Glucocorticoids are ineffective in alcoholic hepatitis: A meta-analysis adjusting for confounding variables

E Christensen, C Glud

Abstract
The aim of this study was to perform a meta-analysis of controlled clinical trials of glucocorticoid treatment in clinical alcoholic hepatitis, adjusting for prognostic variables and their possible interaction with therapy, because these trials have given appreciably different results. Weighted logistic regression analysis was applied using the summarised descriptive data (for example, % with encephalopathy, mean bilirubin value) of the treatment and control groups of 12 controlled trials that gave this information. Despite evidence of publication bias favouring glucocorticoid treatment, its overall effect on mortality was not statistically significant (p=0.20) – the relative risk (steroid/control) was 0.78 (95% confidence intervals 0.51, 1.18). There was indication of interaction between glucocorticoid therapy and gender, but not encephalopathy. Thus, the effect of glucocorticoid treatment may be different (beneficial or harmful) in special patient subgroups. These results do not support the routine use of glucocorticoids in patients with alcoholic hepatitis, including those with encephalopathy. Whether other subgroups may benefit needs further investigation using the individual patient data from the published trials and testing in new randomised trials.

(Gut 1995; 37: 113–118)

Keywords: glucocorticoids, alcoholic hepatitis, meta-analysis.

Clinical alcoholic hepatitis is a serious disease in which no treatment has been established as being clearly effective.1 Many trials evaluating the effect of glucocorticosteroid treatment on short term mortality have been performed, but results have been appreciably different ranging from a pronounced beneficial effect to indications of a harmful effect.2-14 The two latest studies in patients with severe disease13,14 suggest a significantly beneficial effect of glucocorticosteroid. Four meta-analyses15-17 also report a beneficial therapeutic effect especially in patients with encephalopathy.1,15 The published pooled results of the meta-analyses are not valid, however, because the trial results are not homogeneous as required by the methods applied. In fact the heterogeneous trial results call for a comprehensive analysis addressing the following questions:

(1) Are the varying trial results caused by an imbalance between the treatment and control groups with regard to the patient characteristics which – independent of treatment – are being associated with mortality (prognostic variables)?18

(2) Does the magnitude of the therapeutic effect depend on the type of patients included in the trial or equivalently, are the variables describing the patients associated with the magnitude of the therapeutic effect (so called variable-therapy interaction)?19,20

Of these questions, the first is particularly relevant: the risk of an imbalance between the treatment and control groups in respect of patient characteristics that influences the outcome is considerable in small trials21 like many of those performed in this area. In addition, publication bias may have distorted the picture.

The purpose of this investigation was to perform a meta-analysis which adjusts for the influence of patient characteristics that covary with prognosis and therapeutic effect, and to study the evidence of publication bias.

Methods

SELECTION OF TRIALS

All published, randomised clinical trials which evaluated the short term effect (<3 months) on survival of glucocorticosteroid therapy (T) versus placebo or no active drug (control (C)) in patients with alcoholic hepatitis were considered. The completeness of the trial sample was checked by MEDLINE and by a cross-bibliographic check of the reference lists of the published meta-analyses and the individual trial reports. The trials were included in this study only if they provided a description of patient characteristics (as the mean value or percentage) separately for both treatment groups.3-14 Thus, the study of Helman et al2 which presented only a grading of the overall clinical severity in the two treatment groups, could not contribute to the analysis on the influence of confounding variables on the result. For other analyses, we used the variable information for the total group of patients of that study, when feasible. Each trial was also rated using a quality score estimated according to Chalmers et al22. A few basal characteristics of the included trials and the quality score are presented in Table I.
TABLE I Controlled clinical trials evaluating glucocorticoid treatment in alcoholic hepatitis

<table>
<thead>
<tr>
<th>First author (ref)</th>
<th>Year</th>
<th>Drug*</th>
<th>Initial dose (mg/d)</th>
<th>No of patients</th>
<th>Duration of therapy</th>
<th>Follow up</th>
<th>Quality score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helman et al. 37</td>
<td>1971</td>
<td>Pi</td>
<td>40</td>
<td>37</td>
<td>6 wk</td>
<td>3 mth</td>
<td>52</td>
</tr>
<tr>
<td>Porter et al. 35</td>
<td>1971</td>
<td>MP</td>
<td>20</td>
<td>40</td>
<td>50 d</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td>Campa et al. 36</td>
<td>1973</td>
<td>P</td>
<td>45</td>
<td>45</td>
<td>6 wk</td>
<td>40</td>
<td>37</td>
</tr>
<tr>
<td>Blitzer et al. 37</td>
<td>1977</td>
<td>Pi</td>
<td>28</td>
<td>28</td>
<td>6 d</td>
<td>60</td>
<td>69</td>
</tr>
<tr>
<td>Maddrey et al. 37</td>
<td>1978</td>
<td>Pi</td>
<td>55</td>
<td>55</td>
<td>30 d</td>
<td>2 mth</td>
<td>34</td>
</tr>
<tr>
<td>Lesene et al. 37</td>
<td>1978</td>
<td>Pi</td>
<td>14</td>
<td>14</td>
<td>30 d</td>
<td>2 mth</td>
<td>61</td>
</tr>
<tr>
<td>Shumaker et al. 37</td>
<td>1979</td>
<td>Pi</td>
<td>66</td>
<td>32</td>
<td>4 wk</td>
<td>1 mth(?)</td>
<td>61</td>
</tr>
<tr>
<td>Depew et al. 37</td>
<td>1980</td>
<td>Pi</td>
<td>28</td>
<td>28</td>
<td>6 wk</td>
<td>2 mth</td>
<td>61</td>
</tr>
<tr>
<td>Theodosiou et al. 37</td>
<td>1982</td>
<td>MP</td>
<td>1000</td>
<td>55</td>
<td>3 d</td>
<td>&gt;1 mth</td>
<td>42</td>
</tr>
<tr>
<td>Mendenhall et al. 37</td>
<td>1984</td>
<td>Pi</td>
<td>178</td>
<td>30</td>
<td>3</td>
<td>9 yr</td>
<td>71</td>
</tr>
<tr>
<td>Bories et al. 37</td>
<td>1987</td>
<td>Pi</td>
<td>65</td>
<td>65</td>
<td>1 mth</td>
<td>3 mth</td>
<td>47</td>
</tr>
<tr>
<td>Carithers et al. 37</td>
<td>1989</td>
<td>MP</td>
<td>32</td>
<td>66</td>
<td>4 wk</td>
<td>4</td>
<td>82</td>
</tr>
<tr>
<td>Ramond et al. 37</td>
<td>1992</td>
<td>Pi</td>
<td>40</td>
<td>61</td>
<td>4 wk</td>
<td>2 mth</td>
<td>81</td>
</tr>
</tbody>
</table>

*Pi=prednisolone, MP=methylprednisolone, P=prednisone.
†Estimated according to Chalmers et al. 37
‡The mortality after 2-5 months is used in the present meta-analysis.

EFFECT VARIABLE (END POINT) AND STATISTICAL WEIGHTING

Even though, in most of the trials, glucocorticoid treatment was given for about one month, mortality was evaluated over a period of up to 2-5 months after inclusion (presumably roughly corresponding to the maximum duration of the stay in hospital for these severely ill patients). Because of this, mortality up to 2-5 months after randomisation was used as the effect variable. These data are presented in Table II.

END POINT FOR TRIAL GROUPS: DEATH RISK

The mortality of each treatment and control group was summarised as the log Death Risk (DR)=log(\(r/(n-r+0.5)/n\)), where \(n\) is the number of patients and \(r\) is the number of deaths.23 The constant, 0.5, is a correction allowing for \(r\) being zero or equal to \(n\). Each treatment and control group was given a statistical weight (\(w\)), which was the reciprocal of the variance of the log death risk – that is, \(w=\sqrt{n-r+1}\).22 where \(p=\sqrt{DR/(1+DR)}\), DR being the lower log death risk (including negative values), the lower the mortality. A log death risk of zero corresponds to a 50% mortality, positive values to a mortality of more than 50%, and negative values to a mortality of less than 50%.

ENDPOINT FOR TRIALS: RELATIVE RISK

The magnitude of the effect of treatment in each trial was summarised as the log relative risk (or log odds ratio). This was the difference in the log death risk between the treatment and control groups. The statistical weight of the log relative risk is the reciprocal of the sum of the variances of the log death risks for the treatment and control groups.23 The log death risks, the log relative risks with 95% confidence limits (95% CI),23 and the corresponding weights (indicating the confidence of the estimates) are also presented for each trial in Table II.

To indicate a possible publication bias, the log relative risk was plotted as a function of the statistical weight of the trial.

POOLED RELATIVE RISK

Since the test of heterogeneity23 24 between the trial results (relative risks) shows that these are significantly different (\(p<0.05\)), the simple Woolf25 and Mantel-Haenszel23 tests of significance of the pooled relative risk (or weighted average) are not valid. They both, however, suggest a highly statistically significant effect of glucocorticoid treatment. Using the method suggested by DerSimonian and Laird,25 which allows for heterogeneity (different trial results), the pooled relative risk is 0.57 (95% CI: 0.34, 0.97); without the study of Helman et al.2 it is 0.62 (0.38, 1.05), a much less convincing result. These pooled relative risk estimates are not optimal, however, because they disregard the possible influence of descriptive variables on the outcome.

DESCRIPTIVE VARIABLES STUDIED

The mean age, mean serum bilirubin, mean serum albumin, ascorbic (\%\), male gender (\%), and encephalopathy (\%) were given for the treatment and control groups in at least 11 of the 12 trials included and were studied further. The following variables could not be studied for the reasons indicated: fever (\%\) (missing in five trials), renal insufficiency (\%) (only given in two trials), mean serum creatinine (missing in four trials), number of days in hospital before treatment (missing in three trials), mean leukocytes (missing in three trials), mean serum aspartate aminotransferase (SGOT) (different units, normal range not indicated, recalculated to a common unit not considered feasible), mean prothrombin time (appreciably different units, recalculated to a common unit not considered feasible), mean serum alkaline phosphatases (missing in five trials), mean blood haematocrit/haemoglobin (six gave the mean haematocrit in per cent, four gave the mean haemoglobin concentration, and two gave neither; it was not considered feasible to recalculates to haemoglobin concentration because of missing information about mean corpuscular haemoglobin concentration).
ASSOCIATION BETWEEN END POINT AND DESCRIPTIVE VARIABLES
As summarised in the introduction, the end-point (outcome) may show an association with the following: (1) certain descriptive variables irrespective of treatment (prognostic variables); (2) some combinations of descriptive variables and treatment (variable-therapy interaction). These associations were studied using weighted (logistic) regression analysis, applying the calculated empirical weights given in Table II. This analysis describes the endpoint variable Y as the sum of a constant $b_0$ and the descriptive variables $z_1, ..., z_p$, each multiplied by its corresponding regression coefficient $b_1, ..., b_p$:

$$Y = b_0 + b_1z_1 + \ldots + b_pz_p.$$  

Employing this method, two types of analyses were studied:

1. The association between the log relative risk (T/C) and the ratio of the descriptive variables between the treatment (T) and the control (C) groups using the summarised results for each trial ($n=12$).
2. The association between the log death risk and the level of descriptive variables in each treatment and control group ($n=24$).

Here the pairing of the treatment and control groups of each trial was incorporated into the analysis by including $t-1$ dummy trial indicator variables $-1$ being the number of trials. This provides the best estimate of the adjusted overall therapeutic effect. Variable-therapy interaction was studied by including interaction terms (variable * therapy) in the regression model. Further statistical details are explained in the appendix.

A 5% level for statistical significance was used. Because of the limited number of observations — in this study the number of trials (that is, 12) or trial groups (that is, 24) — the statistical power of the analyses is rather small. Therefore, the analyses can give some indications only, they cannot provide very precise results.

Results
INFLUENCE OF SAMPLE SIZE, TRIAL QUALITY, AND TREATMENT REGIMEN
As shown in Figure 1, all large sample trials, having a statistical weight of 2-5 or higher, showed no therapeutic effect (log relative risk very close to zero), while all the trials that showed a therapeutic effect (noticeably negative log relative risk) were small, having a statistical weight less than 2-5. This is highly suggestive of publication bias.

As shown in Figure 2, there was no association between the magnitude of the therapeutic effect and the quality score.

No association was found between the magnitude of the therapeutic effect and the type of corticosteroid, the daily dose or the duration of therapy.

HETEROGENEITY OF PATIENT SAMPLES AND IMBALANCE BETWEEN TREATMENT AND CONTROL GROUPS
Figure 3 shows the noticeable heterogeneity between the trials and the degree of imbalance (greatest in smallest trials) between treatment and control groups with regard to age, male gender, encephalopathy, ascites, bilirubin, and albumin. There is no obvious simple pattern associated with the therapeutic effect (Fig 3, lower right panel).

INFLUENCE OF IMBALANCE ON THE RESULT
Analysis of trial end points (log relative risk)
The best fitting weighted logistic regression model of the treatment effect (log relative risk (T/C)) described by the degree of imbalance, expressed as the (T/C) ratio of descriptive variables between the treatment and the control groups, is shown in Table III. In this model a high therapeutic effect (low log relative risk (T/C)) shows association with imbalance in the direction of a low (T/C) ratio in bilirubin, percentage of men, and percentage with ascites indicative of a better spontaneous prognosis in the glucocorticoid therapy group and thereby a bias in favour of that group.

Analysis of trial group end point (log death risk)
The result of weighted logistic regression analysis for prediction of the log death risk in each treatment and control group adjusting for the combined influence of a possible imbalance in prognostic variables is shown in Table IV. This model provides the best estimate of the adjusted overall therapeutic effect of glucocorticoid treatment. The estimated overall effect is not significant. The estimated
log death risk (T/C) was −0·25 (95% CI −0·67, 0·16); that is, the relative risk (T/C) was 0·78 (0·51, 1·18). The type 2 error risk of overlooking a relative risk of 0·70 or less is only 20%. When the study of Helman et al. 2 (applying the summarised variables for the total group of patients to both treatment groups) was included, the estimated log death risk (T/C) was −0·31, p=0·15 (95% CI −0·75, 0·13); that is the relative risk (T/C) was 0·73 (0·47, 1·14).

The model showed that poorer prognosis was significantly associated with a high prevalence of encephalopathy and a high bilirubin concentration.

INFLUENCE OF INTERACTION AND IMBALANCE
In analyses of the influence of each descriptive variable adjusting for therapy and the variable-therapy interaction (including trial indicators) only a slight indication of males % – therapy interaction was found – that is, the magnitude of the therapeutic effect tended to decrease with the increasing value of males % (p=0·17). This is illustrated in Fig 4. This interaction was practically the same with and without trial indicator variables included in the analysis. The other variables (including encephalopathy) showed no sign of therapeutic interaction in this analysis (p>0·5).

Discussion
The effectiveness of glucocorticosteroid treatment in alcoholic hepatitis has been debated for many years. Although four meta-analyses 15–17 and a comment 28 have concluded that this treatment is beneficial in alcoholic hepatitis, especially in patients with hepatic encephalopathy, 15 the present study shows that the evidence supporting that conclusion is questionable. One late trial, which
suggested a noticeably beneficial effect of glucocorticoid, considered survival up to four weeks only, omitting subsequent survival.\(^{13}\) If that had been included, it is possible that some late fatal side effects in the glucocorticoid group could have reduced the drug’s beneficial effect on survival. In the most recent positive trial,\(^{14}\) the glucocorticoid group was favoured by cases with somewhat less severe disease at randomisation. Adjustment for the hepatic aspects of this by Cox regression analysis increased the risk of a type 1 error (the \(p\) value) from 0·001 to 0·02.\(^{14}\) The magnitude of the adjusted therapeutic effect and the 95% CIs were not presented. A possible imbalance in non-hepatic aspects of disease severity was not adjusted for. Such an imbalance may also have been present, since four patients in the control group (but none in the treatment group) died from pancreatitis or acute respiratory distress syndrome,\(^{14}\) for which corticosteroid is ineffective.\(^{29}\)

A number of factors may influence the results of a controlled clinical trial. In this report we have investigated the influence of some factors which did not receive much attention in earlier studies. For this purpose we have used weighted (logistic) regression analysis.\(^{23\ 24\ 26\ 27}\) The statistical weights are measured related to the statistical power of the trial, they are a function of the number of patients and the outcome.\(^{23}\)

We found that a therapeutic effect was confined to the trials with the least statistical weight. The trials with the highest statistical weights were all clearly negative. This is a strong argument against glucocorticosteroid therapy having a significant overall effect in alcoholic hepatitis. The finding also indicates some publication bias (positive trials being published more frequently than negative) in this area, especially when the fact that smaller trials have the highest risk of being biased by random factors, including imbalance between the treatment and control groups\(^{21}\) which may highly influence the result, is considered. As reported by Conn\(^ {27}\) one negative small trial has not been fully published and we cannot exclude the possibility that other negative smaller trials have been withheld from publication. A publication bias would imply that the therapeutic effect is even less than that found in the previous report for the hepatic aspects of the trial.\(^ {14}\)

We found no significant association between the quality score of Chalmers et al\(^ {22}\) and the therapeutic effect. This is at variance with other results,\(^ {15}\) but these were based on other quality criteria.

The patient samples included in the various trials were noticeably different with regard to many descriptive variables, and in some cases there was a substantial imbalance in some important prognostic variables (encephalopathy, ascites, bilirubin) between the glucocorticoid and control groups.

After adjustment controlled imbalance by weighted logistic regression analysis, the overall therapeutic effect was found to be far from statistically significant. This negative overall result does not, however, necessarily exclude the possibility of a beneficial therapeutic effect in some special patients and a harmful effect in other patients.

For this reason we performed analyses allowing for interaction between descriptive variables and treatment. These analyses did not show an interaction between encephalopathy and therapy. This agrees with the most recent trial\(^ {14}\) but is in contrast with the claim by some meta-analysts that the therapeutic effect is mainly confined to patients with encephalopathy.\(^ {15}\) Although results differ depending on the method and which trials are included,\(^ {15\ 17\ 28}\) Survival figures for the subgroup of patients with encephalopathy at randomisation were not, however, published in some of the reports. This means that the meta-analysts must have obtained these numbers from other sources (directly from the authors?), although this is not described explicitly in their reports.\(^ {15\ 28}\) Furthermore, it is not possible from the published reports\(^ {14\ 28}\) to test if the somewhat smaller subgroups of patients with encephalopathy were balanced between the treatment and control groups in respect of other prognostic variables. Nevertheless, the reported negative result for the encephalopathic patients in the three largest trials having the greatest statistical weight\(^ {28}\) (Fig 2) suggests that the overall therapeutic effect in this subgroup is also close to zero.
The present analysis suggests some interaction between the percentage of men and the effect of glucocorticoid therapy, indicating a better effect in women. Since the results were based on analyses of average variable values from groups of patients they cannot be extrapolated to the individual patient. Nevertheless, the results should encourage the authors of the largest trials11 to make their data available for more elaborate analyses of therapeutic interactions,18 19 using the full raw data set to confirm or refute the indications provided by the present analysis. This should include a study of the value of the discriminant function of Maddrey et al16 in defining good responders. This could not be studied in the present study. Unfortunately, the Copenhagen trial20 30 does not contain a sufficient number of these very severely ill patients to allow a useful analysis of that kind.

Thus, in contrast to the previous meta-analyses15 17 26 and results,11 we conclude that the overall effect of glucocorticoid treatment in patients with chronic alcoholic hepatitis is not statistically significant. This conclusion may not be totally without practical consequences as about 68% of European specialists seem to offer this treatment to patients with alcoholic hepatitis.31 Nevertheless, a beneficial effect (or harmful effect) cannot be excluded in some subgroups. To solve these problems, new analyses of the individual patient data from the largest of the performed trials and testing of the resultant hypotheses in new randomised trials will be necessary.

Appendix

Variable-therapy interaction is studied by including interaction terms (variable * therapy) in the regression model. That is, the log death risk (Y), can be described as a function of the therapy (Zagg 1 for glucocorticoid, 0 for control) and the variable(s) characterising the group (Zaggmean) (mean or %) and the variable-therapy interaction (Zagg*Zaggmean). For one descriptive variable this can be expressed as follows:

\[ Y = b_0 + b_1Zagg + b_2Zaggmean + b_3Zagg*Zaggmean + b_4 + 0 \]

where \( b_4 \) is a constant term.

For illustrative purposes the weighted simple regression lines of the log death risk as a function of the given descriptive variable can also be estimated separately for the treatment and control groups:

\[ Y = b_0 + b_1Zagg + b_2 + 0 \] (II)
\[ Y = b_0 + b_1Zagg + b_2 + 0 \] (III)

Given the defined scoring of therapy and control, the terms in equation (I) can be obtained from the terms in equations (II) and (III) as follows:

\[ b_0 = b_0 + b_2; b_1 = b_1; b_2 = b_2 + b_2; b_3 = b_3 + b_3; b_4 = b_4 \]

Thus, a significant \( b_3 \) in equation (I) is equivalent to a significant difference between \( b_3agg \) and \( b_3agg \) (equations (II) and (III)) and suggests interaction. That is, the effect of treatment depends on the variable in question, which may then be termed "therapeutic."18 19 A significant value of \( b_3agg \) suggests that the variable in question has a prognostic influence.18 19

Glucocorticoids are ineffective in alcoholic hepatitis: a meta-analysis adjusting for confounding variables.
E Christensen and C Gluud

_Gut_ 1995 37: 113-118
doi: 10.1136/gut.37.1.113

Updated information and services can be found at:
http://gut.bmj.com/content/37/1/113

These include:

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections

Hepatitis other (155)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/