44] In an interview lasting approximately 30–45 minutes, each participant was asked to complete the SG exercise and value the three Child–Pugh health states and a fourth health state 'their own health today'. In addition, participants completed an EQ-5D-3L questionnaire and a set of questions on demographic and socioeconomic characteristics (see *Data analysis*). Standard clinical and laboratory variables from patient's medical records were used to calculate the Child–Pugh and MELD scores for each patient.

Data analysis

Data were collected on the utility scores of four health states, estimated using SG methodology, when the value of a health state is:

$$U(\text{state i}) = p \times U(\text{full health}) + (1 - p) \times U(\text{death}), \quad (1)$$

or

$$U(\text{state i}) = p. \quad (2)$$

Results

A sample of 142 people with liver cirrhosis was recruited, 91 with alcoholic liver disease and 51 with viral hepatitis (see *Table 27*). The majority of participants had the mildest form of liver cirrhosis (58%). Also in *Table 27* are the demographic and socioeconomic characteristics of the sample. The majority of respondents were male (75%), 39% were married, 44% were unable to work and 82% were white British. Thirty-one per cent of the sample stated that they had no formal qualifications. The average age was 57 years.

The mean values for each of the health states can be seen in *Table 28*. When health state X represents Child–Pugh A, health state Y represents Child–Pugh B and health state Z represents Child–Pugh C. Mean values follow the expected patterns, for which the most severe health states have the lowest health-state values. Paired sample *t*-tests were used to assess if valuations for state Child–Pugh A (X) were significantly different from Child–Pugh B (Y) and valuations for Child–Pugh B (Y) significantly different from Child–Pugh C (Z). It was expected that these differences would be statistically significant, reflecting the differences in severity between states and the different severities of the disease. These tests were performed for the full sample and for both subsamples. All were found to be statistically significant at the 5% level. Further *t*-tests were conducted to test whether or not there was concordance between own health, as valued by the SG, and own health as valued by the EQ-5D. As expected, no significant differences were found.

Discussion

This study measured QoL directly using the SG and indirectly via the EQ-5D-3L. The aim was to explore health-state values for various stages of disease, represented by different Child–Pugh classifications (A, B and C). However, the collection of SG data from participants with Child–Pugh C was complicated by the presence of encephalopathy in many of these participants, which interfered with the ability of participants to give informed consent and also to understand the concept. This problem accounts for the relatively small sample size in this category.

TABLE 27 Study sample characteristics

Sample characteristics	All (%)	Alcoholic cirrhosis (%)	Viral (%)
Study site			
Mary's London	29 (20)	16 (18)	13 (25)
King's London	33 (23)	20 (22)	13 (25)
Nottingham	7 (5)	3 (3)	4 (8)
Glasgow	34 (24)	27 (30)	7 (14)
Newcastle upon Tyne	39 (27)	25 (27)	14 (27)
Total	142	91	51
Respondents classified in each Child–Pugh classification (missing n = 6)			
A	79 (58)	48 (55)	31 (63)
B	46 (34)	30 (35)	16 (33)
C	11 (8)	9 (10)	2 (4)
Demographic data			
Gender (missing = 2)			
Male	105 (75)	65 (72)	40 (78)
Female	36 (25)	25 (28)	11 (22)

TABLE 27 Study sample characteristics (*continued*)

Sample characteristics	All (%)	Alcoholic cirrhosis (%)	Viral (%)
Marital status (missing = 2)			
Married	54 (39)	39 (44)	15 (29)
Single	35 (25)	17 (19)	18 (35)
Divorced/separated	46 (33)	29 (33)	17 (33)
Widowed	5 (4)	4 (5)	1 (2)
Employment status (missing = 3)			
Employed	30 (22)	17 (19)	13 (26)
Retired	38 (27)	28 (32)	10 (20)
At home/not looking for work	3 (2)	3 (3)	0 (0)
Unable to work	61 (44)	35 (40)	26 (51)
Unemployed	5 (4)	3 (3)	2 (4)
Other	2 (1)	2 (2)	0 (0)
Qualifications (missing = 4)			
None	42 (31)	27 (31)	15 (30)
GCSEs (1–5)	23 (17)	17 (20)	6 (12)
GCSEs (> 5)	14 (10)	10 (12)	4 (8)
A-levels (1–2)	2 (1)	2 (2)	0 (0)
A-levels (> 2)	4 (3)	1 (1)	3 (6)
Degree	10 (7)	8 (9)	2 (4)
Higher degree	10 (7)	7 (8)	3 (6)
NVQ1	4 (3)	3 (3)	1 (2)
NVQ2	5 (4)	1 (1)	4 (8)
NVQ3	2 (1)	2 (2)	0 (0)
NVQ4	3 (2)	0 (0)	3 (6)
Other	19 (14)	9 (10)	10 (20)
Ethnicity (missing = 2)			
White British	115 (82)	79 (89)	36 (71)
White Irish	6 (4)	4 (5)	2 (4)
White other	8 (6)	2 (2)	6 (12)
Mixed white Asian	2 (1)	0 (0)	2 (4)
Asian Indian	2 (1)	0 (0)	2 (4)
Black Caribbean	2 (1)	2 (2)	0 (0)
Black African	2 (1)	1 (1)	1 (2)
Black other	1 (1)	1 (1)	0 (0)
Other	2 (1)	0 (0)	2 (4)

A-level, Advanced level; GCSE, General Certificate of Secondary Education; NVQ, National Vocational Qualification.

TABLE 28 Health state valuations

	Mean	Standard deviation	*p*-value (X > Y and Y > Z)	*p*-value (EQ-5D SG and EQ-5D)
Full sample				
State X ($n = 141$)	0.645	0.305	–	–
State Y ($n = 142$)	0.566	0.567	< 10.000**	–
State Z ($n = 142$)	0.429	0.279	< 0.0001**	–
Own health ($n = 139$) (SG)	0.578	0.311	–	–
Own health ($n = 139$) (EQ-5D)	0.602	0.381	–	0.570
ALD sample				
State X ($n = 91$)	0.632	0.314	–	–
State Y ($n = 91$)	0.537	0.319	< 0.0001**	–
State Z ($n = 91$)	0.417	0.282	< 0.0001**	–
Own health ($n = 89$) (SG)	0.572	0.317	–	–
Own health ($n = 89$) (EQ-5D)	0.6187	0.376	–	0.372
Viral sample				
State X ($n = 50$)	0.669	0.289	–	–
State Y ($n = 51$)	0.616	0.283	0.003*	–
State Z ($n = 51$)	0.451	0.272	< 0.0001**	–
Own health ($n = 50$) (SG)	0.582	0.297	–	–
Own health ($n = 50$) (EQ-5D)	0.572	0.389	–	0.729

*, significantly different at 5% level, X compared with Y and Y compared with Z.
**, significantly different at 1% level, X compared with Y and Y compared with Z.

At the aggregate level, results of health-state valuations were in line with a priori expectations, when Child–Pugh A was preferred to Child–Pugh B and Child–Pugh B preferred to Child–Pugh C. As expected, at the aggregate level, statistically significant differences between the values of health states Child–Pugh A, B and C were found, adding credibility to the results. As disease severity increases, health-state valuations significantly decrease. Convergent validity was tested by examining valuations of an individual's own health today, as measured by the SG and EQ-5D. These values were not significantly different, suggesting high levels of convergent validity between the two measures values. These results hold for the full sample analysis and subsample analyses.

At the individual level there were some inconsistencies in both the ranking of health states and the valuation exercise. Such results are not uncommon at the individual level.[45] Importantly, valuations at the aggregate level are consistent with a priori expectations.

Conclusion

As seen in the systematic review[40] and a search of literature since 2008 there is little empirical research into health-state valuations in liver disease generally, and in alcoholic liver disease specifically, that can be used to derive QALYs for use in economic evaluations and decision-making. This study has measured QoL using the SG method to obtain utility values for three different classifications of disease: Child–Pugh A, B and C, and in two separate populations: those suffering from cirrhosis owing to alcoholic liver disease and those with cirrhosis owing to viral hepatitis. This research offers a unique opportunity to estimate QALYs on future trials for liver cirrhosis when QoL can be mapped to Child–Pugh classification of the trial participant.

Part 2: economic evaluation

The aim of the economic evaluation was to determine which single treatment (PTX or prednisolone), dual treatment (PTX and prednisolone) or standard care (placebo) is the most cost-effective option when treating AH. The following comparisons have been made:

i. PTX vs. no PTX
ii. prednisolone vs. no prednisolone
iii. PTX vs. prednisolone vs. PTX and prednisolone vs. placebo.

Table 29 provides an illustrative example of the 2 × 2 factorial trial design adopted for the STOPAH study. It is important to note that the assumption made with this form of trial design is that there is no interaction between PTX and prednisone.

The perspective (i.e. whose costs and benefit are considered relevant) of the trial is of the health service provider (NHS), hence only costs borne by the NHS were considered. Patient time and travel costs were not collected for this trial as it was anticipated that the response rate would be low and information of poor quality. For the economic evaluation both within-trial and a model-based evaluations were conducted. The role of the latter was to extrapolate from the short trial follow-up over the estimated lifetime of the trial participants. For both the within-trial and model-based analyses, the following outcomes were reported:

- costs to the NHS at 28 days and 1 year
- mortality at 28 days and 1 year
- QALYs over 1 year based on responses to the EQ-5D-3L questionnaire administered at discharge, 90 days and 1 year.

For all outcomes there were missing data after 28 days (the treatment phase) for those participants who were recruited towards the end of study, owing to an extension in the recruitment phase.

Methods overview

This economic evaluation includes a within-trial analysis and a model-based analysis.

First, a cost-effectiveness analysis was conducted that allowed us to calculate the incremental cost per additional survivor at 28 days. Second, we planned to repeat the analysis for a 1-year time horizon; however, as estimates of the incremental cost per additional survivor may be difficult to interpret according to standard decision-making criteria within the NHS,[46] a cost–utility analysis was also performed.

When we came to analyse the data there were few responses ($n = 192$) to the use of medical services questionnaire administered at 1 year across the four treatment arms. If we used these responses, even with imputations for the missing responses, there would be significant concerns that there would have selection biases given the nature and extent of the missing data. However, to extrapolate to 1 year and longer follow-up we used a Markov model. In a Markov model people move between discrete states of health over time. Within a Markov model, movement between states can take place only once every period of time, called a cycle length. In this model we have adopted a cycle length of 1 day, as this will allow us to

TABLE 29 A 2 × 2 factorial design for the STOPAH study

	Placebo for A	Treatment for A
Placebo for B	OO (placebo for A and placebo for B)	AO (treatment for A and placebo for B)
Treatment for B	OB (placebo for A and treatment for B)	AB (treatment for A and treatment for B)

A, prednisolone; B, PTX; O, placebo.

cross-validate estimates with the within-trial analysis data and be sufficiently sensitive to changes over time. In this model the costs and utility values were collected over the treatment phase and 90 days to estimate the total average cost per day and average QoL utility per day for each treatment arm.

Incremental cost-effectiveness

In both the within-trial analysis and the model-based analysis, data on costs and effects of the interventions were combined to obtain an incremental cost-effectiveness ratio (ICER). For the within-trial analysis, this was performed by calculating the mean difference in costs between the interventions and control group over the difference in effect between the interventions and control group. This gives us the cost per additional unit of effectiveness (e.g. survivors or QALYs) gained – a more effective but more costly intervention relative to a less costly but less effective intervention. A similar approach was used to estimate the ICER for the model-based analysis.

Cost data collection

NHS costs and frequency of use of health-care services

The costs to the NHS depended on the use of NHS resources during the initial treatment phase and during the 1-year follow-up. The frequency of resource use was calculated for each participant to generate the average cost per patient in each arm of the trial. Any potential differences in costs and outcomes between patient groups were identified from this analysis. All unit costs were collected from routine sources, for example NHS Reference Costs.[47]

During the initial admission, the discharge visit form captured information on the length of admission and the procedures that were performed during this time. After participants were discharged, their resource use was captured using the use of medical services self-reported questionnaires completed at 90 days and 1 year, and the day-90 and 1-year visit forms. The use of the medical services questionnaire allowed us to divide resource use into secondary care, primary care and social care. Secondary care resources included inpatient stays, day-case visits, accident and emergency visits, and outpatient visits. Primary care resources included general practitioner (GP) practice and home visits, nurse practice and home visits, and home and office visits with a social worker or care worker. If a participant was prescribed a medication, the number of prescriptions they received was recorded. To summarise, data collection and costs can be split into thee areas:

1. initial treatment costs
2. secondary care costs
3. primary and other community care costs.

Initial treatment costs

The treatments being examined during the STOPAH trial were relatively low-cost medications. The cost of each treatment is based on the dosage and length of treatment (28 days), specified in the trial protocol. The cost for PTX and prednisolone came from the *British National Formulary* (BNF).[48]

Other costs incurred prior to discharge were collected via the discharge form and incorporated into the total cost of initial treatment. This form contained information on length of initial admission (based on trial recruitment and discharge dates), medical procedures and medical events that could occur during the treatment phase. A number of participants had a trial start date after their date of discharge, so we assumed the length of their initial admission was 0 days. If a participant was missing their discharge date then assumptions on the date of discharge were informed by data in their records. The length of initial admission for participants without a discharge date but whose date of death was relatively close to the end of their treatment phase was taken as 28 days. Participants who withdrew from the trial or had a date of death at a later stage of the trial were given a length of admission that was the average of their patient group. This simple form of imputation was used because there were relatively few people (9.5% of the trial population) in this category.

The discharge form also contained information on whether or not a participant was admitted to an intensive care unit (ICU) during the initial treatment phase and the length of time that they spent there. The initial length of admission was multiplied by the NHS cost per night for an inpatient stay on a general ward. A higher cost was used for an ICU stay and the initial length of admission was reduced by the number of nights in ICU to prevent any double counting. In terms of procedures, the discharge form included information on imaging, including use of computerised tomography and magnetic resonance imaging, as well as a wide range of procedures. All these events were costed using the NHS Reference Costs.[47] These reference costs were, however, amended by removing any element of hospitalisation, as these costs were already included. This was done by estimating the average cost per day from the cost of the relevant non-elective long-stay reference cost, then multiplying this cost per day by the average length of stay for the national tariff. This hospitalisation cost was then subtracted from the procedure cost to give a cost of the procedure without any length-of-stay component. An example of this calculation is the procedure cost for banding varices. This is captured in the NHS tariff 'GI with single intervention'. The non-elective cost of a long stay is £1582 and the average length of stay for this procedure is 4.0 days. The average length of stay is multiplied by £265, the cost of a non-elective bed stay. This results in a procedure cost of £522. This is a rough estimate of the procedure cost including staff costs. The same calculation was used to estimate the cost adverse events procedures.

A participant may receive some procedures more than once. If this happened then the number of procedures that a participant had was multiplied by the cost per procedure. Unfortunately, no information was captured on the number of standard procedures performed on the discharge form, which may have caused an underestimation of the cost of discharge procedures. In some cases, additional information on the number of procedures performed was available in the 'Comments' data set and when this was the case, these data were included in the costs for that participant and used to estimate the total cost of discharge procedures per randomised arm (note: the comments data set contained participant information that was provided by participants and clinicians). For participants for whom the discharge information was completely missing, the median costs of discharge procedures were imputed on the data available for the randomised arm. The median estimate was used instead of the mean, as this was felt more representative as the mean value was skewed because a number of participants in each arm had extremely high discharge procedure costs caused by an extended admission to ICU.

Nutritional supplements were provided for participants who were admitted into hospital. Assumptions about the nature and type of these supplements were informed by clinical advice from a dietitian. It was assumed that participants with this illness would need 1.5 kcal/ml sip feed and a suitable nutritional substitute would have 300 kcal and 12 g of protein in a 200-ml bottle. It was assumed that participants took the nutritional supplement five times per day every day for the length of their initial admission; we have made the assumption that participants would take this nutritional drink during only their initial admission as their sole source of nutrition. Sensitivity analysis was used to test this assumption. The unit cost of this nutritional supplement was taken from the BNF.[48]

We planned to cost GI bleed, sepsis and renal failure as common adverse events. The data to estimate the cost for the treatments for sepsis and GI bleed were captured in the concomitant medication form and procedures to diagnose sepsis, and to diagnose and treat GI bleeds, were captured in the discharge and adverse events forms. Renal failure was based on the NHS tariff for 'hospital haemofiltration'. Using clinical advice it was assumed that the haemofiltration would be performed continuously. This information was collected during the treatment phase at 7-day intervals and as a result we assumed that the participant had the treatment for a period of 7 days continuously. Using the treatment day forms at days 7, 14, 21 and 28 we could determine if the renal failure was resolved or if the participant needed further treatment. If the patient had renal failure at day 28 and it was not resolved, we assumed that he/she continued the treatment for another 7 days. Sensitivity analysis was conducted around this assumption in our model-based analysis.

Laboratory tests were routinely performed on participants during the treatment phase and follow-up period to monitor their disease progression and the effects of the initial medical treatments being examined in the trial. The laboratory tests include, but are not limited to, FBC, glucose, urea and PT. Every participant was due to have these laboratory tests during their initial admission and at each follow-up visit. There could be a discrepancy between randomised arms if one group had a shorter/longer initial admission on average than the other arms. The unit costs for these laboratory tests were collected from the Royal Victoria Infirmary's Newcastle University Hospitals costing tool.[49] Within the discharge visit laboratory test data, there were additional data for some participants who had 'subsequent visit' laboratory test information. The information collected on laboratory tests at this visit was excluded from our analysis because it was not consistently recorded across sites.

Adverse events

Participants' adverse event information was collected via the adverse event form. Within this form, adverse events that occurred were categorised as: no action taken; change in concomitant medication; investigations and hospitalisation. Only investigations and hospitalisations were costed, as medications were captured in the concomitant medication form. Some of the investigations recorded included being 'observed' and 'getting vital signs checked', and we made the assumption that these costs would form part of standard care when in hospital. All other investigations were costed using national tariffs. The calculation that was used for generating procedure costs was used here, unless it was explicitly stated in the adverse event form or 'comments' data set that the participant was treated in a different health-care setting (e.g. outpatient). One participant was referred to their GP but it was assumed this cost would be captured in the use of medical services questionnaire. We divided the adverse events into the three time periods: 28 days, 90 days and 1 year. We presented the average resource use of the adverse event investigations performed during the 28-day treatment phase. This time period was used because there was limited information about who was missing data at 90 days and 1 year. Information on adverse events that required hospitalisation was collected via the SAE form. This form recorded the length of admission if a participant was hospitalised and the severity of their admission. If a person died while in hospital we used the participant's date of death to estimate their length of admission. For some participants, data on hospitalisation were duplicated and, when this could be verified, the duplicate data were removed from our analysis. We assumed, based on clinical advice, that unless stated otherwise in the adverse event or 'comments' data set, a participant was treated on a general ward rather than in ICU. Sensitivity analysis was conducted around this assumption to determine what effect an ICU admission would have on total average costs; this was conducted in the model-based analysis.

Appendix 4 provides a detailed description of all the unit costs used, where they were sourced from and what CRF they relate to. *Appendix 5* provides a detailed description on the average number of adverse event investigations performed at 28 days.

Concomitant medications

During the treatment phase and 1-year follow-up, additional medications prescribed to the participant were collected via the concomitant medication form. Many participants used a large number of additional medications for a wide variety of indications, including, for example, fungal nail infection treatments, paracetamol for headaches, vitamin supplements and medications for alcohol withdrawal. There were over 22,000 additional medications reported by participants. These were broken down into categories so the results could be presented in an interpretable manner. We used a dummy variable for final indication, for which final indication was based on the 15 BNF categories, with a further three additional categories added for medications that did not fall under these BNF categories: emergency treatment of poisoning, wound dressing and procedure medication. This allowed us to segregate our medications based on established indications. All costs of concomitant medications were sourced from the BNF when possible. Imputations for participants with a reduced follow-up period were based on the final indication of concomitant medications instead of the frequency of use of individual medications.

Secondary care costs

The secondary care resources that were used by participants were accident and emergency visits, inpatient admissions, day admissions and outpatient visits (*Table 30*). Each hospital visit/admission incured a different tariff depending on the length of stay and type of admission. When possible, tariffs directly related to the condition were used (e.g. hepatology outpatient visit), as it is a more accurate indication of costs. The frequency of use of these health-care resources were multiplied by the national tariff to get the total follow-up cost per patient. In the 'comments' data set it was reported that some participants were too ill to complete their use of medical services questionnaire but patient records were used when possible to provide this information. As a result, many participants have information on only secondary care resource use and imputations were made for the use of primary care services. If participants were admitted during the follow-up period, the ward that they were admitted to was not collected in the use of medical services questionnaire. We decided to remove participant-reported inpatient admissions from our primary analysis, as this information duplicated that collected in the CRFs and that the CRF data would be a more accurate representation of length of admissions, as it was not subject to recall bias. Sensitivity analysis was conducted around the different wards participants could potentially be admitted to: general ward and ICU.

Primary care and social care costs

The health-care resources available to participants at a primary care setting included GPs, practice and district nurses, social workers and care workers (see *Table 30*). All visits with these practitioners could occur at the health-care practice or at the participant's house. We distinguished between the different locations of each consultation to account for the different costs associated with each consultation.

If a participant attended a health-care practitioner during the follow-up period they could be given a prescription. Information on the type of medication prescribed was not recorded in the use of medical services questionnaire but the number of prescriptions each participant received was recorded. We applied a generic cost for prescriptions from the Personal Social Services Research Unit.[50] This is calculated as the average cost of prescriptions prescribed by GPs over the previous year. We have not incorporated

TABLE 30 Unit costs: follow-up

Costs	Unit	Unit cost (£)	Source
Follow-up costs: hospital			
A&E visit	Per visit	115	Resource-use questionnaire/NHS tariffs
Day case	Per day	693	Resource-use questionnaire/NHS tariffs
Inpatient: general ward	Per night	265	Resource-use questionnaire/NHS tariffs
Outpatient	Per visit	213 (hepatology department)	Resource-use questionnaire/NHS tariffs
Primary care			
GP: at practice	Per visit	45	Resource-use questionnaire/PSSRU[50]
GP: at home	Per visit	114	Resource-use questionnaire/PSSRU[50]
GP nurse: at practice	Per visit	11.37	Resource-use questionnaire/PSSRU[50]
GP nurse: at home	Per visit	39	Resource-use questionnaire/PSSRU[50]
Prescription cost	Per script	7.73	Resource-use questionnaire/PSSRU[50]
Social services			
Social worker: at home (including travel costs)	Per visit	190.80	Resource-use questionnaire/PSSRU[50]
Social worker: at office	Per visit	159	Resource-use questionnaire/PSSRU[50]

A&E, accident and emergency; PSSRU, Personal Social Services Research Unit.

participants' costs if they had to pay for a prescription but we have accounted for the cost associated with administering prescriptions. The number of prescriptions was multiplied by this cost for each participant.

A care worker collected information on visits but owing to misinterpretations about care workers this cost was excluded from our analysis. A number of patients interpreted a care worker visit to their home as home help and they had reported daily visits. Using the 'comments' data set we could identify the different interpretations to the care worker question (e.g. alcohol counselling, daily support from a sponsor) and as a result we could not accurately assign a unit cost for this resource use and it was hence omitted from our analysis. The average resource use for this service will be presented but take caution in its interpretation, as some people classed it as home care. Social workers tend to work on an individual level with clients, so to estimate the length of appointments was difficult. Using expert opinion we assumed a visit with a social worker, regardless of location, would have a duration of 1 hour. If a visit was a home visit we included an extra 12 minutes to incorporate travelling time, using the Personal Social Services Research Unit GP average travelling time as a guide.[50]

We used a simple imputation of including the upper interquartile range for participants' use of medical services information. We used this range, as the mean value was very skewed because of a small number of participants who used very high quantities of medical services and the majority of participants who responded to the questionnaires had little resource use. This assumption was tested in the sensitivity analysis in our model-based analysis.

Information on alcohol counselling was collected at the 90-day and 1-year visits. There was no information on the frequency of counselling attendances, just whether or not the participant had attended and, if they had attended, whether or not they attended more than once. Information reported in the 'comments' data set was used when many participants reported the number of times they attended counselling sessions. These data were used to estimate an average number of visits, which was then assigned to every participant who reported that they had seen an alcohol counsellor. A sensitivity analysis was conducted around this assumption. For participants with missing counselling information we assumed a proportion of each treatment arm had at least one counselling session, dependent on the number of people in each arm who reported having at least one counselling session. This was explored in our sensitivity analysis by assuming that everyone had at least one counselling session and assuming that no one had a counselling session, unless it was reported.

All health-care resource use was captured via the use of a medical services questionnaire at 90 days and 1 year. We initially planned on presenting the combined average health-care resource use over the complete follow-up period of 1 year; however, owing to the extent of missing data (caused by the reduction in the trial follow-up period because of the recruitment extension) this resulted in participants having a length of follow-up considerably shorter than 1 year. As a result of this we presented the average 90-day health-care resource use and average 1-year health-care resource use separately.

Effectiveness measures

The economic evaluation was conducted with different outcome measures at 28 days (mortality only) and 1 year (mortality and QALYs).

Mortality rates

The initial economic evaluation considered the incremental cost per additional survivor, with mortality rates assessed at 28 days and 1 year. For participants recruited later in the trial, we have information only on their mortality status until the end of the trial follow-up period, which was 31 March 2014.

Quality-adjusted life-years based on responses to the European Quality of Life-5 Dimensions-3 Levels

Quality-adjusted life-years were generated from the utility values derived from responses to the EQ-5D-3L during the follow-up period. For participants without a 1-year follow-up, their probability of mortality was estimated with a survival function. QALYs were based on the utility estimates from the EQ-5D-3L at

discharge, 90 days and 1 year during the follow-up period using the areas under the curve approach. The EQ-5D-3L is commonly used to assess the QoL for participants with different medical conditions. It allows us to generate a QoL profile for participants and make comparisons between different groups of patients. The EQ-5D-3L questionnaire describes health status in five dimensions. These dimensions are mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each of these dimensions has three levels, regressing from the best possible situation in the dimension to the worse possible situation in the dimension. The number of dimensions and levels within each dimension provide 243 possible health states for participants to choose from.

The responses to the EQ-5D-3L questionnaire were transformed using a standard algorithm[51] to produce a health-state utility score at scheduled intervals during the follow-up for each participant in the four treatment groups. Utility values for perfect health and death are 1 and 0, respectively. The area under the curve approach puts a time weight onto each utility score. The time-weighted average of the scores based on the responses to the EQ-5D-3L measured at discharge, 90 days and at the end of measurement period (1 year) allows us to generate QALY values for each participant.[52] Missing baseline EQ-5D-3L scores occurred as some participants were unable to complete the questionnaires at their time of admission. Missing utility values at 90 days and 1 year were estimated based on additional information provided by the participant's status from the 'comments' data set, with imputations based on utility values reported by participants in similar health states.

Baseline utility values were estimated based on how this patient group is usually admitted to hospital (typically as emergency cases, and unconscious and needing urgent medical treatment). Some participants have their discharge extremely close to their start date, so the assumed baseline utility values can be tested in a sensitivity analysis using these utility values. However, an issue may arise with the utility scores generated from the EQ-5D-3L because it allows participants to have negative utility values, that is, a health state worse than death. If this appeared to be the case a small positive utility value (0.01) was imputed for baseline utility values to prevent participants experiencing a utility gain if they die.

The average utility scores for each treatment arm were presented for the three data collection points: discharge, 90 days and 1 year. The results presented are only for survivors who had completed the EQ-5D-3L questionnaire. It is important to note that for the initial presentation of results, no utility value was imputed for participants who died or did not complete the EQ-5D-3L questionnaire. Fewer than 2% of participants at each data collection point partially completed the questionnaire. The missing values for some participants could be estimated using information in the 'comments' data set, for example one participant was missing information on 'usual activities' but in the 'comments' data set it was written that they were 'in a care home', hence it was assumed that they were unable to perform their usual activities as a result of this. An extreme imputation method was used for participants with missing information and no additional information in the 'comments' data set; a 'worst case scenario' was assumed for these missing values. This imputation method was explored in the sensitivity analysis.

In the economic analysis it was assumed that if a participant died during the trial, their utility value was assumed to be 0 for every utility data collection point after their date of death. To estimate the utility values of the remaining participants we used the interquartile range of the utility values provided by participants still alive at each data collection point. A number of participants were 'too ill' to complete the EQ-5D-3L questionnaire at the different visits and hence we imputed their utility value as the lower interquartile range estimate of their treatment arm. If a participant 'did not attend' a scheduled visit we assumed that these participants were too ill to attend the visit and hence could not complete the questionnaire. We assumed that these participants would have a slightly greater utility value than participants who were 'too ill' to complete the questionnaire, as they would be able to complete the questionnaire if they had attended their scheduled visit. As a result we estimated their utility value to be the higher value of the interquartile range of their treatment arm. For all other missing utility values a multiple imputation technique was applied. These assumptions were explored in our sensitivity analysis in our model-based analysis.

The methods, results and sensitivity analyses will be presented separately into two sections: within-trial analysis and model-based analysis.

Within-trial analysis

Our main analysis was the relative cost-effectiveness of the study treatments at 28 days.

Methods

The cost-effectiveness analysis at 28 days will be replicated for three analyses using the methods below.

Four-arm comparison: within-the-table analysis

A four-arm comparison was undertaken as the primary analysis. This comparison considers each treatment arm of the 2 × 2 factorial trial design and compares it as if they were mutually exclusive individual strategies; PTX *or* prednisolone, PTX *and* prednisolone or standard care (placebo). This allows us to identify the effects of PTX and prednisolone when there is no interaction between the two treatments. The within-the-table analysis looked at the costs, outcomes and net benefits for the four treatment arms. The results were presented on the cost-effectiveness plane from which a cost-effectiveness frontier was generated. The cost-effectiveness frontier allows us to determine the treatment option that maximises net benefits at our chosen ceiling ratio. For example, in a cost–utility analysis, the National Institute for Health and Care Excellence typically adopted a threshold value for society's willingness to pay of £30,000 per QALY;[46] however, it is unclear what the ceiling ratio is for the cost per additional survivor.

As noted in the previous paragraph, the proposed primary is unbiased but is not efficient; the trial was not powered to detect differences between the four separate arms. A four-arm comparison is arguably more useful for the economic analysis but lacks statistical power so conclusions are based on the balance of probabilities.

At-the-margins analysis

An at-the-margins approach was used to present a two-arm comparison between PTX and no PTX, and prednisolone and no prednisolone. It is essentially analysing the data as if it were two overlapping, but independent, two-arm randomised controlled trials. Costs and outcomes were presented for all those participants who received PTX versus those not receiving PTX (prednisolone and placebo plus prednisolone plus PTX vs. placebo and placebo plus placebo and PTX) and all those participants who received prednisolone versus those not receiving prednisolone (placebo and PTX plus prednisolone and PTX vs. placebo and placebo plus prednisolone and placebo) (see *Table 29*). The simple effect of each treatment is the mean difference in outcomes across all participants who have received the treatment compared with all the participants who have not received the treatment. This analysis assumes there is no interaction between treatments. This analysis is efficient but the results will be biased if the interaction of the two treatments does not equal 0.

Sensitivity analysis

For the within-trial analysis, stochastic sensitivity analyses were performed. A stochastic sensitivity analysis was undertaken to allow presentation of the level of variance around outcome measures included in the cost–utility analysis. Uncertainty surrounding the cost-effectiveness ratio was addressed using the bootstrapping technique. The results of the bootstrapping simulation were presented on the 'cost-effectiveness plane', which highlights the preferred treatment option. If the results lie in the north-west or south-east quadrants the preferred treatment is clear, as one option dominates the other (i.e. is less costly and more effective). If the results lie in the north-east or south-west quadrants, the decision as to which is the preferred treatment is less clear (i.e. one option may be less costly but also less effective, or more effective but at greater cost); the ICER may aid this decision. The bootstrapping was also used to estimate CIs for both costs and effects from the four-arm and at-the-margins analyses. A cost-effectiveness acceptability curve was also used to present the probability of a treatment being cost-effective based on a range of values for society's willingness to pay. In the four-arm comparison we compared the bootstrapped results of the mean costs and mean outcomes for each treatment option against every other treatment option.

An extreme sensitivity analysis was conducted on our base-case analysis removing the most costly 10% of the participants from each treatment arm. There was a number of participants across the treatment arms that had an extended period of admission and this appears to be a key cost driver in our analysis.

Results

There were 1103 participants recruited into the STOPAH trial. Four participants who withdrew from the trial and withdrew consent for their information to be used were excluded from the analysis, seven participants who were randomised incorrectly were also excluded; this was consistent with the statistical analysis. One participant was randomised after they died so they were also excluded from the health economics analysis. Of the 1091 participants left in our analysis, 223 were affected by the trial extension for recruitment and we have information only on their 28-day status. Nearly 400 participants died during the trial period and only 223 participants were alive and completed the 1-year trial follow-up.

The 1091 participants that were used in the health economic analyses were evenly dispersed across the four treatment arms; 272 receiving standard care (placebo), 274 receiving prednisolone and placebo, 273 receiving PTX and placebo and 272 receiving the dual treatment (prednisolone and PTX).

Costs were estimated during the initial admission and during the follow-up period. We initially planned on presenting the combined total average cost of follow-up for each treatment arm, however, owing to the reduction in the follow-up period because of the extended recruitment period we decided to present follow-up costs individually.

Tables 31 and *32* describe the average health-care resource use by the randomised arms at 90 days for both liver-related and non-liver-related use of services. There were 415 use of medical services questionnaires returned at day 90, and of these, five were completely blank. For our analysis we assumed

TABLE 31 Follow-up resource use at 90 days: liver related

Follow-up resource use – day 90	OO		AO		OB		AB	
	n[a]	Mean (SD)	*n*	Mean (SD)	*n*	Mean (SD)	*n*	Mean (SD)
A&E	106	1.05 (5.65)[b]	104	1.55 (9.01)	94	0.52 (1.12)	107	0.62 (1.01)
Inpatient	106	12.17 (23.17)	104	12.11 (24.73)	94	10.36 (17.76)	107	11.38 (21.52)
Day case	106	0.10 (0.39)	104	1.08 (8.83)	94	0.31 (1.25)	107	0.53 (3.54)
Outpatient	106	1.91 (3.58)	104	2.87 (9.66)	94	1.72 (3.41)	107	1.22 (1.57)
GP practice	106	2.38 (8.92)	104	1.31 (1.78)	94	3.45 (13.02)	107	1.51 (1.96)
GP home	106	0.23 (1.10)	104	0.09 (0.28)	94	0.12 (0.67)	107	0.15 (0.55)
Practice nurse	106	1.75 (9.62)	104	0.62 (1.73)	94	0.33 (1.14)	107	0.77 (2.41)
District nurse	106	0.17 (1.04)	104	0.13 (0.71)	94	0.79 (5.08)	107	0.03 (0.22)
Social worker: home	106	0.23 (1.96)	104	0.05 (0.29)	94	0.00 (0.00)	107	0.21 (1.10)
Social worker: office	106	0.05 (0.40)	104	0.01 (0.98)	94	0.07 (0.55)	107	0.11 (0.59)
Care worker: home	106	0.23 (1.12)	104	0.66 (2.93)	94	0.74 (3.54)	107	0.79 (5.12)
Care worker: office	106	0.89 (5.90)	104	0.07 (0.53)	94	0.57 (2.61)	107	0.25 (1.30)
Prescription	106	5.43 (18.56)	104	2.10 (4.46)	94	2.54 (9.34)	107	2.06 (5.07)

A, prednisolone; A&E, accident and emergency; B, PTX; O, placebo; SD, standard deviation.
a *n* is the number of participants in each arm who responded to the questionnaire.
b How to interpret the table: the average A&E attendances from those in the placebo group who responded to the questionnaire was 1.05 visits.

TABLE 32 Follow-up resource use at 90 days: non-liver related

Follow-up resource use – day 90	OO		AO		OB		AB	
	n[a]	Mean (SD)	*n*	Mean (SD)	*n*	Mean (SD)	*n*	Mean (SD)
A&E	106	0.77 (5.706)	104	0.19 (1.080)	94	0.14 (0.521)	107	0.28 (0.799)
Inpatient	106	2.72 (11.600)	104	3.01 (14.390)	94	1.20 (9.387)	107	0.86 (3.583)
Day case	106	0.04 (0.191)	104	0.02 (0.138)	94	0.00 (0.000)	107	0.04 (0.191)
Outpatient	106	1.10 (8.753)	104	0.59 (1.568)	94	0.14 (0.499)	107	0.33 (1.106)
GP practice	106	0.38 (0.990)	104	0.81 (1.817)	94	0.23 (0.663)	107	0.66 (1.447)
GP home	106	0.06 (0.361)	104	0.05 (0.256)	94	0.00 (0.000)	107	0.04 (0.305)
Practice nurse	106	0.98 (8.746)	104	0.65 (3.547)	94	0.39 (2.515)	107	0.37 (2.365)
District nurse	106	0.35 (2.143)	104	0.10 (0.704)	94	0.11 (0.695)	107	0.05 (0.319)
Social worker: home	106	0.08 (0.686)	104	0.08 (0.534)	94	0.03 (0.230)	107	0.01 (0.097)
Social worker: office	106	0.14 (1.199)	104	0.05 (0.490)	94	0.03 (0.309)	107	0.00 (0.000)
Care worker: home	106	0.08 (0.329)	104	1.05 (8.891)	94	0.20 (1.380)	107	0.45 (2.819)
Care worker: office	106	0.15 (0.903)	104	0.08 (0.534)	94	0.01 (0.103)	107	0.06 (0.408)
Prescription	106	3.40 (15.077)	104	1.04 (3.424)	94	0.43 (1.092)	107	0.56 (1.319)

A, prednisolone; A&E, accident and emergency; B, PTX; O, placebo; SD, standard deviation.
a *n* is the number of participants in each arm who responded to the questionnaire.

that any participant who partially completed the questionnaire left their other responses blank, as they were not applicable to them. This assumption was explored in the sensitivity analysis. For participants who had died during the 90-day follow-up period, their resource use was automatically imputed as 0. This could cause an underestimation in our costs as they may have used some services within the data collection period but before they died. In our preliminary analysis of the health-care resource use data we discovered some unintuitive results reported by participants; an example is reporting 215 inpatient nights in the last 90 days. For our analysis we decided to cap these unintuitive responses at 90.

Tables 33 and *34* present similar data to *Tables 31* and *32*, but for the 1-year follow-up period.

Table 31 highlights that the prednisolone only treatment arm has a higher, on average, liver-related secondary care resource use. The number of inpatient nights for non-liver-related conditions for the prednisolone-only treatment arm is also the highest out of all treatment arms (see *Table 32*). The primary care resource use for the PTX/placebo treatment arm is relatively low compared with the other treatment arms, especially in relation to GP visits and office visits with a social worker or care advisor. The standard care treatment arm has the highest prescription use reported for both liver- and non-liver-related causes. For all other health-care resource use the pattern of resource use is similar for all treatment arms. A similar pattern of resource use is observed in *Tables 33* and *34*.

Utility data were collected at discharge, day 90 and 1 year but were used only in our 1-year analysis. Preliminary utility scores for each treatment arm suggest that, on average, the standard treatment arm has the highest utility values at discharge. At 90 days, the PTX-only arm has the highest average utility score but at 1 year, PTX only has the lowest average utility score. At 1 year, the standard treatment arm again has the highest average utility score (see *Table 35*). When interpreting *Table 35* caution needs to be taken as the results are presented only for survivors who completed the EQ-5D-3L; this is because this table is also used as a data source for the model-based economic evaluation. The numbers with utility data in each treatment arm decrease by approximately 30% from discharge to the 90-day visit, with a further 40%

TABLE 33 Follow-up resource use at 1 year: liver related

Follow-up resource use – 1 year	OO		AO		OB		AB	
	n[a]	Mean (SD)	*n*	Mean (SD)	*n*	Mean (SD)	*n*	Mean (SD)
A&E	52	0.83 (1.593)[b]	51	1.16 (2.838)	44	1.07 (2.688)	45	0.53 (1.392)
Inpatient	52	12.69 (23.772)	51	7.63 (16.949)	44	8.36 (26.049)	45	5.04 (13.588)
Day case	52	0.46 (1.540)	51	0.75 (2.719)	44	0.95 (4.529)	45	0.24 (0.679)
Outpatient	52	2.71 (2.637)	51	2.53 (3.331)	44	2.48 (2.724)	45	2.51 (2.873)
GP practice	52	1.77 (2.784)	51	2.61 (4.336)	44	2.45 (4.196)	45	3.40 (6.257)
GP home	52	0.83 (4.993)	51	0.12 (0.711)	44	0.52 (2.140)	45	0.13 (0.457)
Practice nurse	52	0.98 (5.020)	51	0.88 (2.414)	44	1.09 (2.400)	45	0.42 (1.288)
District nurse	52	0.02 (0.139)	51	0.37 (1.777)	44	0.45 (2.107)	45	0.02 (0.149)
Social worker: home	52	0.19 (1.253)	51	0.73 (5.040)	44	1.98 (10.866)	45	0.04 (0.208)
Social worker: office	52	0.12 (0.832)	51	0.00 (0.000)	44	0.00 (0.000)	45	0.00 (0.000)
Care worker: home	52	0.12 (0.583)	51	0.02 (0.140)	44	2.00 (10.920)	45	0.31 (2.087)
Care worker: office	52	0.83 (5.044)	51	0.47 (2.845)	44	0.20 (1.357)	45	4.62 (27.062)
Prescription	52	9.38 (37.795)	51	4.35 (10.451)	44	3.66 (6.463)	45	3.98 (7.617)

A, prednisolone; A&E, accident and emergency; B, PTX; O, placebo; SD, standard deviation.
a *n* is the number of participants in each arm who responded to the questionnaire.
b How to interpret the table: the average A&E attendances from those in the placebo group who responded to the questionnaire was 0.83 visits.

TABLE 34 Follow-up resource use at 1 year: non-liver related

Follow-up resource use – 1 year	OO		AO		OB		AB	
	n[a]	Mean (SD)	*n*	Mean (SD)	*n*	Mean (SD)	*n*	Mean (SD)
A&E	52	0.33 (0.585)	51	0.31 (0.812)	44	0.30 (1.133)	45	0.27 (0.580)
Inpatient	52	4.04 (20.981)	51	1.41 (4.904)	44	2.52 (10.551)	45	1.18 (3.695)
Day case	52	0.10 (0.358)	51	0.18 (0.623)	44	0.05 (0.211)	45	0.04 (0.208)
Outpatient	52	0.90 (1.902)	51	1.14 (1.960)	44	0.86 (1.534)	45	0.78 (2.055)
GP practice	52	1.77 (3.246)	51	1.96 (3.013)	44	1.61 (2.879)	45	1.11 (2.102)
GP home	52	0.21 (1.391)	51	0.18 (0.888)	44	0.66 (2.167)	45	0.27 (1.372)
Practice nurse	52	0.73 (2.011)	51	0.31 (0.678)	44	2.23 (10.828)	45	0.24 (0.645)
District nurse	52	1.73 (8.047)	51	0.29 (1.501)	44	0.59 (2.182)	45	0.24 (1.495)
Social worker: home	52	0.08 (0.436)	51	0.00 (0.000)	44	0.45 (2.715)	45	0.02 (0.149)
Social worker: office	52	0.02 (0.139)	51	0.00 (0.000)	44	0.02 (0.151)	45	0.02 (0.149)
Care worker: home	52	0.00 (0.000)	51	0.18 (1.260)	44	0.09 (0.603)	45	0.89 (5.373)
Care worker: office	52	0.21 (1.177)	51	0.00 (0.000)	44	1.00 (6.633)	45	1.00 (5.502)
Prescription	52	8.04 (37.996)	51	1.02 (2.596)	44	0.77 (2.371)	45	0.89 (2.665)

A, prednisolone; A&E, accident and emergency; B, PTX; O, placebo; SD, standard deviation.
a *n* is the number of participants in each arm who responded to the questionnaire.

TABLE 35 Utility values at discharge, 90 days and 1 year for survivors only

EQ-5D results during the trial	*n*[a]	Discharge (SD)	*n*	90 days (SD)	*n*	1 year (SD)
OO	143	0.654 (0.316)	103	0.582 (0.371)	46	0.673 (0.306)
AO	147	0.615 (0.329)	100	0.545 (0.362)	48	0.566 (0.381)
OB	119	0.616 (0.347)	83	0.604 (0.326)	36	0.477 (0.376)
AB	128	0.635 (0.332)	91	0.561 (0.353)	40	0.604 (0.324)

A, prednisolone; B, PTX; O, placebo; SD, standard deviation.
a *n* is number of participants in each treatment arm.

decrease between the 90-day visit and 1-year visit. Imputations were made for missing data in the economic analysis at 1 year but they are not represented in *Table 35*. The assumptions made in the economic analysis at 1 year were explored in the sensitivity analysis and a multiple imputation method was adopted. For participants who partially completed the questionnaire we used the 'worst-case scenario' method of imputation. This assumption was explored and had little to no effect on the average utility scores for each treatment arms owing to the low number of participants it affected.

Primary analysis

The primary analysis conducted for the health economics focused on the costs and outcomes at the end of the treatment phase (day 28). The costs included in the 28-day analysis were the intervention costs, length of initial admission (including time spent in ICU), nutritional supplements, discharge procedures and adverse event investigations. Discharge procedures were considered only if a participant was discharged within the treatment phase. If a participant was discharged after the treatment phase, the length of their initial admission was capped at 28 days.

For participants whose discharge form included some information, but for whom complete data were not available, we assumed that what was recorded represented all procedures performed. If a participant was in hospital on any of the days during the treatment phase but had no laboratory test information, we assumed that the laboratory tests that were specified in the protocol were performed during and after the treatment phase. These laboratory tests included LFTs (AST, ALT, albumin, ALP, bilirubin and total protein), creatinine, urea, PT and FBC (haemoglobin, WBC, neutrophils and platelets). Other laboratory tests (e.g. glucose) were available and performed on other participants during the 28-day treatment phase, however, these additional tests were not specified in the trial protocol and hence not included in our imputation. This assumption was explored in the sensitivity analysis in our model-based analysis. This could result in a slight underestimation of laboratory costs for all treatment groups. The investigations performed as part of an adverse event were recorded in the adverse event form. We identified the time at which the investigation was performed based on each patient's start date and the date of the investigation. Only investigations that were performed during the 28-day treatment phase were included in this analysis.

On average, standard treatment was the most costly treatment over 28 days. This was because of the longer initial admission and more frequent ICU admissions, on average, compared with the other treatment arms. The total cost on average for standard care was just under £4900 over the 28-day treatment phase. PTX alone was the only treatment to have a higher probability of death at 28 days than standard care and was, hence, on average, more costly and less effective (i.e. dominated) by prednisolone only and dual treatment. Both prednisolone only and dual treatment (prednisolone and PTX) dominated standard care and PTX only. At 28 days, prednisolone had on average a 0.023 higher probability of survival than standard care, while the dual treatment arm had an average 0.032 higher probability of survival (*Table 36*).

TABLE 36 A 2 × 2 factorial design: within-the-table cost-effectiveness analysis

Parameter	OO ($n = 272$)	AO ($n = 274$)	OB ($n = 273$)	AB ($n = 272$)
Mean costs, £ (SD)	4869 (8131)	3618 (4052)	4194 (2810)	3827 (2711)
Incremental cost/patient vs. OO, £	N/A	–1251	–675	–1042
Patients in each arm who died during the treatment phase, *n* (%)	45 (16.7)	38 (14.3)	50 (19.4)	35 (13.5)
Probability of mortality at 28 days (95% CI)	0.167 (0.1287 to 0.2169)	0.143 (0.1058 to 0.1878)	0.194 (0.1355 to 0.2344)	0.135 (0.0882 to 0.1654)
Probability of survival at 28 days	0.833	0.856	0.806	0.865
Incremental survival vs. OO	N/A	0.023	–0.027	0.032
Incremental cost/survival vs. OO	N/A	Dominant	Dominated	Dominant

A, prednisolone; B, PTX; N/A, not applicable; O, placebo.
Patients whose status at 28 days was 'unknown' were excluded from this analysis.

An incremental analysis was performed across all treatment arms. If one treatment was not dominated by another treatment, then we calculated the ICER to compare these treatments. The least costly option was prednisolone. Only dual treatment was more effective than prednisolone alone but this was more costly and the incremental cost per additional survivor was over £26,125 (*Table 37*). This number is difficult to interpret but with other things remaining equal, every additional survivor at 28 days would need to gain almost 0.77 QALYs each over their remaining lifetime for the incremental cost per QALY to be £30,000 [note: PTX and prednisolone (dual treatment) has a 0.009 higher probability of an additional survivor at day 28]. Assuming the difference in cost remains the same then assuming the maximum willingness to pay per QALY is £30,000, then the QALY gain for each additional survivor would need to be 0.77 QALYs (209/30,000/0.009 = 0.77).

Figure 10 was produced by bootstrapping estimates of the mean costs and mean probability of dying across the four treatment arms. It is an illustrative example of the changes in the probability of each treatment being cost-effective at a range of willingness-to-pay values. It is analogous to a one-sided CI in that if it is judged that society is willing to pay an upper limit for the cost per survivor, then society will certainly pay a lower value.

TABLE 37 Incremental cost per additional survivor at 28 days

Intervention	Cost (£)	Survival	Incremental cost per additional survivor (£)	Probability that intervention is cost-effective for different threshold values for society's WTP for an additional survivor		
				£0	£5000	£10,000
AO	3618	0.857	–	0.79	0.79	0.60
AB	3827	0.865	26,125	0.20	0.13	0.08
OB	4194	0.806	Dominated by AO and AB	0.01	0.08	0.30
OO	4869	0.833	Dominated by AO and AB	0.00	0.00	0.02

A, prednisolone; B, PTX; O, placebo; SD, standard deviation; WTP, willingness to pay.
Assumes no interaction between PTX and prednisolone.

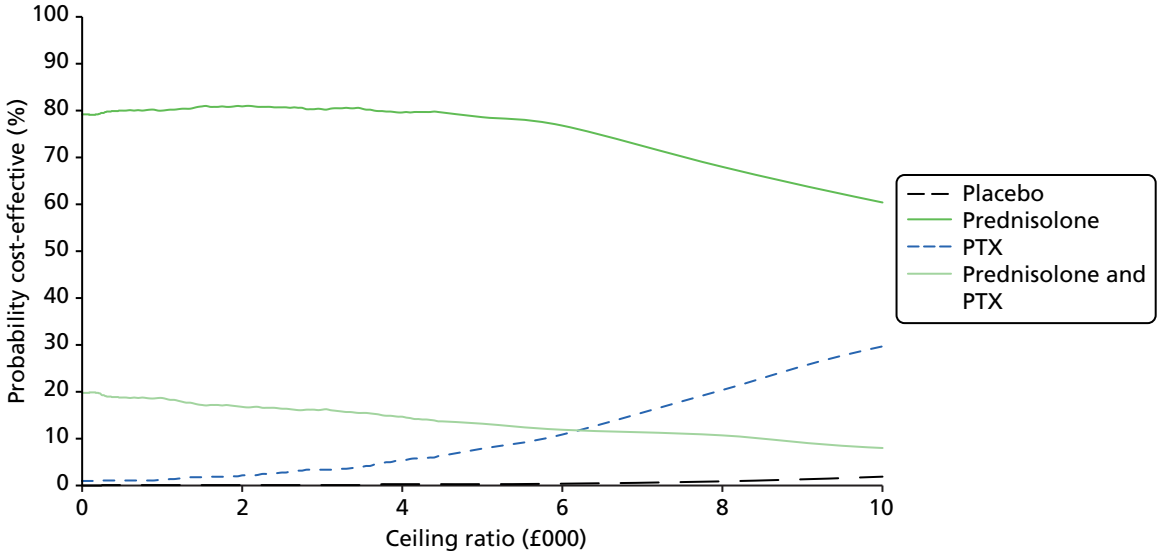

FIGURE 10 Cost-effectiveness acceptability curve: placebo vs. prednisolone vs. PTX vs. prednisolone and PTX.

At-the-margins analysis

The effects of both active treatments, prednisolone and PTX, were compared individually in two subsequent analyses [(prednisolone and placebo plus prednisolone and PTX vs. placebo and PTX plus placebo and placebo) and (prednisolone and PTX plus placebo and PTX vs. placebo and placebo plus prednisolone and placebo)]. Since prednisolone appeared to be the most cost-effective treatment at 28 days when compared with the other three treatments, our first analysis determined the cost-effectiveness of prednisolone vs. no prednisolone (see *Table 38*).

Table 38 shows the results of the at-the-margins analyses comparing prednisolone vs. no prednisolone. These results support the findings in the four-arm comparison, as prednisolone dominates no prednisolone (prednisolone and placebo plus prednisolone and PTX vs. placebo and PTX plus placebo and placebo). The results of this analysis were bootstrapped to show that the statistical imprecision surrounding the estimates of incremental cost and incremental survivors (see *Figure 11*) and probability of prednisolone being cost-effective compared with no prednisolone (see *Figure 12*). In *Figure 11*, caution needs to be taken when interpreting our results because we are looking at the incremental difference in mortality between both treatment arms. The majority of iterations produced from the bootstrapping analysis are in the south-west quadrant and this clearly illustrates that prednisolone is dominant because it has a lower cost on average and a lower probability of mortality compared with no prednisolone. *Figure 12* highlights that as society's willingness to pay for an additional survivor increases, the likelihood that prednisolone compared with no prednisolone would be considered cost-effective decreases, but it still remains the most likely to be considered cost-effective.

TABLE 38 Cost-effectiveness analysis: prednisolone vs. no prednisolone

Intervention	Cost, £ (SD)	Survival	Incremental cost per additional survivor	Probability that intervention is cost-effective for different threshold values for society's WTP for an additional survivor		
				£0	£5000	£10,000
Prednisolone ($n = 546$)	3722 (3448)	0.861	–	1.00	0.98	0.81
No prednisolone ($n = 545$)	4531 (6082)	0.800	Dominated	0.00	0.02	0.19

A, prednisolone; B, PTX; O, placebo; SD, standard deviation; WTP, willingness to pay.
Assumes no interaction between PTX and prednisolone.

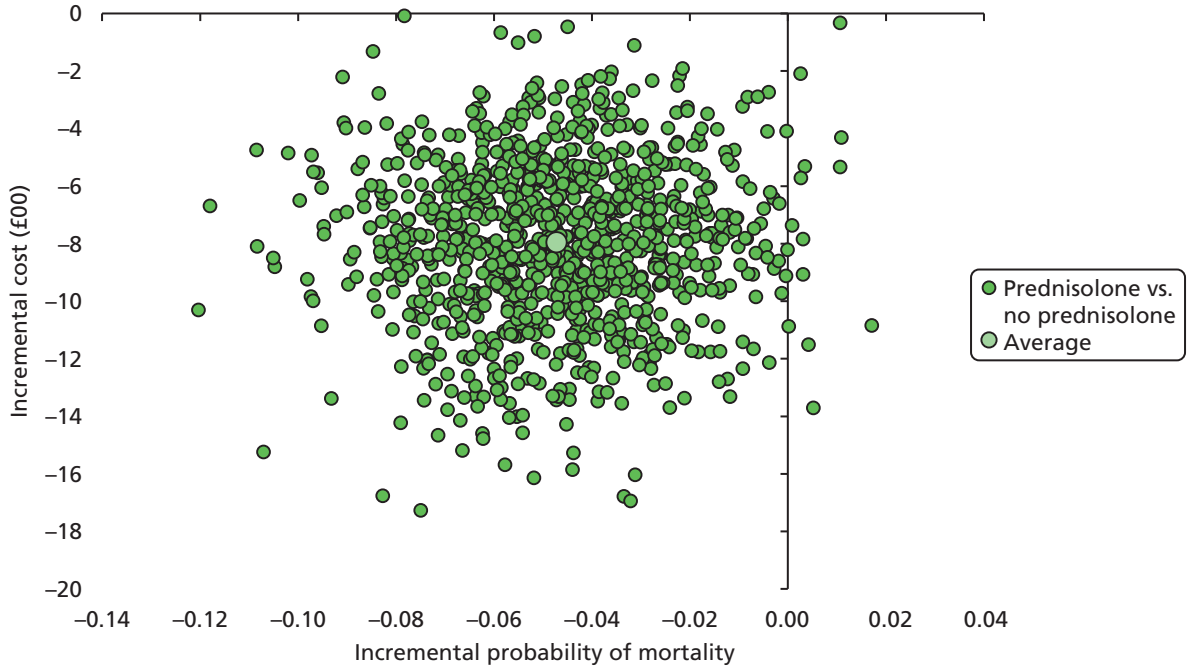

FIGURE 11 Cost-effectiveness scatterplot: prednisolone vs. no prednisolone at 28 days.

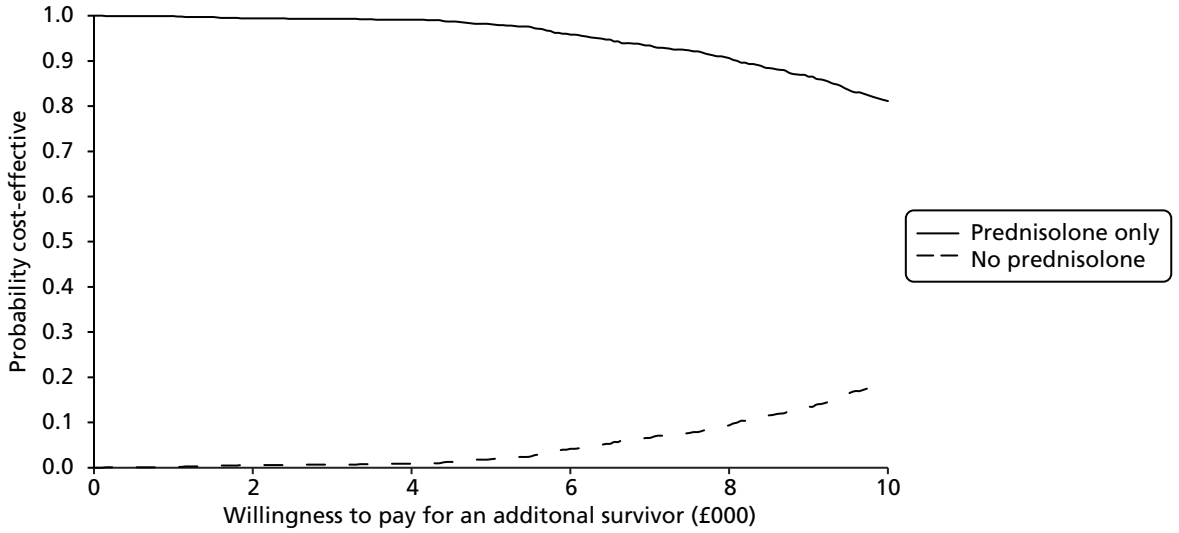

FIGURE 12 Cost-effectiveness acceptability curve: prednisolone vs. no prednisolone at 28 days.

The next at-the-margins analysis we conducted looked at PTX vs. no PTX. The incremental and bootstrapped results are presented in *Table 39* and *Figures 13* and *14*.

The PTX arm was slightly less effective but less costly than the no PTX arm. The incremental cost per additional survivor for no PTX compared with PTX was in excess of £25,000 (see *Table 39*). The plots of bootstrapped mean costs and differences in survival are shown in *Figure 13*. This figure illustrates the statistical imprecision surrounding estimates of survival but that it is highly likely that PTX is less costly than no PTX. The situation of the average iterations in the south-east quadrant suggests that despite the cost savings generated by PTX it has a very slightly higher probability of mortality compared with no PTX. The cost-effectiveness acceptability curve is illustrated in *Figure 14*. This figure illustrates that there is approximately a 75% chance that PTX would be considered cost-effective over the range of threshold values for society's willingness to pay for an additional survivor.

TABLE 39 Cost-effectiveness analysis: PTX vs. no PTX

Intervention	Cost, £ (SD)	Survival	Incremental cost per additional survivor, £	Probability that intervention is cost-effective for different threshold values for society's WTP for an additional survivor		
				£0	£5000	£10,000
PTX ($n = 545$)	4012 (2765)	0.836	–	0.76	0.77	0.74
No PTX ($n = 546$)	4241 (6441)	0.845	25,444	0.24	0.23	0.26

A, prednisolone; B, PTX; O, placebo; SD, standard deviation; WTP, willingness to pay.
Assumes no interaction between PTX and prednisolone.

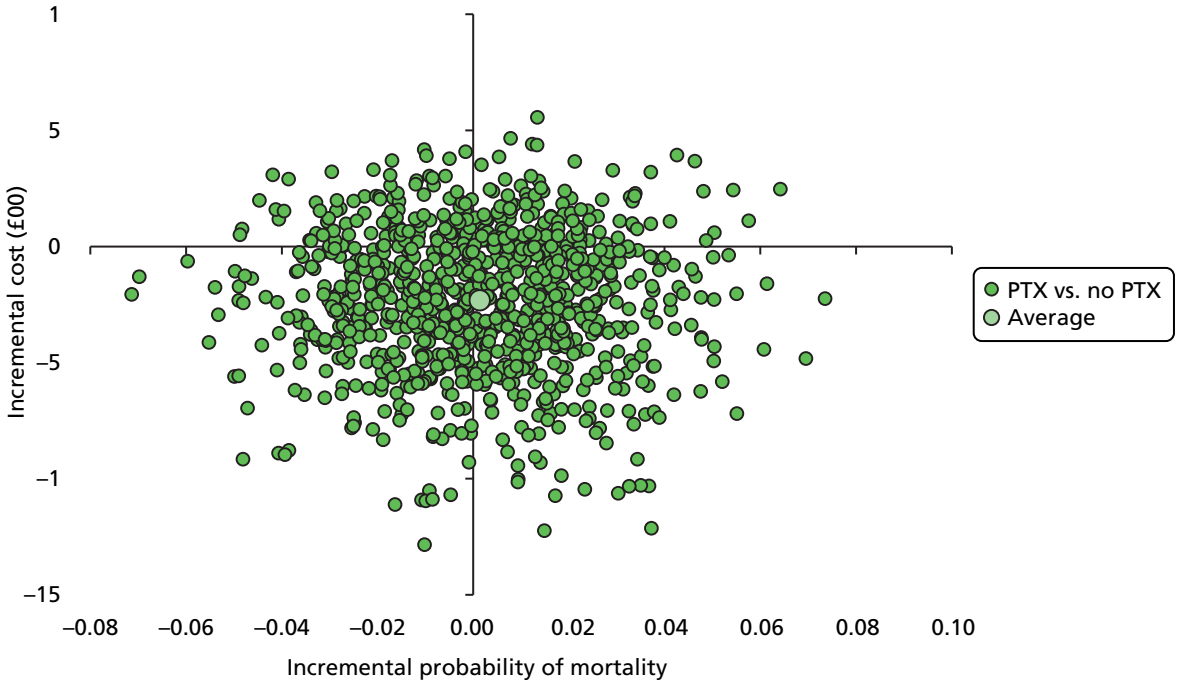

FIGURE 13 Cost-effectiveness scatterplot: PTX vs. no PTX at 28 days.

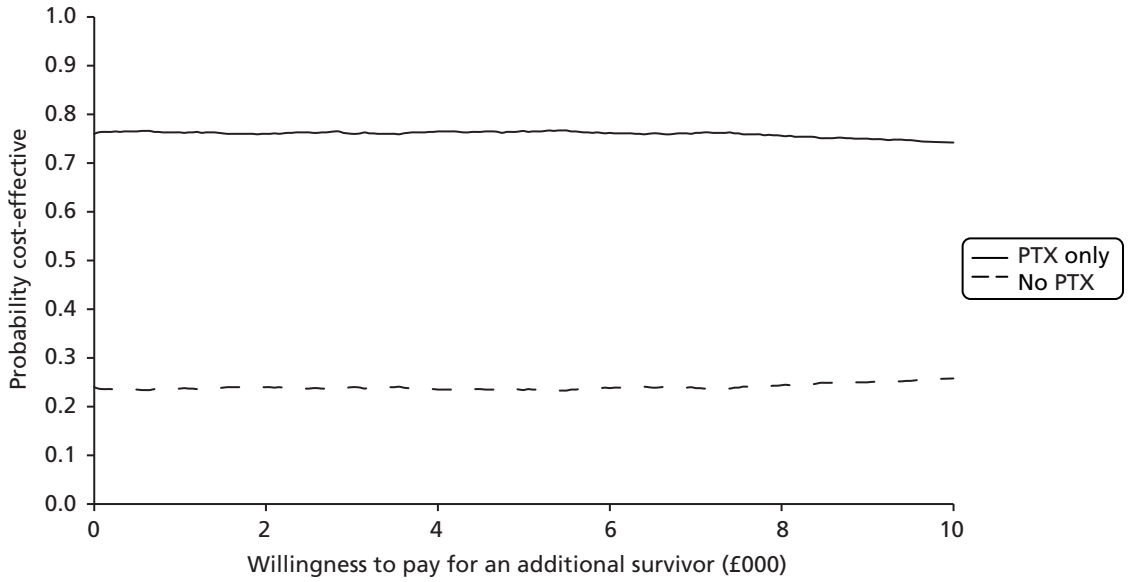

FIGURE 14 Cost-effectiveness acceptability curve: PTX vs. no PTX at 28 days.

Sensitivity analysis

In conjunction with our stochastic sensitivity analyses we also conducted other sensitivity analyses to reduce uncertainty from our base-case analysis. In our initial analysis we could identify a small number of patients in each treatment arm who were having a high impact on the total average cost in each treatment arm. We removed the most costly 10% of patients in each treatment arm and reran our base-case analysis. Unsurprisingly the average total cost across treatment arms was lower but overall differences in costs remained similar across arms, as there was an even spread of high-cost patients. One key difference to note is that 14% of high-cost patients who were removed from the analysis died during the treatment phase. This resulted in prednisolone being the dominant treatment compared with every other treatment option. *Table 40* highlights the results from the sensitivity analysis.

In our probability of survival we excluded patients whose status at 28 days was unknown. If we assumed that they were still alive at 28 days, prednisolone only is still the favourable treatment option. Dual treatment has a slightly higher probability of survival in this case but the ICER comparing dual treatment to prednisolone alone suggests that the cost per additional survivor is in excess of £100,000 [355/(0.1463 – 0.1434) = 355/0.003 = £11,8333].

Our sensitivity analyses were conducted for the within-the-table analysis described above. We have previously discussed how we estimated discharge procedures and adverse event investigations. Some of the unit costs generated did not appear reflective of the procedure performed (e.g. chest drain).

TABLE 40 Cost-effectiveness sensitivity analysis with the most costly 10% of patients removed from each arm

Parameter	AO ($n = 246$)	AB ($n = 244$)	OB ($n = 245$)	OO ($n = 245$)
Mean costs, £ (SD)	2898 (1789)	3253 (2049)	3624 (2265)	3924 (2299)
Incremental cost/patient vs. AO	N/A	355	726	1026
Probability of survival at 28 days[a]	0.890	0.849	0.818	0.791
Incremental survival vs. AO	N/A	–0.041	–0.072	–0.099
Incremental cost/survival vs. AO	Dominant	Dominated	Dominated	Dominated

A, prednisolone; B, PTX; N/A, not applicable; O, placebo; SD, standard deviation.
a Patients whose status at 28 days was 'unknown' were excluded from this analysis.

We conducted extreme sensitivity analyses around the costs of most concern. We initially removed these costs completely from our analysis and then we doubled the cost of these procedures in our analysis. Both extreme analyses did not affect the overall results from our primary analysis; prednisolone was still the most cost-effective treatment option.

Model-based analysis

Our model was used to estimate the cost-effectiveness of the treatment arms over a 1-year period and over a patient's lifetime. The following section presents details on the methods we adopted, our results and our sensitivity analysis.

Methods

As previously mentioned (see *Methods overview*) we used a Markov model to estimate the relative cost-effectiveness of each of the four treatment arms over 1 year and over the patient's lifetime time horizon. A daily cycle length was adopted as this patient group has a high mortality rate. The parameters used in the model were: cost of the treatment medications as our initial cost for each treatment arm; average total cost per day, if alive, of management; average utility per day if alive; and probability of mortality per day. By definition, both daily cost and utility were taken as 0 if the individual was dead.

The daily management cost per treatment arm incorporated: initial admission, ICU admission, discharge procedures, adverse events, liver biopsy, alcohol counselling, laboratory test, renal failure support, liver transplant, nutritional support and resource use at 90 days. The methods to determine these costs are described in *Methods*. After we estimated the total average resource use cost per treatment arm over 90 days we estimated the daily total average cost per treatment arm by dividing this figure by 90.

We then estimated the average utility per day by treatment arm for people alive on a given day. The utility score for someone who had died was taken as 0. To calculate the utility score by treatment arm per day we used the data reported in *Table 35* and assumed that the baseline utility was –0.402. This value was chosen to represent the health state of patients with AH when they are admitted to hospital and enter the study. However, if a patient was discharged within 10 days we used the discharge utility score as the baseline score; this was to prevent us overestimating the treatment effect. This assumption was explored in a sensitivity analysis.

To estimate the utility score assigned to each day we first estimated the quality-adjusted life-days (QALDs) over the first 90 days of follow-up. This was accomplished using the trapezoid method.

$$\begin{aligned}\text{QALDs} = (&\textbf{\textit{EQ5D}}\textit{baseline} \times \text{days to discharge}) + \{[(\textbf{\textit{EQ5D}}\textit{discharge} - \textbf{\textit{EQ5D}}\textit{baseline})/2] \\ &\times \text{days to discharge}\} + [(\textbf{\textit{EQ5D}}\textit{discharge} \times (90 - \text{days to discharge})] \\ &+ \{[(\text{EQ5D90 day} - \textbf{\textit{EQ5D}}\textit{discharge})/2] \times (90 - \text{days to discharge})]\}. \end{aligned} \quad (3)$$

To give an average utility score for the 90 days of follow-up, QALDs were then divided by 90. This value was used in the economic model as the utility score assigned for each day spent alive.

A small number of participants (< 1%) had their discharge utility score collected after 90 days, which was excluded from our analysis. This allowed us to maintain consistency with the assumptions made when estimating costs over 90 days. If a participant had a later discharge (i.e. later than 90 days) we assumed that this was 'missing' and used the multiple imputation method to estimate their utility score. We also reduced their length of discharge to 89 days for our QALD calculation so there was a difference of 1 day and preventing their utility score being multiplied by 0.

The probability of mortality on a daily cycle was estimated using the patients' status at 1 year. We used information only on the participants that we had information for at 1 year. If the participant's status was unknown (i.e. late randomisation or lost to follow-up), he/she was excluded from this calculation. We estimated the probability of mortality at 1 year on those who died, divided by the total number of

patients with their status available at 1 year. This probability was converted into a rate, and this rate was converted into a daily probability of mortality for each treatment arm using the following method: $[(1 + \text{rate})^{(1/365)+1}]$. We assumed the patients without a 1-year status would have the same probability of dying. We also included the annual probability of all-cause mortality and this was converted into a daily rate using the same methods as described above.

Figure 15 is an illustration of the model structure used to estimate 1-year results and extrapolate the trial results, and *Table 41* reports the parameters used in our model.

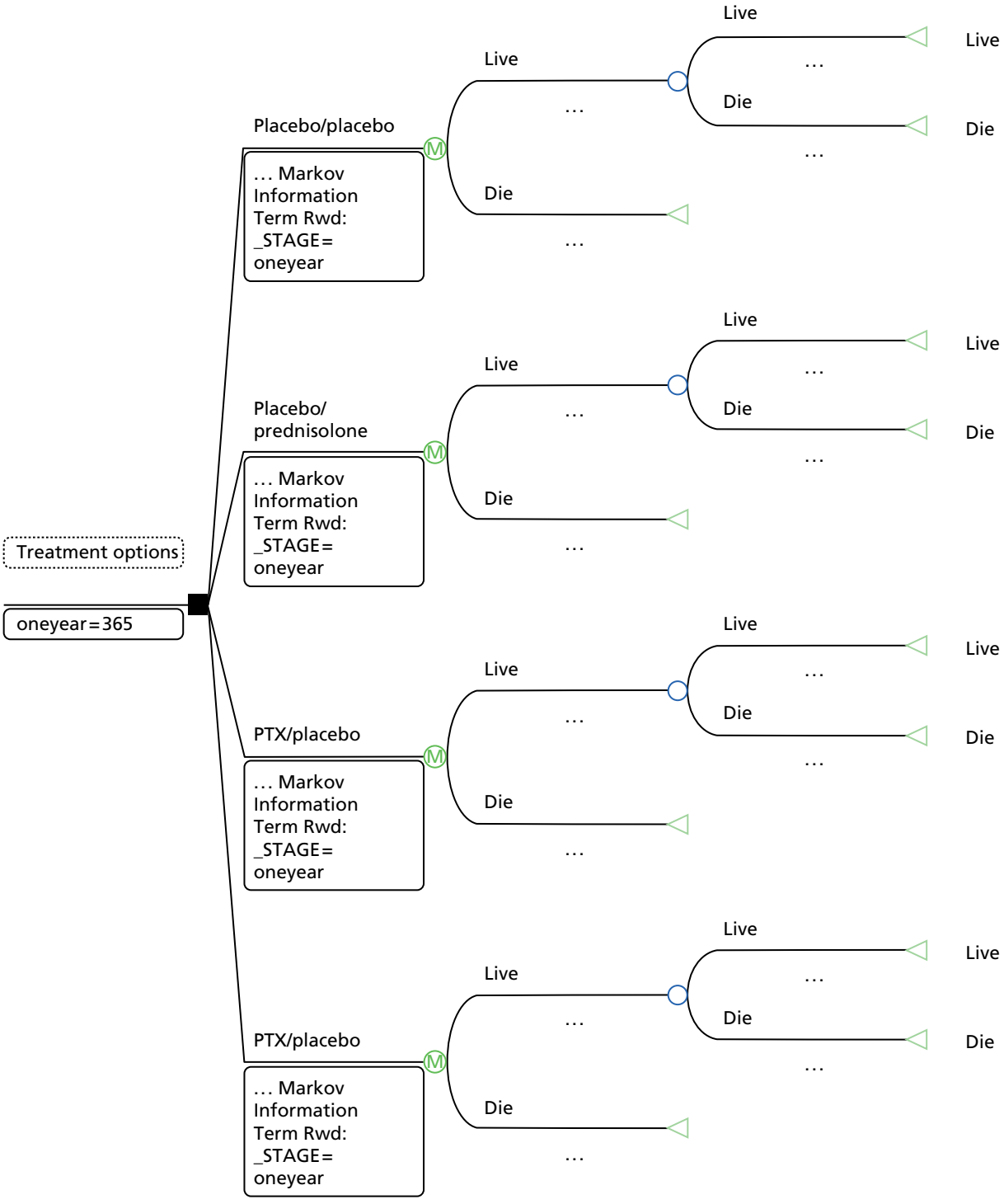

FIGURE 15 Economic model.

TABLE 41 Parameters used in the economic model

Parameters	OO	AO	OB	AB
Daily probability of dying from AH	0.00162	0.00171	0.00168	0.00162
Daily probability of dying from all causes[a]	0.000007–0.000961	0.000007–0.000961	0.000007–0.000961	0.000007–0.000961
Initial cost (£)	0.00	10.48	18.10	28.58
Incremental cost per day spent alive, £ (SD)	95 (133)	80 (91)	78 (67)	80 (91)
Initial utility score	–0.402	–0.402	–0.402	–0.402
Incremental utility score for each day alive (SD)	0.3476 (0.3089)	0.3551 (0.2980)	0.2700 (0.3235)	0.3477 (0.2969)
Distribution form for costs	Gamma	Gamma	Gamma	Gamma
Distributions for cost/day – alpha	0.51020	0.77285	1.35531	1.46924
Distributions for cost/day – lambda	0.00537	0.00966	0.01737	0.01836
Distribution form for utilities	Beta	Beta	Beta	Beta
Distributions for utilities – alpha	0.47851	0.56062	0.23851	0.54692
Distributions for utilities – beta	0.89810	1.01814	0.64487	1.02604
Distribution form for daily probability of dying from AH	Beta	Beta	Beta	Beta
Distribution for daily probability of dying from AH – alpha	0.44064	0.46854	0.45864	0.44064
Distribution for daily probability of dying from AH – beta	272	274	273	272

A, prednisolone; B, PTX; O, placebo.
a The range of the daily probability of all-cause mortality is estimated based on the assumption that the cohort has a starting age of 48 years and the daily rate changes on an annual basis until the cohort reach the age of 100 years or die.

The initial cost for each treatment arm is the cost of the initial medications. The distributions for incremental costs were estimated using a gamma distribution using the mean and variance to estimate the variables needed to define the gamma distribution. The incremental utility distributions were estimated using a beta distribution (again the mean and variance were used to estimate the parameters needed to define the shape of the beta distribution). The distribution of the daily probability of death owing to AH was estimated based on the number of people dying in each arm on a daily basis. Of note in this analysis is that the choice of parameter values is conservative for prednisolone alone with respect to mortality in particular.

Sensitivity analysis

For the model-based analysis, probabilistic sensitivity analysis (PSA) was conducted on our modelled results using the Monte Carlo simulation. This allowed us to vary all of our parameters simultaneously to determine what effect this had on the overprobability of one treatment being cost-effective relative to the others. Distributions for each model parameter, along with the information to define those distributions,

are presented in *Table 41*. The results of this analysis are presented in a similar fashion to those from the within-trial bootstrapped analysis. For the model-based analysis for up to 1-year follow-up, discounting was not to be carried out.

Deterministic sensitivity analysis was carried out to test for the effect of our assumptions and variability, such as an exploration of alternative unit costs applied to the different resources used and the number of visits a participant had with a counsellor. We chose to conduct the deterministic sensitivity analysis as part of our model-based analysis because, in addition to the assumptions made in our base-case analysis, we also made assumptions with regard to costs estimated over the 90-day period. By conducting this analysis as part of the model analysis we can capture all of these assumptions and make amendments in one analysis.

Modelling results

The Markov model was run for 365 cycles in the first instance to present the results for the 1-year analysis. In this analysis, QALDs, as predicted by the model, have been converted into QALYs by dividing by 365.

The least costly and least effective option is PTX alone (*Table 42*). The next least costly option is prednisolone alone, which is associated with an incremental cost per QALY gained of £6924 compared with PTX alone. No treatment and PTX treatment are, on average, less effective and more costly than combination treatment (PTX and prednisolone) and unlikely to be cost-effectiveness compared with prednisolone alone.

Table 43 reports the undiscounted incremental cost-effectiveness over a 10-year horizon (which, given mortality rates, is equivalent to a lifetime time horizon). In this analysis, PTX alone is the least costly and least effective intervention, which is consistent with the lower daily cost assumed for this intervention. Dual treatment is dominated by prednisolone only as it is more costly on average and less effective. Prednisolone only has an ICER of £2867 so would be considered cost-effective compared with PTX only. No treatment is unlikely to be considered worthwhile. This analysis represents an extreme sensitivity

TABLE 42 Analysis at 1 year comparing the four possible treatment arms

Intervention	Cost, £	Incremental cost, £	QALYs	Incremental QALYs	Incremental cost per QALY gained (ICER), £
OB	21,223	–	0.200	–	–
AO	21,653	430	0.2621	0.0621	6924
AB	21,992	339	0.2604	–0.0017	Dominated by AO
OO	26,082	4429	0.2604	0.00	Dominated by AO

A, prednisolone; B, PTX; O, placebo.

TABLE 43 Lifetime (10-year) Markov results: undiscounted

Intervention	Cost, £	Incremental cost, £	QALYs	Incremental QALYs	Incremental cost per QALY gained, £ (ICER)
OB	46,136	–	0.4375	–	–
AO	46,503	367	0.5655	0.128	2867
OO	58,232	11729	0.5837	0.0182	644,451
1AB	46,644	–11588	0.5554	–0.0283	Dominated by AO

A, prednisolone; B, PTX; O, placebo.

analysis representing a scenario when prednisolone alone might be cost-effective. *Table 44* reports a similar analysis but this time both costs and QALYs are discounted at the recommended 3.5% rate. The findings of this analysis are similar. In this case, standard treatment is dominated by dual treatment. This is because the standard treatment group have a marginally higher probability of survival but when the additional QALYs gained over a longer period of time are discounted, the average QALY gain for that treatment arm is lower than prednisolone only. Despite dual treatment not being dominated it would be highly unlikely to be considered cost-effective when compared with prednisolone only. What these two analyses illustrate is that relatively small differences in mortality, should they exist, could change conclusions about cost-effectiveness when a lifetime time horizon has been taken. Overall, prednisolone only appears to be the most favourable treatment option but caution needs to be taken interpreting these results.

Sensitivity analysis

As noted in the previous paragraph a PSA was conducted on our model parameters. The results of this analysis at both 1 year and a lifetime were very similar in that the likelihood of any one treatment being considered cost-effective was very similar (*Figures 16* and *17*). This reflects the considerable imprecision surrounding estimates used within the model and again represents a cautious interpretation of the evidence available. Taken together, the results of the deterministic and probabilistic economic evaluation would suggest that treatment with prednisolone alone is a promising treatment but there is insufficient economic evidence to be conclusive.

TABLE 44 Lifetime (10-year) Markov results: discounted

Intervention	Cost, £	Incremental cost, £	QALYs	Incremental effect	Incremental cost per QALY gained, £ (ICER)
OB	42,899	–	0.4068	–	–
AO	43,275	376	0.5263	0.1195	3146
AB	45,517	2242	0.5420	0.0157	142,803
OO	54,052	8535	0.5418	–0.0002	Dominated by AB

A, prednisolone; B, PTX; O, placebo.

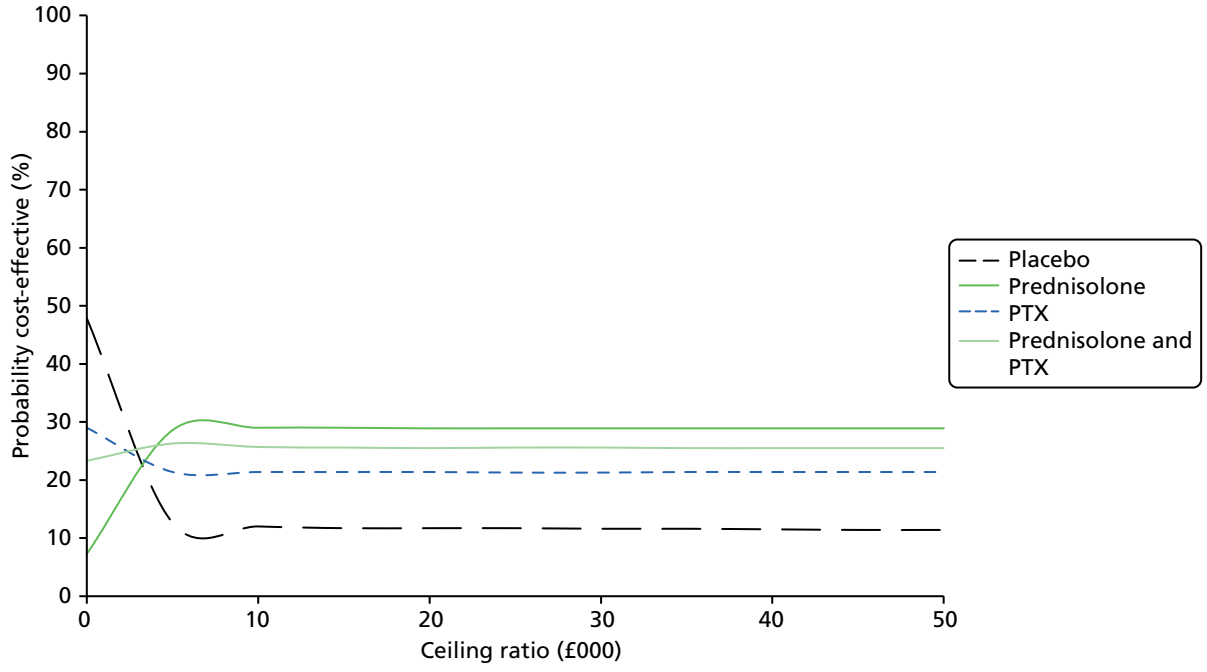

FIGURE 16 One-year cycle PSA results.

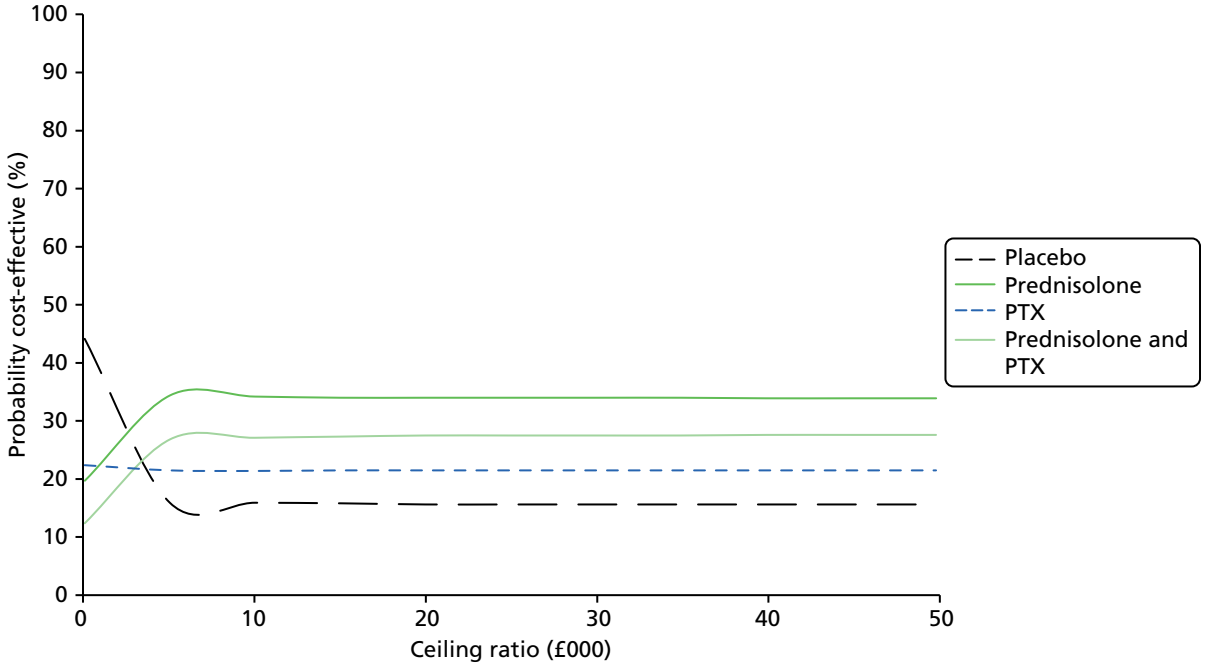

FIGURE 17 Lifetime (10-year) cycle PSA results.

We conducted a deterministic sensitivity analysis around some of the assumptions previously mentioned in the chapter that were used to estimate the average daily cost for each treatment option. In this analysis we assumed that all additional hospital admissions outside of the initial treatment phase (28 days) were in ICU. We used multiple imputations to estimate all missing utility values at discharge and 90 days. *Table 45* presents the updated parameters that were used in our model and the results at 1 year after these changes were made. The key cost driver in this analysis was the cost of ICU admissions. This cost alone undermined any of the cost savings made by changing all of our other assumptions. In this analysis, despite producing similar results to our 1-year analysis, PTX alone would not be considered cost-effective because of the higher ICER (£85,427) associated with it, owing to the increase in costs. If patients with AH are treated in an ICU setting compared with a general ward, consideration needs to be taken on the effect this will have on overall cost-effectiveness of prednisolone.

TABLE 45 Deterministic sensitivity analysis with 1-year Markov results

Parameter	OB	AO	AB	OO
Average daily cost, £ (SD)	349 (308)	360 (344)	373 (345)	407 (372)
Average daily utility score (SD)	0.2911 (0.2925)	0.3283 (0.2830)	0.3278 (0.2713)	0.3177 (0.2991)
Total average costs at 1 year, £	94,897	97,400	102,435	111,741
Incremental cost difference at 1 year, £	N/A	2503	5034	9306
Total average QALY score at 1 year	0.2129	0.2422	0.2455	0.2379
Incremental QALY difference at 1 year	N/A	0.0293	0.0033	–0.0077
Incremental cost/QALY at 1 year, £	–	85,427	1,525,455	Dominated by AO and AB

A, prednisolone; B, PTX; N/A, not applicable; O, placebo.

Summary

This chapter has reported both a preference elicitation exercise based on a SG experiment and an economic evaluation. The aim was to estimate the QoL associated with the three Child–Pugh classifications (A, B and C). Given the severity of the disease, relatively fewer people with Child–Pugh C were able to complete the SG. The results of the analysis were in line with expectations in that higher utilities were recorded for those in less severe health states compared with more severe states. Convergent validity was tested by examining valuations of an individual's own health today, as measured by the SG and EQ-5D. These values were not significantly different, suggesting high levels of convergent validity between the two measures values. These results hold for the full sample analysis and subsample analyses.

With respect to the economic evaluation, both within-trial and model-based economic evaluations were conducted. The economic evaluation itself was complicated both by the nature and extent of the data that needed to be considered, and more specifically the truncated follow-up of a substantial proportion of the trial participants. The consequence of this truncated follow-up was that several of the pre-planned analyses were not sensible to conduct given the very large quantities of missing data that existed. Nevertheless, both a cost-effectiveness and a cost–utility analysis were conducted. The former was conducted as part of a within-trial-based analysis and the latter as part of an economic model. It should be noted that because of decisions made about the choice of parameter values used in the model, the model-based evaluation is highly conservative against the prednisolone alone treatment strategy. Overall, the results of the within-trial analysis suggest that prednisolone might be considered cost-effective compared with the other treatment options. The deterministic model-based cost–utility analysis results at a 1-year time horizon support this. However, under conservative assumptions these results did not hold over the estimated lifetime of patients in the PSA. These and the results of an accompanying PSA suggest that there is considerable uncertainty remaining over long-term cost-effectiveness and that, while apparently promising, treatment conclusions about the use of prednisolone alone in the longer term remain inconclusive.

Chapter 6 Discussion and conclusions

Summary of findings

Therapeutic effects

Although the trial results show that prednisolone at a dose of 40 mg daily for 28 days reduces the mortality at 1 month of severe AH by approximately 30% (odds ratio 0.72, 95% CI 0.52 to 1.01; $p = 0.056$), this fails to reach the conventional threshold for statistical significance. In multivariate analyses, which allow small variations in baseline variables to be taken into account, the beneficial effects of prednisolone are much clearer (odds ratio 0.609, 95% CI 0.409 to 0.909; $p = 0.015$). Although the sample size is not powered for a multivariate logistic regression analysis, the adjusted odds ratio provides a more precise estimate of the treatment effect. Taken together with previous studies, summarised in Mathurin's meta-analysis,[17] it is reasonable to conclude that prednisolone does have a therapeutic benefit up to 28 days. However, after 28 days the survival curves for prednisolone-treated and prednisolone-untreated patients converge, and at 90 days and 1 year there are no survival benefits to receiving steroids. Survival beyond 28 days may be influenced by adverse events related to steroid use (see *Adverse events*) or, alternatively, may be influenced by other factors and, in particular, recidivism. We were able to show a major effect of recidivism at 90 days or mortality at 1 year, but do not have the data to demonstrate an effect at earlier time points.

Pentoxifylline has no impact on the mortality at 28 days, 90 days or 1 year. In previous studies,[18] PTX has conferred survival benefit through protection of renal function because of acute kidney injury. In this trial the rate of acute kidney injury was low, so although there was numerically fewer acute kidney injuries in the PTX-treated patients this failed to reach statistical significance. The combination of prednisolone with PTX had no effect on patients' survival at 28 days, 90 days or 1 year. Without being able to demonstrate an effect on survival or acute kidney injury it is difficult to claim any therapeutic benefit for PTX.

Adverse events

Prednisolone, as expected, was associated with an increased risk of infection, with 13% of patients experiencing an infection-related SAE in the prednisolone group compared with 7% in the control group. Approximately 35% of infection-related SAEs resulted in mortality but this proportion did not differ between treatment groups. The higher risk of an infection-related SAE continued in the prednisolone group after the 28 days but by 90 days the risk was equivalent, irrespective of the treatment arm. Infection is therefore likely to have accounted for the excess deaths after 28 days in the prednisolone-treated group that caused the convergence of the mortality curves. However, it is difficult to be sure that infection is the key cause of death in this group of patients, when mortality is usually associated with multiorgan failure. Drawing on previous studies[12] as well as results from the STOPAH trial, it is possible to conclude that patients who fail to respond to steroids after 7 days, defined by a Lille score of > 0.45, may not benefit from further treatment and, therefore, steroids should be stopped to avoid further risk of infection. This is coherent with current European Association for Study of the Liver (EASL) guidelines.[53]

Although PTX has previously been reported to protect against hepatorenal syndrome or acute kidney injury in patients with AH, this benefit could not be confirmed in the STOPAH trial as only 1.5% of patients in the PTX-treated group and 2.5% in the controls experienced acute kidney injury.

Prognostic scores

The three prognostic scores based on baseline clinical and laboratory variables did not perform well for the prediction of mortality at 28 days, with deteriorating performance for mortality at 90 days and 1 year. None of the scores gave area under the ROC curve > 0.75. The scores do not inform physicians on the likely response to prednisolone and, therefore, how to select patients who would benefit from the intervention. At present, no scoring system assesses the risk of infection, which would be helpful in initial treatment selection.

The 2011 trial of early liver transplantation for patients with AH relied on the Lille score to identify prospective candidates.[54] However, the data in the STOPAH trial suggest that patients with an adverse Lille score (> 0.45) have an expected survival of around 50% compared with the 25% in the original study. This rate of survival without transplantation does not justify the use of scarce donor organs, meaning that a new system for selecting transplant candidates should be considered.

Health economics

The health economic component of this study comprised two elements: a preference elicitation exercise using a SG experiment and an economic evaluation.

The preference elicitation component measured QoL directly using the SG, and indirectly using the EQ-5D-3L. The aim was to collect information on the health state valuation of Child–Pugh classification A, B and C. The results of the SG showed good face validity, with the Child–Pugh classification A health state being preferred to Child–Pugh classification B health state, and the Child–Pugh classification B health state being preferred the Child–Pugh classification C. The analysis also suggested high levels of convergent validity with the EQ-5D-3L responses.

These data would be useful for future studies exploring the cost-effectiveness of other treatments for alcoholic cirrhosis.

With respect to the economic evaluation, both prednisolone and PTX are very low-cost medications but the costs of managing AH within the UK NHS are high. Furthermore, the trial addressed the possibility that one or other, or indeed both, treatments may help alleviate some of the considerable morbidity and early morbidity associated with the condition. Given the profile of the condition it was plausible, however, that a more effective treatment might not be cost-saving simply because those who survive longer may need considerable further costly care. The economic evaluation was expressly designed to explore these trade-offs.

At 28 days, the total average cost of treatment varied between £3618 per patient (prednisolone only) to £4869 per patient (placebo patients). On average, in a within-the-table analysis, which compares all four arms, prednisolone only was the least costly treatment and slightly less effective than dual treatment. However, prednisolone was most likely to be considered cost-effective over the range of values considered for society's willingness to pay for an extra survivor. Much stronger results in favour of prednisolone were found in the at-the-margins analysis, when prednisolone was compared with no prednisolone. PTX was unlikely to be considered cost-effective, except in the at-the-margins analysis comparing PTX vs. no PTX, when PTX was most likely to be considered cost-effective, primarily because the estimates of effects were slightly in its favour (see *Figure 14*).

In terms of the incremental cost per QALY, the results tended again to favour the use of prednisolone. However, in this case any conclusions would be tentative because the assumptions made in the model around survival worked slightly against prednisolone and also there were very limited data available up to 1 year.

Strengths and weaknesses of the study

The STOPAH trial had a number of strengths; primarily the size of the study, which gave the trial adequate power to detect even a relatively modest treatment effect. Furthermore, the speed at which the trial recruited ensured that there were no significant changes in clinical practice during the study period, which might have influenced background rates of mortality. Clinical practice guideline developers have a preference for trials conducted at multiple sites rather than those focused on single centres. In this respect, the STOPAH trial excelled, having recruited in 65 sites across the UK. This ensured that recruitment took place in many district general hospitals as well as in liver units and teaching hospitals. In the UK, severe AH is not routinely an indication for transfer to tertiary centres and it is, therefore, important to have assessed the impact of treatment in the type of environment where the condition is usually treated.

The STOPAH trial was deliberately designed as a pragmatic clinical trial with broad inclusion criteria. Other trials in this field[22,55] have required liver biopsy and histological confirmation of the diagnosis prior to inclusion; however, in the UK, the USA and much of Europe, liver biopsy is rarely performed for diagnostic confirmation and reserved only for diagnostic uncertainty. This strategy is vindicated by a number of studies[22,55–57] that show that clinicians can accurately judge the diagnosis by the recent use of alcohol and also applying of strict criteria on the duration of jaundice. A trial requiring liver biopsy could not have recruited this number of subjects, as many centres do not have access to the transjugular technique that would be required in this group of patients or would have considered the procedure to be unethical.

Although the trial design allowed for the recruitment of patients with GI bleeding, sepsis and renal failure, once these conditions had been stabilised, the number of patients with these conditions was relatively small. Nevertheless, the results show no differences in the mortality rates or in the response to treatment in these patients. This finding should be used to ensure these patients are not excluded from future studies.

Reporting on the rate at which patients completed at least 75% of prescribed medication was required in order to define a 'per-protocol' population. Unfortunately, this issue was not reported well and was probably impaired after patients were discharged from hospital. It is well recognised that this group of patients is poorly adherent with follow-up procedures and reporting.

The SG exercise represents one of the very largest studies of its kind conducted in this area to date. It was based on best practice methods and benefited from being conducted alongside the STOPAH trial. The original aim of the SG was to estimate health-state valuations for those with cirrhosis owing to alcoholic liver disease and cirrhosis owing to viral hepatitis. Unfortunately, only 50% of the targeted recruitment was reached for viral hepatitis, thus limiting the applicability of findings to this group. Furthermore, the majority of participants (58%) had the mildest form of liver cirrhosis and some of the participants with Child–Pugh classification C found the SG difficult to complete or had limited ability to give informed consent. Taken together, valuation of the Child–Pugh classification C health state is subject to some uncertainty, as there were relatively few people with Child–Pugh classification C able to provide valuations.

The economic evaluation greatly benefited from being embedded in a rigorously conducted randomised controlled trial in which considerable efforts were made to collect all relevant data. These data were combined within an economic evaluation that followed explicit and rigorous methodology. Nevertheless, the economic evaluation was not without its challenges. Among these was the decision to curtail patient follow-up in order to maximise patient recruitment and sample size for the primary trial end point. This decision required careful consideration of the relative importance of effectiveness, safety and economic to all stakeholders. For the economic evaluation, the consequence of this decision was that not all the pre-planned analyses were conducted. This is because longer-term data up to 1 year was not available. This limitation was overcome in part by performing the economic evaluation on costs and survival at 28 days and attempting to model the incremental cost per QALY over 1 year and a patient's lifetime. The results of all these analyses are supportive of the use of prednisolone, but because of the limited long-term data, from an economic perspective, conclusions can only be tentative. However, taking the totality of evidence from the trial, of which the economics is only one part, the findings suggest the use of prednisolone reduces short-term mortality and that a decision-maker needs to form a judgement whether or not the suggested evidence on cost-effectiveness is sufficient.

Comparison with results of other studies

Mortality rates in AH have dropped since 1990. One can speculate that two factors have resulted in this change: (1) patients with severe liver injury are being looked after more effectively now than when the placebo-controlled trials were conducted 30 years ago, and (2) the patients with severe AH are not as nutritionally compromised as those that were treated 30 years ago.

The mortality rates in the STOPAH trial appear to be substantially lower than those reported in earlier studies.[14,15,17,58,59] It should be recalled that the placebo-controlled trials in severe AH were mainly conducted between 1971 and 1990, and when the treatment of hepatic failure, renal failure and sepsis were substantially less effective than they are today. The overall mortality in the STOPAH trial was 16% at 28 days, with 17% mortality in the placebo group and 14% mortality in the prednisolone only group. This is entirely in keeping with other studies. Similar mortality at 28–30 days in steroid-treated patients were found in other studies. Park *et al.*[27] found a 12% mortality rate, Mathurin *et al.*[23] found around a 12% rate and Nguyen-Khac *et al.*[31] found a 15% rate. In recent times there have been few placebo-controlled studies. However, mortality in the Akriviadis *et al.*[18] study was around 20% and in Boetticher *et al.*[60] was 22%. The mortality is clearly different from some older studies such as Ramond *et al.*,[14] who found 38% mortality.

The power calculation for the STOPAH trial was based on an average mortality rate of 35% in placebo-treated control subjects, whereas the observed mortality rate in placebo-treated subjects was actually 16.7%. Inevitably this will mean that future studies should use STOPAH trial mortality figures for power calculations.

In Mathurin's meta-analysis of placebo-controlled trials in patients with a DF ≥ 32, the hazard ratio for survival in the prednisolone-treated group was 0.43 (95% CI 0.27 to 0.7) compared with controls.[17] However, in these studies the mortality rate in the placebo arms was 34.3% compared with the 16.7% in the STOPAH trial. Nevertheless it should be noted that when patients with outlying (absolute standardised residual > 2) laboratory values are filtered out from the STOPAH data set, the odds ratio for mortality in the prednisolone-treated group was 0.427 (95% CI 0.253 to 0.721; $p = 0.001$). It is therefore likely that the magnitude of the steroid effect is highly comparable.

In view of the lower mortality figures, we compared the highly objective laboratory results with those reported in previous trials. In the STOPAH trial, the levels of creatinine were slightly lower than those seen in the Nguyen-Khac *et al.*[31] study but PTs, bilirubin, MELD and DF scores were highly comparable.

The Lille score, which includes a component based on response to treatment, also performed poorly even when the analysis was restricted to those patients who had received prednisolone treatment. A Lille score of > 0.45 has been used to select patients who might be candidates for early liver transplantation.[54] However, in this study approximately 50% of patients with this adverse Lille score survived to 1 year, indicating that it would be a poor criterion for transplant selection.

The incidence of renal failure and infection observed in the STOPAH patients appeared to be lower than that reported in other studies.[61] This may be partially accounted for through a lack of comprehensive reporting. No other systematic issue has been identified to explain the lower incidence of these complications. However, it should be noted that the STOPAH trial was conducted in all types of hospital and therefore did not suffer from the potential bias of a tertiary referral centre population.

Implications for services and future research

Prednisolone is already indicated for the treatment of severe AH in the guidelines published by the American Association for the Study of Liver Diseases and EASL.[62] Based on the STOPAH trial results, prednisolone may now be considered as the standard of care in patients with AH and a DF ≥ 32; however, the temporary nature of the benefit needs to be emphasised and steps taken to minimise the risk of infection.

As one of the major drawbacks to steroids, highlighted in this study as well as previous trials,[63] is the increased susceptibility to potentially fatal sepsis, future research should focus on strategies to reduce this risk. In patients who do not respond to prednisolone (no fall in bilirubin after 1 week of treatment) steroids should be stopped as indicated currently in the EASL guidelines. However, the optimal duration of steroid treatment has not been adequately explored and consideration should be given to shorter courses such as 14 days.

A second strategy to reduce infection would be to use *N*-acetylcysteine (NAC) alongside steroids. This is based on the observation of lower rates of sepsis in the Nguyen-Khac *et al.* trial,[31] which compared steroids alone against steroids plus NAC. In the Nguyen-Khac *et al.* trial,[31] mortality at 28 days was 24% in the steroid only group compared with 8% in the steroid plus NAC-treated group and infection rates were 25% versus 12%, respectively. NAC appears to have a therapeutic benefit on immune function, which has not yet been explained.[31] An alternative strategy is to combine the use of steroids with granulocyte colony stimulating factor, which has previously been evaluated in a small recently published trial.[64]

The STOPAH trial results do not demonstrate any benefit from the use of PTX in severe AH. The low incidence of acute kidney injury in the STOPAH trial prevents a definitive conclusion being drawn on the ability of PTX to prevent hepatorenal syndrome. Nevertheless, any renal benefit must be considered as a secondary outcome compared with reduction of mortality.

Despite improvements in short-term (28-day) mortality in patients treated for severe AH, there is still a relentless loss of patients after the initial admission, owing to alcoholic liver disease. The majority of patients admitted with AH already have underlying liver cirrhosis and the observed mortality appears to relate to complications arising from this condition. It is already well recognised that abstinence from alcohol has a major impact on survival in patients with alcohol-related cirrhosis[59] and the data in the STOPAH trial confirm that abstinence at 90 days strongly influences survival at 1 year. Nevertheless, the rates of complete abstinence in this cohort are relatively low. Future studies should concentrate on strategies to promote abstinence post admission. Linking patients into consultant-led addiction services while in hospital has been recommended in the National Confidential Enquiry into Patient Outcomes and Death report on ALD but has not been widely implemented.[63] Research is required to evaluate whether or not a full implementation of the National Confidential Enquiry into Patient Outcomes and Death recommendations would improve the outcomes of patients with AH or other ALD.

The use of anticraving medication to promote abstinence is sparse, despite clinical trial evidence of benefit.[65] Increased commissioning of alcohol addiction services in acute hospitals providing care for patients with ALD will be required to maximise the opportunity for patients to maintain abstinence after an admission for severe AH. In addition, we recommend a trial of anticraving medication to evaluate the impact on mortality and hospital readmission rates.

The STOPAH trial results found that the existing prognostic scoring systems are weak. Clinical management including the selection of patients for liver transplantation require more informative prognostic systems. The use of metabonomic, genomic and cytokine profiling methods needs to be evaluated in this population of patients.

Acknowledgements

Funding

The authors acknowledge funding from the UK National Institute for Health Research Health Technology Assessment Board, the National Institute for Health Research Clinical Research Network for providing research nurse support and the Imperial College Biomedical Research Centre.

Sponsor

University Hospital Southampton NHS Foundation Trust.

Clinical Trials Unit

Mrs Jennifer Silk and Mrs Jessica Comper.

Data Monitoring and Ethics Committee

Professor Christopher Hawkey, Professor Philip Newsome and Mrs Debbie Hepworth.

Trial Steering Committee

Dr Stephen Falk, Dr Andrew Protheroe and Professor Sally Stenning.

Patient representative

Mr Colin Standfield.

Investigators

Dr Aileen Smith, Dr John Christie, Dr Keith George, Dr Christopher Shorrock, Professor Peter Mills, Dr Ian Shaw, Dr Hyder Hussaini, Dr Earl Williams, Dr Dara De Las Heras, Dr George Abouda, Dr Sulleman Moreea, Dr David Parker, Dr Anurag Agrawal, Dr Gurjit Singh, Dr Daniel Forton, Dr Richard Aspinall, Professor Graham Foster, Dr Mathis Heydtmann, Dr Nicholas Sharer, Dr Matthew Foxton, Dr Julia Maltby, Dr Sambit Sen, Dr Keith Dear, Dr Sushma Saksena, Dr John Dillon, Dr Chin Lye Ch'ng, Dr Joanne Topping, Dr Andrew Yeoman, Dr Andrew Holt, Dr Harriet Mitchison, Dr Judith Morris, Dr Mazn Karmo, Dr Anthoor Jayaprakash, Dr Subrata Saha, Dr Carole Collins, Dr John Greenaway, Dr Rudi Matull, Dr Andrew Davis, Professor Aftab Ala, Dr Saket Singhal, Dr Sharat Misra, Dr Alison Brind, Dr Sundaramoorthy Bharathi, Dr Mark Aldersley, Dr Gavin Wright, Dr Peter Bramley, Dr Aravamuthan Sreedharan, Dr Balasubramaniam Vijayan, Dr Sam Thompson, Dr Michael Perry, Dr Mark Welfare, Dr Andrew Fraser, Dr Gary James, Dr Hugh Dalziel, Dr Jonathan Mitchell, Dr Matthew Cramp, Dr Stephen Stewart, Dr Martin Phillips and Professor Mark Deverill.

Contributions of authors

Mark Thursz (Professor, Hepatology) was the chief investigator of the STOPAH trial.

Mark Thursz, **Ewan Forrest** (Consultant Physician, Gastroenterology), **Christopher Day** (Professor, Hepatology), **Andrew Austin** (Consultant Physician, Hepatology and Gastroenterology), **John O'Grady** (Professor, Hepatology), **Stephen Ryder** (Consultant Physician, Hepatology), **Michael Allison** (Consultant Hepatology), **Dermot Gleeson** (Professor, Hepatology), **Anne McCune** (Consultant Physician, Hepatology), **David Patch** (Consultant Physician, Hepatology) and **Mark Wright** (Consultant Physician, Gastroenterology) participated in the conception and design of the study.

Paul Roderick (Professor, Public Health), **Steven Masson** (Consultant Physician, Hepatologist), **Paul Richardson** (Consultant Physician, Hepatology), **Jane Mellor** (Clinical Trial Manager, SCTU), **Louise Stanton** (Senior Statistician, SCTU) and **Megan Bowers** (Senior Statistician, UoSCTU) designed the trial.

Luke Vale (Professor, Health Economics), **Scott Kirkman** (Health Economist), **Tara Homer** (Health Economist) and **Laura Ternent** (Health Economist) carried out the health economic analysis.

Ian Ratcliffe (Statistician UoSCTU), **Louise Stanton** and **Megan Bowers** designed the statistical analysis plan.

Nichola Downs (Clinical Trial Co-ordinator) was the trial co-ordinator and carried out the site monitoring.

Outcome data were analysed by **Megan Bowers** and **Ian Ratcliffe**.

Publications

Forrest E, Mellor J, Stanton L, Bowers M, Ryder P, Austin A, *et al.* STeroids Or Pentoxifylline for Alcoholic Hepatitis (STOPAH): study protocol for a randomised controlled trial. *Trials* 2013;**14**:262.

Thursz MR, Forrest EH, Ryder S, STOPAH investigators. Prednisolone or pentoxifylline for alcoholic hepatitis. *N Engl J Med* 2015;**373**:282–3.

Data sharing statement

Data can be obtained from the corresponding author after approval by the TMG.

References

1. Forrest E, Morris AJ, Stewart S, Phillips M, Oo YH, Fisher NC, *et al.* The Glasgow alcoholic hepatitis score identifies patients who may benefit from corticosteroids. *Gut* 2007;**56**:1743–6. http://dx.doi.org/10.1136/gut.2006.099226
2. Royal College of Physicians. *Alcohol Misuse: Can the NHS Afford it? Recommendations for a Coherent Alcohol Strategy for Hospitals: A Report of the Working Party of the Royal College of Physicians*. London: Royal College of Physicians; 2001.
3. Fisher NC, Hanson J, Phillips A, Rao JN, Swarbrick ET. Mortality from liver disease in the West Midlands, 1993–2000: observational study. *BMJ* 2002;**325**:312–13. http://dx.doi.org/10.1136/bmj.325.7359.312
4. Scottish Executive Health Economics Unit. *Cost to Society of Alcohol Misuse in Scotland, An Update to Alcohol Misuse in Scotland Trends and Costs*. Edinburgh: Scottish Government; 2004.
5. Leon DA, McCambridge J. Liver cirrhosis mortality rates in Britain from 1950–2002: an analysis of routine data. *Lancet* 2006;**367**:52–6. http://dx.doi.org/10.1016/S0140-6736(06)67924-5
6. Baker A, Rooney C. Recent trends in alcohol-related mortality and the impact of ICD-10 on the monitoring of deaths in England and Wales. *Health Stat Q* 2003;**17**:5.
7. Hislop WS, Bouchier IA, Allan JG, Brunt PW, Eastwood M, Finlayson ND, *et al.* Alcoholic liver disease in Scotland and northeastern England: presenting features in 510 patients. *Q J Med* 1983;**52**:232–43.
8. Maddrey WC, Boitnott JK, Bedine MS, Weber FL, Mezey E, White RI. Corticosteroid therapy of alcoholic hepatitis. *Gastroenterology* 1978;**75**:193–9.
9. Carithers RL Jr, Herlong HF, Diehl AM, Shaw EW, Combes B, Fallon HJ, *et al.* Methylprednisolone therapy in patients with severe alcoholic hepatitis. A randomized multicenter trial. *Ann Int Med* 1989;**110**:685–90. http://dx.doi.org/10.7326/0003-4819-110-9-685
10. Mathurin P, Mendenhall CL, Carithers RL Jr, Ramond MJ, Maddrey WC, Garstide P, *et al.* Corticosteroids improve short-term survival in patients with severe alcoholic hepatitis (AH): individual data analysis of the last three randomized placebo controlled double blind trials of corticosteroids in severe AH. *J Hepatol* 2002;**36**:480–7. http://dx.doi.org/10.1016/S0168-8278(01)00289-6
11. Morgan MY. Treatment of alcoholic hepatitis. *Alcohol Alcohol* 1996;**21**:117–34. http://dx.doi.org/10.1093/oxfordjournals.alcalc.a008123
12. Rambaldi A, Saconato HH, Christensen E, Thorlund K, Wetterslev J, Gluud C. Systematic review: glucocorticosteroids for alcoholic hepatitis – a Cochrane Hepato-Biliary Group systematic review with meta-analyses and trial sequential analyses of randomized clinical trials. *Aliment Pharmacol Ther* 2008;**27**:1167–78. http://dx.doi.org/10.1111/j.1365-2036.2008.03685.x
13. Mendenhall CL, Anderson S, Garcia-Pont P, Goldberg S, Kiernan T, Seeff LB, *et al.* Short-term and long-term survival in patients with alcoholic hepatitis treated with oxandrolone and prednisolone. *N Engl J Med* 1984;**311**:1464–70. http://dx.doi.org/10.1056/NEJM198412063112302
14. Ramond MJ, Poynard T, Rueff B, Mathurin P, Théodore C, Chaput JC, *et al.* A randomized trial of prednisolone in patients with severe alcoholic hepatitis. *N Engl J Med* 1992;**326**:507–12. http://dx.doi.org/10.1056/NEJM199202203260802

15. Mathurin P, Duchatelle V, Ramond MJ, Degott C, Bedossa P, Erlinger S, *et al.* Survival and prognostic factors in patients with severe alcoholic hepatitis treated with prednisolone. *Gastroenterology* 1996;**110**:1847–53. http://dx.doi.org/10.1053/gast.1996.v110.pm8964410

16. Christensen E, Gluud C. Glucocorticoids are ineffective in alcoholic hepatitis: a meta-analysis adjusting for confounding variables. *Gut* 1995;**37**:113–18. http://dx.doi.org/10.1136/gut.37.1.113

17. Mathurin P, O'Grady J, Carithers RL, Phillips M, Louvet A, Mendenhall CL, *et al.* Corticosteroids improve short-term survival in patients with severe alcoholic hepatitis: meta-analysis of individual patient data. *Gut* 2011;**60**:255–60. http://dx.doi.org/10.1136/gut.2010.224097

18. Akriviadis E, Botla R, Briggs W, Han S, Reynolds T, Shakil O. Pentoxifylline improves short-term survival in severe acute alcoholic hepatitis: a double-blind, placebo-controlled trial. *Gastroenterology* 2000;**119**:1637–48. http://dx.doi.org/10.1053/gast.2000.20189

19. Sidhu SS, Goyal O, Singla M, Bhatia KL, Chhina RS, Sood A. Pentoxifylline in severe alcoholic hepatitis: a prospective, randomised trial. *J Assoc Phys India* 2012;**60**:20–2.

20. Whitfield K, Rambaldi A, Wetterslev J, Gluud C. Pentoxifylline for alcoholic hepatitis. *Cochrane Database Syst Rev* 2009;**4**:CD007339. http://dx.doi.org/10.1002/14651858.cd007339.pub2

21. Parker R, Armstrong MJ, Corbett C, Rowe IA, Houlihan DD. Systematic review: pentoxifylline for the treatment of severe alcoholic hepatitis. *Aliment Pharmacol Ther* 2013;**37**:845–54. http://dx.doi.org/10.1111/apt.12279

22. Singal AK, Salameh H, Singal A, Jampana SC, Freeman DH, Anderson KE, *et al.* Management practices of hepatitis C virus infected alcoholic hepatitis patients: a survey of physicians. *World J Gastrointest Pharmacol Ther* 2013;**4**:16–22. http://dx.doi.org/10.4292/wjgpt.v4.i2.16

23. Mathurin P, Louvet A, Duhamel A, Nahon P, Carbonell N, Boursier J, *et al.* Prednisolone with vs. without pentoxifylline and survival of patients with severe alcoholic hepatitis: a randomized clinical trial. *JAMA* 2013;**310**:1033–41. http://dx.doi.org/10.1001/jama.2013.276300

24. Sidhu SS, Goyal O, Singla P, Gupta D, Sood A, Chhina RS, *et al.* Corticosteroid plus pentoxifylline is not better than corticosteroid alone for improving survival in severe alcoholic hepatitis (COPE trial). *Dig Dis Sci* 2012;**57**:1664–71. http://dx.doi.org/10.1007/s10620-012-2097-4

25. De BK, Gangopadhyay S, Dutta D, Baksi SD, Pani A, Ghosh P. Pentoxifylline versus prednisolone for severe alcoholic hepatitis: a randomized controlled trial. *World J Gastroenterol* 2009;**15**:1613–19. http://dx.doi.org/10.3748/wjg.15.1613

26. Garcoa J, Hernandez G, Lopez A, Barrera C, Vargas L. Pentoxifilina versus esteroide en la sobrevivencia a corto plazo enhepatitis aguda alcoholica severa. *Med Interna Mex* 2012;**28**:228.

27. Park SH, Kim DJ, Kim YS, Yim HJ, Tak WY, Lee HJ, *et al.* Pentoxifylline vs. corticosteroid to treat severe alcoholic hepatitis: a randomised, non-inferiority, open trial. *J Hepatol* 2014;**61**:792–8. http://dx.doi.org/10.1016/j.jhep.2014.05.014

28. Louvet A, Diaz E, Dharancy S, Coevoet H, Texier F, Thévenot T, *et al.* Early switch to pentoxifylline in patients with severe alcoholic hepatitis is inefficient in non-responders to corticosteroids. *J Hepatol* 2008;**48**:465–70. http://dx.doi.org/10.1016/j.jhep.2007.10.010

29. Forrest E, Mellor J, Stanton L, Bowers M, Ryder P, Austin A, *et al.* Steroids or pentoxifylline for alcoholic hepatitis (STOPAH): study protocol for a randomised controlled trial. *Trials* 2013;**14**:262.

30. Pocock SJ. Current controversies in data monitoring for clinical trials. *Clin Trial* 2006;**3**:513–21. http://dx.doi.org/10.1177/1740774506073467

31. Nguyen-Khac E, Thevenot T, Piquet MA, Benferhat S, Goria O, Chatelain D, *et al.* Glucocorticoids plus *n*-acetylcysteine in severe alcoholic hepatitis. *N Engl J Med* 2011;**365**:1781–9. http://dx.doi.org/10.1056/NEJMoa1101214

32. Mathurin P, Lucey MR. Management of alcoholic hepatitis. *J Hepatol* 2012;**56**(Suppl. 1):S39–45. http://dx.doi.org/10.1016/S0168-8278(12)60005-1

33. Forrest EH, Evans CD, Stewart S, Phillips M, Oo YH, McAvoy NC, *et al.* Analysis of factors related to mortality in alcoholic hepatitis and the derivation and validation of the Glasgow alcoholic hepatitis score. *Gut* 2002;**54**:1174–9. http://dx.doi.org/10.1136/gut.2004.050781

34. Kind P, Hardman G, Leese B. Measuring health status: information for primary care decision making. *Health Policy* 2005;**71**:303–13. http://dx.doi.org/10.1016/j.healthpol.2004.02.008

35. SF-36.org. *Short Form 36.* URL: www.sf-36.org (accessed 1 October 2014).

36. Younossi ZM, Guyatt G, Kiwi M, Boparai N, King D. Development of a disease specific questionnaire to measure health related quality of life in patients with chronic liver disease. *Gut* 1999;**45**:295–300. http://dx.doi.org/10.1136/gut.45.2.295

37. Gralnek IM, Hays RD, Kilbourne A, Rosen HR, Keeffe EB, Artinian L, *et al.* Development and evaluation of the Liver Disease Quality of Life instrument in persons with advanced, chronic liver disease – the LDQOL 1.0. *Am J Gastroenterol* 2000;**95**:3552–65.

38. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the Economic Evaluation of Health Care Programmes*. 3rd edn. Oxford: Oxford University Press; 2005.

39. Von Neumann J, Morgenstern O. *Theory of Games and Economic Behaviour*. 3rd edn. New York, NY: John Wiley & Sons; 1953.

40. McLernon DJ, Dillon J, Donnan PT. Health-state utilities in liver disease: a systematic review. *Med Decis Mak* 2008;**28**:582–92. http://dx.doi.org/10.1177/0272989X08315240

41. Child CG, Turcotte JG. Surgery and Portal Hypertension. In: Child CG, editor. *The Liver and Portal Hypertension*. Philadelphia, PA: WB Saunders; 1964. pp. 1–85.

42. Pugh RN, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. *Br J Surg* 1973;**60**:646–9. http://dx.doi.org/10.1002/bjs.1800600817

43. Mason H, Donaldson N, McColl E, Day C. *Quality of Life and Alcoholic Cirrhosis of the Liver: An Economic Approach*. Newcastle: Health Economists Study Group; 2004.

44. Furlong W, Feeny D, Torrance G, Barr R, Horsman J. *Guide to Design and Development of Health State Utility Instrumentation*. Hamilton, ON: McMaster University; 1990.

45. Dolan P, Stalmeier P. The validity of time trade-off values in calculating QALYs: constant proportional time trade-off versus the proportional heuristic. *J Health Econ* 2003;**22**:445–58. http://dx.doi.org/10.1016/S0167-6296(02)00120-0

46. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal 2013*. London: National Insitute for Health and Care Excellence; 2013.

47. Department of Health. *NHS Reference Costs 2011–12*. London: Department of Health; 2012.

48. Joint Formulary Committee. *British National Formulary* (online). London: BMJ Group and Pharmaceutical Press; 2014. URL: www.medicinescomplete.com (accessed July 2014).

49. Newcastle Biomedicine. *Newcastle Joint Research Office*. URL: www.newcastlejro.org.uk/wp.../11/RD-Valid-Submission-Checklist.doc

50. Curtis L. *Unit Costs of Health and Social Care 2012*. Personal Social Services Research Unit: University of Kent; 2012.

51. Brazier J, Ratcliffe J, Tsuchiya A, Salomon JA. *Measuring and Valuing Health Benefits for Economic Evaluation*. Oxford: Oxford University Press; 2007.

52. Matthews JN, Campbell MJ, Altman DG, Royston P. Analysis of serial measurements in medical research. *BMJ* 1990;**300**:230–5.

53. European Association for the Study of Liver. EASL clinical practical guidelines: management of alcoholic liver disease. *J Hepatol* 2012;**57**:399–420.

54. Mathurin P, Moreno C, Samuel D, Dumortier J, Salleron J, Durand F, *et al.* Early liver transplantation for severe alcoholic hepatitis. *N Engl J Med* 2011;**365**:1790–800. http://dx.doi.org/10.1056/NEJMoa1105703

55. Forrest E, Gleeson D. Is a liver biopsy necessary in alcoholic hepatitis? *J Hepatol* 2012;**56**:1428–9. http://dx.doi.org/10.1016/j.jhep.2011.12.028

56. Hamid R, Forrest EH. Is histology required for the diagnosis of alcoholic hepatitis: a review of published randomized controlled trials. *Gut* 2011;**60**:A233.

57. Elphick Dube AK, McFarlane E, Jones J, Gleeson D. Spectrum of liver histology in presumed decompensated alcoholic liver disease. *Am J Gastroenterol* 2007;**102**:780–8.

58. Cabre E, Rodriguez-Iglesias P, Caballeria J, Quer JC, Sánchez-Lombraña JL, Parés A, *et al.* Short- and long-term outcome of severe alcohol-induced hepatitis treated with steroids or enteral nutrition: a multicenter randomized trial. *Hepatology* 2000;**32**:36–42. http://dx.doi.org/10.1053/jhep.2000.8627

59. Potts J, Goubet S, Heneghan M, Verma S. Determinants of long term outcome in severe alcoholic hepatitis. *Aliment Pharmacol Ther* 2013;**38**:584–95. http://dx.doi.org/10.1111/apt.12427

60. Boetticher NC, Peine CJ, Kwo P, Abrams GA, Patel T, Aqel B, *et al.* A randomized, double-blinded, placebo-controlled multicenter trial of etanercept in the treatment of alcoholic hepatitis. *Gastroenterology* 2008;**135**:1953–60.

61. Louvet A, Wartel F, Castel H, Dharancy S, Hollebecque A, Canva-Delcambre V, *et al.* Infection in patients with severe alcoholic hepatitis treated with steroids: early response to therapy is the key factor. *Gastroenterology* 2009;**137**:541–8. http://dx.doi.org/10.1053/j.gastro.2009.04.062

62. O'Shea RS, Dasarathy S, McCullough AJ. Alcoholic liver disease. *Hepatology* 2009;**51**:307–28.

63. National Confidential Enquiry into Patient Outcome and Death. *Alcohol Related Liver Disease: Measuring the Units (2013)*. 2013. URL: www.ncepod.org.uk/2013arld.htm (accessed 14 October 2014).

64. Singh V, Sharma AK, Narasimhan RL, Bhalla A, Sharma N, Sharma R. Granulocyte colony-stimulating factor in severe alcoholic hepatitis: a randomized pilot study. *Am J Gastroenterol* 2014;**109**:1417–23.

65. Addolorato G, Leggio L, Ferrulli A, Cardone S, Vonghia L, Mirijello A, *et al.* Effectiveness and safety of baclofen for maintenance of alcohol abstinence in alcohol-dependent patients with liver cirrhosis: randomised, double-blind controlled study. *Lancet* 2007;**370**:1915–22. http://dx.doi.org/10.1016/S0140-6736(07)61814-5

66. Petts G, Lloyd K, Vergis N, Kudo H, Quaglia A, Forrest E, *et al.* Utility of liver biopsy in alcoholic hepatitis: Data from the Steroids or Pentoxyfilline in Alcoholic Hepatitis (STOPAH) clinical trial. *J Hepatol* 2015;**62**(Suppl. 1):S776–7.

67. Antoniades CG, Khamri W, Abeles RD, Taams LS, Triantafyllou E, Possamai LA, *et al.* Secretory leukocyte protease inhibitor: a pivotal mediator of anti-inflammatory responses in acetaminophen-induced acute liver failure. *Hepatology* 2014;**59**:1564–76.

68. Antoniades CG, Quaglia A, Taams LS, Mitry RR, Hussain M, Abeles R, *et al.* Source and characterization of hepatic macrophages in acetaminophen-induced acute liver failure in humans. *Hepatology* 2012;**56**:735–46.

69. Bernsmeier C, Pop OT, Singanayagam A, Triantafyllou E, Patel VC, Weston CJ, *et al.* Patients with acute-on-chronic liver failure have increased numbers of regulatory immune cells expressing the receptor tyrosine kinase MERTK. *Gastroenterology* 2015;**148**:603–15.

70. NHS. *NHS Blood and Transplant*. URL: www.nhsbt.nhs.uk (accessed 21 May 2014).

Appendix 1 Prognostic score formula

Discriminant function

$$DF = 4.6 \times (\text{prothrombin time (PTPATIENT} - \text{PTCONTROL)} + \text{serum bilirubin (}\mu\text{mol/l)}/17.1. \quad (4)$$

(PTCONTROL is defined as the mean value at each centre; this mean value may be updated on a weekly or monthly basis.)

TABLE 46 Glasgow Alcoholic Hepatitis Score

Parameter	1	2	3
Age (years)	< 50	≥ 50	–
WBC ($\times 10^9$)	< 15	≥ 15	–
Urea (mmol)	< 5	≥ 5	–
Prothrombin ratio	< 1.5	1.5–2.0	> 2.0
Bilirubin (μmol)	< 125	125–250	> 250

Lille score

$$\begin{aligned}\text{Lille score} = {} & 3.19 - 0.101 \times (\text{age in years}) + 0.147 \times (\text{albumin day 0 in g/l}) + 0.0165 \\ & \times(\text{evolution in bilirubin level in }\mu\text{mol}) - 0.206 \times (\text{renal insufficiency}) - 0.0065 \\ & \times(\text{bilirubin day 0 in }\mu\text{m}) - 0.0096 \times (\text{prothrombin time in seconds}). \end{aligned} \quad (5)$$

Model for end-stage liver disease score

$$MELD = 9.57 \times \text{Ln (creatinine)} + 3.78 \times \text{Ln(bilirubin)} + 11.2 \times \text{Ln(INR)} + 6.43. \quad (6)$$

Appendix 2 Substudies

Central pathology review

In patients with AH, liver biopsy is rarely performed in routine clinical practice as strict application of clinical criteria accurately define the condition in the majority of cases. Nevertheless, some units routinely practice liver biopsy for confirmation and a number of biopsies are performed owing to diagnostic uncertainty when an alternative diagnosis may be implicated. In many clinical trials, liver biopsy confirmed diagnosis of AH has been an inclusion criterion. The STOPAH trial used a clinical (rather than histological) diagnosis as an inclusion criterion. We therefore sought to validate the clinical criterion by auditing the histology diagnoses on those cases for which routine liver biopsy had been performed.

Two experienced pathologists (Professor Robert Goldin and Professor Alberto Quaglia) sent liver histology sections for central review. In total, 200 liver biopsies were received, including 120 that had been performed as part of routine clinical practice and 80 that had been performed for diagnostic uncertainty. All biopsies were scored for the following characteristics:

- quality of biopsy
- fibrosis
- steatosis
- inflammation
- hepatocyte damage
- bilirubinostasis
- megamitochondria
- miscellaneous findings, for example iron, dysplasia, malignancy
- advanced/not advanced cirrhosis
- active/not very active
- consistent/not consistent with alcohol
- biopsy diagnostic of alcoholic steatohepatitis.

Among biopsies performed as routine clinical practice, the diagnosis of AH was confirmed in 95% of cases. Other diagnoses included advanced cirrhosis without hepatitis, drug-induced liver injury and a case of hepatocellular carcinoma.

This review confirmed that a clinical diagnosis of AH is accurate in the majority of cases. A publication describing these findings and evaluating the ability of histological characteristics to predict the outcome and prognosis will be submitted for publication in December 2014.[66]

Serum samples

Serum samples were collected from all patients enrolled in the STOPAH trial. The serum was processed from clotted venous blood samples at each centre and stored at –80 °C and transferred to the Health Technology Assessment (HTA)-licensed biobank at Imperial College. The TMG gave approval for the serum to be used in two studies:

1. An analysis of metabolomic biomarkers using 1H-magnetic resonance spectroscopy and ultra high-performance liquid chromatography coupled to mass spectrometry. These analyses will seek metabolite or metabolic profiles that correlate with prognosis and outcome of treatment. In addition, serum samples from a group of patients with decompensated cirrhosis will be used as comparators to identify novel biomarkers for use in distinguishing AH from decompensated cirrhosis.
2. Serum samples taken from patients who underwent liver biopsy during the trial will be used to identify serum markers of liver fibrosis in the Nottingham University Hospitals NHS Trust and National Institute for Health Research Biomedical Research Unit.

Genetic susceptibility in alcoholic hepatitis

Investigators were asked to collect 5 ml of ethylenediaminetetraacetic acid whole blood from all consenting STOPAH patients. These samples were transported to Imperial College and stored in the HTA-licensed biobank. Genomic deoxyribonucleic acid (DNA) was extracted from 325 samples and sent for whole exome single nucleotide polymorphism genotyping at the Sanger Centre in Cambridge. The same number of DNA samples from a control group of patients with alcohol use disorders who did not have any history or laboratory evidence of liver disease were also sent for genotyping and a genome-wide association study was conducted after quality control procedures. Evidence of genetic association at a threshold of $p < 10^{-5}$ has been identified at a number of loci. DNA from the remaining patients and an additional cohort of controls is now being used in genotyping assays focused on candidate loci to replicate the associations.

Monocyte analysis

As demonstrated in the STOPAH study, as well as previous trials in AH,[63] patients with this condition are highly susceptible to infection, which is associated with higher rates of mortality. Previous work by Dr Antoniades and Professor Thursz[67–69] has demonstrated a defect in the function of circulating monocytes in acute liver failure, suggesting that a similar problem may explain infection susceptibility among patients with AH. Peripheral blood samples collected from patients in the STOPAH trial were used to phenotype and functionally characterise the circulating monocyte population. To date, it has been discovered that monocytes in AH are capable of phagocytosing bacteria, but they are unable to produce the respiratory burst in the lysosomal compartments that are necessary to kill ingested organisms. The oxidative burst defect was most pronounced in those patients who went on to develop infection. Further investigations are under way to identify the molecular basis of the oxidative burst defect and to find therapeutic interventions to improve monocyte function.

Past medication analysis

It is not known what triggers AH in heavy drinkers. One supposition is that specific drugs may either precipitate or be protective of acute injury. Thirteen STOPAH sites were open to enter patients into the STOPAH trial and the GP substudy; however, participation in this substudy was not compulsory. It involved gaining an electronic download of all medications prescribed in the last year from the patient's GP. Owing to a poor response, it was decided by the TMG on 4 February 2013 that the substudy should not continue.

Appendix 3 Sensitivity analysis of the primary outcome

TABLE 47 Observed 28-day mortality (sensitivity analysis)[a]

	PTX		
	No	Yes	Total
Prednisolone			
No	16.0% (39/244) (placebo/placebo)	19.6% (45/230) (placebo/PTX)	17.7% (84/474)
Yes	14.3% (34/237) (prednisolone/placebo)	12.7% (29/228) (prednisolone/PTX)	13.5% (63/465)
Total	15.2% (73/481)	16.2% (74/458)	15.7% (147/939)

a Includes all randomised patients who have at least 28 days of data or died prior to or on day 28 excluding any patients who were randomised and followed up but were later found to be ineligible.

TABLE 48 Logistic regression analysis for 28-day mortality (sensitivity analysis)[a]

	Prednisolone ($n = 465$)	No prednisolone ($n = 474$)	PTX ($n = 458$)	No PTX ($n = 481$)
28-day mortality,% (*n*/*N*)	13.5% (63/465)	17.7% (84/474)	16.2% (74/458)	15.2% (73/481)
Adjusted odds ratio[b] (95% CI)	0.71 (0.50 to 1.02)		1.08 (0.76 to 1.54)	
p-value	0.065		0.666	

a Includes all randomised patients who have at least 28 days of data or died prior to or on Day 28 excluding any patients who were randomised and followed up but were later found to be ineligible.
b Adjusted for true risk at baseline and factorial design.

Notes

Odds ratios of < 1 represent a favourable outcome for the corresponding intervention.
The interaction between the interventions was investigated as a secondary analysis and was found to be non-significant at the 5% significance level (interaction coefficient = –0.403, (95% CI –1.120 to 0.313; *p*-value = 0.270).
Logistic regression model: 28-day mortality = intercept + prednisolone indicator + PTX indicator + risk.

Appendix 4 Unit cost: intervention, initial admission and adverse events costs

TABLE 49 Unit cost: intervention, initial admission and adverse events costs

	Unit	Cost, £	Source
Drug costs: intervention			
Prednisolone	40 mg × 1 daily (× 28)	10.48	CRF/BNF
PTX	400 mg × 3 daily (× 28)	18.10	CRF/BNF
Inpatient costs			
General ward	Per night: non-elective inpatient (long stay) excess bed-days	265	Discharge visit form/NHS tariffs
HDU	Per night: critical care, one organ supported	981	Discharge visit form/NHS tariffs
ICU	Per night: critical care, three organs supported	1617	Discharge visit form/NHS tariffs
Nutritional supplements (1.5 kcal per ml)	Per bottle	2.06	Discharge visit form/BNF
Lab tests			
Creatinine	Per test	2.96	Treatment-day forms/discharge form/visit forms/lab tests
AST	Per test	2.96	Treatment-day forms/discharge form/visit forms/lab tests
ALT	Per test	2.96	Treatment-day forms/discharge form/visit forms/lab tests
Sodium	Per test	2.96	Treatment-day forms/discharge form/visit forms/lab tests
Phosphate	Per test	2.96	Treatment-day forms/discharge form/visit forms/lab tests
Glucose	Per test	2.96	Treatment-day forms/discharge form/visit forms/lab tests
Albumin	Per test	2.96	Treatment-day forms/discharge form/visit forms/lab tests
Urea	Per test	2.96	Treatment-day forms/discharge form/visit forms/lab tests
ALP	Per test	2.96	Treatment-day forms/discharge form/visit forms/lab tests
Bilirubin	Per test	2.96	Treatment-day forms/discharge form/visit forms/lab tests
Potassium	Per test	2.96	Treatment-day forms/discharge form/visit forms/lab tests
Calcium	Per test	2.96	Treatment-day forms/discharge form/visit forms/lab tests
Total protein	Per test	2.96	Treatment-day forms/discharge form/visit forms/lab tests

continued

TABLE 49 Unit cost: intervention, initial admission and adverse events costs (*continued*)

	Unit	Cost, £	Source
INR	Per test	7.49	Treatment-day forms/discharge form/visit forms/lab tests
FBC	Per test	4.94	Treatment-day forms/discharge form/visit forms/lab tests
PT	Per test	5.34	Treatment-day forms/discharge form/visit forms/lab tests
Discharge procedures			
CT scan – one area	One area no contrast	92	Discharge visit form/NHS tariffs
CT scan – two areas	Two areas no contrast	111	
MRI scan			Discharge visit form/NHS tariffs
Endoscopy	Diagnostic endoscopic upper GI tract procedures	184.20	Discharge visit form/NHS tariffs
Colonoscopy			Discharge visit form/NHS tariffs
Renal replacement	Hospital haemofiltration, with access via haemodialysis catheter	147	Discharge visit form/NHS tariffs
Other discharge procedures and adverse event procedures			
Ultrasound	One site (assumed less than 20 minutes)	61	Discharge visit form/adverse event form/NHS tariffs
Ultrasound	Two sites (assume greater than 20 minutes)	64	Discharge visit form/adverse event form/NHS tariffs
Ultrasound Doppler/portal system	Two sites (assume greater than 20 minutes)	64	Discharge visit form/adverse event form/NHS tariffs
Echocardiogram		72	Discharge visit form/adverse event form/NHS tariffs
Chest drain		331.80	Discharge visit form/adverse event form/NHS tariffs
Chest X-ray	Skeletal X-ray as it includes chest X-ray	31	Discharge visit form/adverse event form/UCL provider to provider tariffs
Spinal X-ray		43	Discharge visit Form/adverse event form/UCL provider to provider tariffs
Other X-ray	Skeletal X-ray	31	Discharge visit form/adverse event form/UCL provider to provider tariffs
Sigmoidoscopy	Diagnostic sigmoidoscopy	520.05	Discharge visit form/adverse event form/NHS tariffs
MRI scan	One area no contrast	165	Discharge visit form/adverse event form/NHS tariffs
Colonoscopy	Diagnostic colonoscopy	238.95	Discharge visit form/adverse event form/NHS tariffs
Draining ascites	Minor therapeutic/diagnostic general abdominal procedures	416.35	Discharge visit form/adverse event form/NHS tariffs
Gastroscopy	Diagnostic endoscopic upper GI tract procedures	184.20	Discharge visit form/adverse event form/NHS tariffs
MRCP	MRI scan, one area, no contrast	169	Discharge visit form/adverse event form/NHS tariffs

TABLE 49 Unit cost: intervention, initial admission and adverse events costs (*continued*)

	Unit	Cost, £	Source
Liver biopsy	Minor endoscopic or percutaneous hepatobiliary or pancreatic procedures	351.40	Discharge visit form/adverse event form/NHS tariffs
TIPS procedure	Intermediate endoscopic or percutaneous hepatobiliary or pancreatic procedures w/CC score 0–1	747.70	Discharge visit form/adverse event form/NHS tariffs
ERCP	Minor diagnostic ERCP	699.90	Discharge visit form/adverse event form/NHS tariffs
Hepatic wedge pressure	Intermediate endoscopic or percutaneous hepatobiliary or pancreatic procedures w/CC score 0–1	164	Discharge visit form/adverse event form/NHS tariffs
ECG	ECG monitoring and stress testing	122	Discharge visit form/adverse event form/NHS tariffs
PICC line	Liver failure disorders, with single intervention	622.90	Discharge visit form/adverse event form/NHS tariffs
Central line	Liver failure disorders, with single intervention	622.90	Discharge visit form/adverse event form/NHS tariffs
Pleural tap	Minor thoracic procedures	751.96	Discharge visit form/adverse event form/NHS tariffs
Urinary tract infections	Kidney or urinary tract infections, without interventions, with CC score 0–1	329.40	Discharge visit form/adverse event form/NHS tariffs
Blood transfusion		123	Discharge visit form/adverse event form/www.nhsbt.nhs.uk[70]
Ultrasound guided aspiration		124	Discharge visit form/adverse event form/UCL provider to provider tariffs
Endoscopic ultrasound cost	Minor therapeutic or diagnostic general abdominal procedures	416.35	Discharge visit form/adverse event form/NHS tariffs
Blood test	Blood culture	17	Discharge visit form/adverse event form/lab costs
Stool test	Faecal test microbiology	13.20	Discharge visit form/adverse event form/lab costs
Urine test	Urine culture	8.51	Discharge visit form/adverse event form/lab costs
Hepatitis E test		35.98	Discharge visit form/adverse event form/lab costs
HDU	Critical care: one organ supported	981	Discharge visit form/adverse event form/NHS tariffs
ICU	Critical care: three organs supported	1617	Discharge visit form/NHS tariffs
IVU	Dynamic studies of urinary tract	380	Discharge visit form/adverse event form/NHS tariffs
NG tube	Nutritional disorders, with interventions	127	Discharge visit form/adverse event form/NHS tariffs
Counselling	Alcohol health worker (A&E), 55-minute consultation	48	Discharge visit form/adverse event form/PSSRU[50]
Psychological consultation	Consultant-led outpatient attendances – non-admitted face-to-face first appointment	264	Discharge visit form/adverse event form/NHS tariffs

continued

TABLE 49 Unit cost: intervention, initial admission and adverse events costs (*continued*)

	Unit	Cost, £	Source
Blood test	Sodium, potassium, urea, creatinine, glucose	14.80	Discharge visit form/adverse event form/lab costs
LFT		6.80	Discharge visit form/adverse event form/lab costs
Blood test	FBC	4.94	Discharge visit form/adverse event form/lab costs
Oesophageal surgery (perforation)	Complex oesophageal, stomach or duodenum procedures, 19+ years	3713	Discharge visit form/adverse event form/NHS tariffs
Duplex scan	Ultrasound mobile scan or intraoperative procedures, 20–40 minutes	71	Discharge visit form/adverse event form/NHS tariffs
Dietitian consultation	45-minute consultation (assumed first visit)	23.25	Discharge visit form/adverse event form/PSSRU[50]
Renal blood test	Urea	2.96	Discharge visit form/adverse event form/lab costs
Removal of tooth	Surgical removal of tooth, ≥ 19 years	1580	Discharge visit form/adverse event form/NHS tariffs
Surgical consultation	Consultant-led outpatient attendances – non-admitted face-to-face first appointment	153	Discharge visit form/adverse event form/NHS tariffs
Ophthalmology attendance	Outpatient appointment	86	Discharge visit form/adverse event form/NHS tariffs
Septic screen	FBC, urea and electrolytes, CRP, blood cultures, urine sample and chest X-ray	70.25	Discharge visit form/adverse event form/NHS tariffs/lab costs
Fluid swab (ascites fluid)	Wound and fluid samples – microbiology	20.33	Discharge visit form/adverse event form/lab costs
Sputum test	Sputum culture – microbiology	25.74	Discharge visit form/adverse event form/lab costs
Haematologist consultation	Consultant-led outpatient attendances – non-admitted face-to-face first appointment	209	Discharge visit form/adverse event form/NHS tariffs
Immunology test (blood)	Immunology autoantibodies	20.71	Discharge visit form/adverse event form/lab costs
Alcohol test (blood test)	Alcohol – whole blood – biochemistry	7.78	Discharge visit form/adverse event form/lab costs
Neurology consultation	Consultant-led outpatient attendances – non-admitted face-to-face first appointment	209	Discharge visit form/adverse event form/NHS tariffs
Dermatology consultation	Consultant-led outpatient attendances – non-admitted face-to-face first appointment	110	Discharge visit form/adverse event form/NHS tariffs
Cystoscopy	Diagnostic flexible cystoscopy, ≥ 19 years	675.45	Discharge visit form/adverse event form/NHS tariffs
Gynaecology exam		149	Discharge visit form/adverse event form/NHS tariffs
Spinal tap		130	Discharge visit form/adverse event form/NHS tariffs
Vascular consultation	Consultant-led outpatient attendances – non-admitted face-to-face first appointment	133	Discharge visit form/adverse event form/NHS tariffs

TABLE 49 Unit cost: intervention, initial admission and adverse events costs (*continued*)

	Unit	Cost, £	Source
Bone marrow aspirate	Biopsy of bone marrow – haematology	301.50	Discharge visit form/adverse event form/NHS tariffs
Anal procedure	Major anal procedures	1219.15	Discharge visit form/adverse event form/NHS tariffs
Eye infirmary consultation	Directly accessed diagnostic services. Non-Surgical ophthalmology, without interventions, with CC score 0–1	110	Discharge visit form/adverse event form/NHS tariffs
Banding of varices	GI with single intervention	522	Discharge visit form/adverse event form/NHS tariffs
Bronchial lavage	Fibre optic bronchoscopy, ≥ 19 years	299.40	Discharge visit form/adverse event form/NHS tariffs

A&E, accident and emergency; CC, complications and comorbidities; CRP, C-reactive protein; CT, computerised tomography; ECG, electrocardiogram; ERCP, endoscopic retrograde cholangiopancreatography; HDU, high-dependency unit; IVP, intravenous pyelogram; IVU, intravenous urogram; MRCP magnetic resonance cholangiopancreatography; MRI, magnetic resonance imaging; NG nasogastric; PICC, peripherally inserted central catheter; PSSRU, Personal Social Services Research Unit; TIPS, transjugular intrahepatic portosystemic shunt; UCL, University College London; w/CC, with complications and comorbidities.

Appendix 5 Average adverse event investigations per treatment arm

TABLE 50 Average adverse event investigations per treatment arm

	OO		AO		OB		AB	
Investigations	*n*	Mean (SD)	*n*	Mean (SD)	*n*	Mean (SD)	*n*	Mean (SD)
Stool test[a]	272	0.018 (0.160)	274	0.018 (0.134)	273	0.022 (0.147)	272	0.011 (0.135)
CT scan	272	0.007 (0.086)	274	0.022 (0.170)	273	0.007 (0.085)	272	0.011 (0.105)
MRI scan	272	0.007 (0.086)	274	0.004 (0.060)	273	0.004 (0.061)	272	0.000 (0.000)
Echocardiogram	272	0.007 (0.121)	274	0.004 (0.060)	273	0.007 (0.121)	272	0.000 (0.000)
Psychological consultation	272	0.000 (0.000)	274	0.004 (0.060)	273	0.000 (0.000)	272	0.000 (0.000)
Blood test	272	0.018 (0.135)	274	0.018 (0.159)	273	0.022 (0.147)	272	0.000 (0.000)
Other X-ray	272	0.015 (0.148)	274	0.004 (0.060)	273	0.004 (0.061)	272	0.004 (0.061)
Chest X-ray	272	0.052 (0.280)	274	0.058 (0.264)	273	0.055 (0.333)	272	0.052 (0.409)
Spinal X-ray	272	0.004 (0.061)	274	0.000 (0.000)	273	0.004 (0.061)	272	0.000 (0.000)
Doppler ultrasound	272	0.000 (0.000)	274	0.007 (0.85)	273	0.000 (0.000)	272	0.004 (0.061)
Gastroscopy	272	0.029 (0.190)	274	0.044 (0.293)	273	0.015 (0.120)	272	0.044 (0.254)
Urine test	272	0.011 (0.105)	274	0.007 (0.085)	273	0.007 (0.085)	272	0.007 (0.086)
Blood culture	272	0.037 (0.331)	274	0.029 (0.208)	273	0.018 (0.134)	272	0.029 (0.208)
Duplex scan	272	0.007 (0.121)	274	0.000 (0.000)	273	0.000 (0.000)	272	0.000 (0.000)
Endoscopy	272	0.004 (0.061)	274	0.007 (0.085)	273	0.015 (0.120)	272	0.004 (0.061)
Renal blood test	272	0.000 (0.000)	274	0.000 (0.000)	273	0.000 (0.000)	272	0.004 (0.061)
Ascites draining	272	0.040 (0.301)	274	0.033 (0.198)	273	0.018 (0.134)	272	0.033 (0.217)
Sigmoidoscopy	272	0.004 (0.301)	274	0.007 (0.085)	273	0.007 (0.085)	272	0.011 (0.105)
Septic screen	272	0.022 (0.226)	274	0.000 (0.000)	273	0.000 (0.000)	272	0.000 (0.000)
ECG	272	0.015 (0.121)	274	0.026 (0.180)	273	0.026 (0.314)	272	0.015 (0.120)
Fluid swab	272	0.011 (0.135)	274	0.000 (0.000)	273	0.004 (0.061)	272	0.000 (0.000)
Blood transfusion	272	0.000 (0.000)	274	0.000 (0.000)	273	0.004 (0.061)	272	0.000 (0.000)
Catheter	272	0.004 (0.061)	274	0.000 (0.000)	273	0.000 (0.000)	272	0.000 (0.000)
Immunology test	272	0.000 (0.000)	274	0.000 (0.000)	273	0.000 (0.000)	272	0.004 (0.060)
Alcohol test	272	0.000 (0.000)	274	0.000 (0.000)	273	0.004 (0.061)	272	0.000 (0.000)
Dermatology appointment	272	0.000 (0.000)	274	0.000 (0.000)	273	0.007 (0.085)	272	0.000 (0.000)
Cystoscopy	272	0.004 (0.061)	274	0.000 (0.000)	273	0.004 (0.061)	272	0.000 (0.000)
Gynaecology exam	272	0.004 (0.061)	274	0.000 (0.000)	273	0.000 (0.000)	272	0.000 (0.000)
Banding of varices	272	0.000 (0.000)	274	0.000 (0.000)	273	0.015 (0.120)	272	0.015 (0.171)

continued

TABLE 50 Average adverse event investigations per treatment arm (*continued*)

	OO		AO		OB		AB	
Investigations	*n*	Mean (SD)	*n*	Mean (SD)	*n*	Mean (SD)	*n*	Mean (SD)
Ultrasound one site	272	0.044 (0.499)	274	0.007 (0.085)	273	0.029 (0.189)	272	0.007 (0.086)
TIPS procedure	272	0.004 (0.061)	274	0.000 (0.000)	273	0.000 (0.000)	272	0.000 (0.000)
Chest drain	272	0.004 (0.061)	274	0.000 (0.000)	273	0.000 (0.000)	272	0.000 (0.000)

A, prednisolone; B, PTX, CT, computerised tomography; ECG, electrocardiogram; MRI, magnetic resonance imaging; SD, standard deviation; TIPS, transjugular intrahepatic portosystemic shunt.

a How to interpret the table: one person having a stool test would equal a probability between 0.018 (OO), 0.018 (AO), 0.022 (OB) and 0.011(AB), dependent on treatment arms.

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library