An Intelligent Aide for Interpreting a Patient's Dialysis Data Set

Derek Sleeman[1], Nick Fluck[2], Elias Gyftodimos, Laura Moss, and Gordon Christie

Departments of Computing Science & Medicine,
The University of Aberdeen
Aberdeen AB24 3FX
Scotland UK

Abstract. Many machines used in the modern hospital settings offer real time physiological monitoring. Haemodialysis machines combine a therapeutic treatment system integrated with sophisticated monitoring equipment. A large array of parameters can be collected including cardiovascular measures such as heart rate and blood pressure together with treatment related data including relative blood volume, ultrafiltration rate and small molecule clearance. A small subset of this information is used by clinicians to monitor treatment and plan therapeutic strategies but it is not usually analysed in any detail. The focus of this paper is the analysis of data collected over a number of treatment sessions with a view to predicting patient physiological behaviour whilst on dialysis and correlating this with clinical characteristics of individual patients.

One of the commonest complications experienced by patients on dialysis is symptomatic hypotension. We have taken real time treatment data and outline a program of work which attempts to predict when hypotension is likely to occur, and which patients might be particularly prone to haemodynamic instability. This initial study has investigated: the rate of change of blood pressure versus rate of change of heart rate, rate of fluid removal, and rate of uraemic toxin clearance. We have used a variety of machine learning techniques (including hierarchical clustering, and Bayesian Network analysis algorithms). We have been able to detect from this dataset, 3 distinct groups which appear to be clinically meaningful. Furthermore we have investigated whether it is possible to predict changes in blood pressure in terms of other parameters with some encouraging results that merit further study.

1 Introduction

The human renal system is responsible for a number of different roles, including: Water Balance, Electrolyte Balance (e.g. sodium, potassium), Toxin Removal, Acid/Base Balance, and Blood Pressure Regulation, [Daugirdas et al, 2001]. Patients with

[1] Corresponding author: Derek Sleeman, dsleeman@csd.abdn.ac.uk www.csd.abdn.ac.uk/ research/~sleeman
[2] Renal Medicine Section, Aberdeen Royal Infirmary & School of Medicine, The University of Aberdeen.

R. Bellazzi, A. Abu-Hanna, and J. Hunter (Eds.): AIME 2007, LNAI 4594, pp. 57–66, 2007.

advanced renal failure where overall excretory renal function is below 15% often require renal replacement therapy. The most common form of treatment in the UK is unit based haemodialysis. Most patients receive standard haemodialysis 3 times a week for between 4-5 hours per session. A haemodialysis machine pumps blood from the patient through a high surface area semi-permeable dialysis membrane before returning the blood back to the patient. Fluid is removed through a process of ultrafiltration due to the pressure differential across the membrane and toxins are removed by diffusion across the membrane into a controlled dialysis solution. A range of treatment refinements have been developed that try to improve patient stability and therapeutic efficacy. They include the use of ultrafiltration alone or in combination with dialysis and dialysis treatment profiling where fluid removal or diasylate sodium concentration is dynamically changed during the treatment session. The entire dialysis process requires careful regulation and monitoring and is handled by an on board computer. Internal monitoring equipment records both machine parameters as well as patient physiological variables. Real time data is often used by clinicians to assess current physiological status but long term analyses of the whole treatment data-sets are unusual. We have used a combination of statistical, data mining [Hand et al, 2001], and theory refinement [Craw & Sleeman, 1990] approaches to address a range of clinical issues such as:

- Can major intra-dialytic complications such as hypotension be predicted?
- Are there distinct groups of patients with specific physiological behaviour characteristics?
- What are the major differences between patients who are stable on dialysis and those that are not?
- Is it possible to create patient specific optimum dialysis strategies [Sleeman et al., 2004].

2 Methods

Three dialysis machines (AK200 Ultra S, Gambro) installed at a hospital based satellite unit (Dr Gray's hospital, Elgin, Grampian, Scotland) have been fitted with a data collection interface node wirelessly linked to a server running data collection software (Exalis, Gambro). Data was routinely collected at each dialysis session delivered by the enabled machines. With this configuration data can be collected on a maximum of 12 patients per week of treatment with storage of approximately 144 hours of real time data. Initial pilot investigations have taken 288 hours of complete data sets for analysis. Collected parameters are detailed in Table 2.1.

The initial suite of programs in this series was implemented in the Neurological ICU at Western General Hospital in Edinburgh to monitor patients who have had traumatic head injuries, [Howells, 1994; McQuatt et al, 1999]. They implemented a monitor system which allows clinicians to review real-time data sets, as well as earlier data, for a number of important parameters on a monitor at the patient's bedside. Additionally, they implemented a BROWSER system for use by clinicians and data analysts to view the same data sets in an off-line mode. Further, both of these systems allow the clinician/administrator to define a series of insult levels for each of the

Table 2.1. Parameters collected during dialysis sessions and their frequency of collection

Parameter Name	Frequency Parameter Collected
Heart Rate	30mins
Systolic Pressure	30mins
Diastolic Pressure	30mins
Actual Weight Loss Rate	1min
Blood Volume	1min
Plasma Conductivity (AK)	30mins
Total Blood	1min
Actual Time	1min
NA Concentration	Once per session
KT/V	1min
Ionic Effective Dialysance	1min
Actual Total Weight Loss	1min

"channels"; for instance in the case of the NICU data sets the colours are normal – white, slight raised – yellow, considerably raised – orange, extreme value – red.

We, at Aberdeen, have extended the system to enable it to deal with not just head injury data sets but with a range of data sets including: dialysis data sets from Pavia (Italy), Aberdeen & Elgin, and an ITU data set. Further types of data sets can be added relatively easily and usually require a further specialized input routine. We took the opportunity to recode the system in Java – hence the new name, JAB (Java version of the Aberdeen Browser). Additionally the user is able to decide which of a range of parameters he/she wishes to have displayed at any time; the system also gives the user the chance to choose between several display formats. Further, the ability to define insult levels has been extended. This implementation thus reads initially details of the parameters collected, followed by the appropriate data set. Figure 2.1 shows the UI for this system displaying part of the Elgin / Aberdeen data set.

3 Results

In the introduction we outlined a number of top-level questions which we hope to address using data sets collected from dialysis patients. Our pilot study aimed to address two specific questions:

a. Is it possible to identify distinct clinical subgroups of patients from the dialysis data alone?
b. Can one predict the blood pressure or change in blood pressure during a dialysis session using the other collected variables?

We selected 72 sessions with complete data from 9 patients (i.e. 8 sessions per patient). This included a range of patients with differing clinical and demographic features. Some were recognised as stable patients whilst others were known to be more challenging to treat effectively.

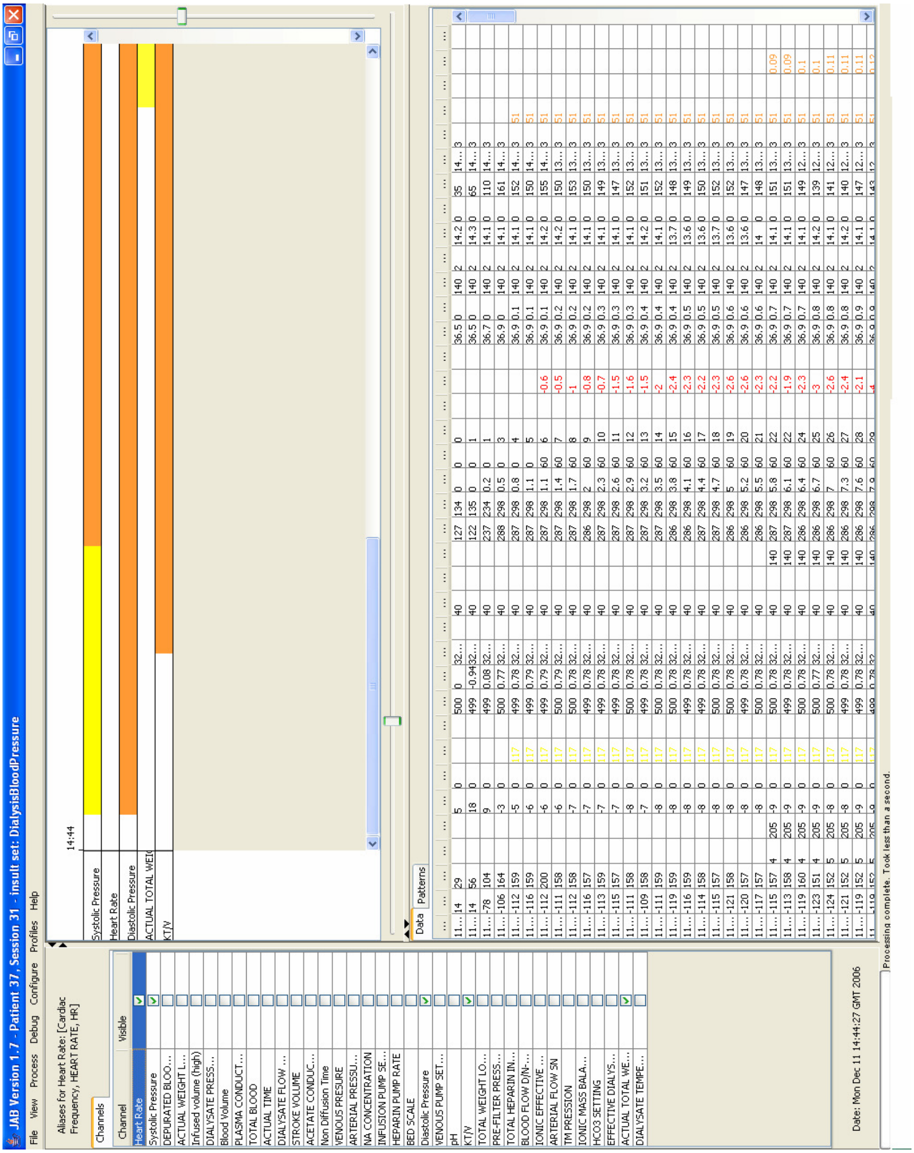

Fig. 2.1. JAB's UI displaying a portion of a dialysis patient's data set

3.1 Determining Distinct Patient Groups in the Data Set: Clustering Analysis and Sub-group Discovery

We have performed an analysis on the group of 9 patients using several clustering algorithms. We used two different methods, namely hierarchical clustering using Ward's method with Euclidean distance [Ward 1963] and k-means clustering (with

Table 3.1. Allocation of patient data records to the three main clusters using hierarchical clustering. Each row contains the patient ID, number of records (percentage in brackets) in each cluster and total number of records for that patient.

Patient	LHS cluster	Centre cluster	RHS cluster	Total
20	48(80%)	0(0%)	12(20%)	60
23	1(2%)	9(19%)	38(79%)	48
25	0(0%)	52(93%)	4(7%)	56
26	0(0%)	59(97%)	2(3%)	61
28	3(5%)	5(9%)	48(86%)	56
30	0(0%)	56(87.5%)	8(12.5%)	64
33	53(95%)	0(0%)	3(5%)	56
35	0(0%)	36(64%)	20(36%)	56
37	0(0%)	0(0%)	56(100%)	56

varying number of clusters), [MacQueen 1967]. Each instance was described by a vector of 10 real-valued attributes: Systolic pressure, diastolic pressure, heart rate, blood volume, and the absolute changes of these four attributes in the previous 30 minute interval, as well as rate of fluid removal (weight loss) and rate of toxin removal (KT/v). So for a patient undergoing a 4-hour dialysis session, this procedure produces 7 such vectors: one at the end of the second 30 minute period when the values are compared with those at the end of the first 30 minute period, one at the end of third 30 minute period when the values are compared with those at the end of the second 30 minute period, etc.

The three clusters of the dendrogram produced by the hierarchical clustering algorithm are shown in Figure 3.1. Labels have the form *pp:ss* where *pp* is patient ID and *ss* is session number. For six patients (ID numbers 20, 28, 30, 33, 35, 37), it can be seen that a high percentage of their instances fall within a single cluster. The algorithm has identified three main sub-clusters: (a) the left-hand side (LHS) cluster in the dendrogram, dominated by records from patients 20 and 33, (b) the centre cluster dominated by patients 25, 26, 30 and 35, and (c) the right-hand side (RHS) cluster dominated by patients 23, 28 and 37. The allocation of patients to these three clusters is shown in more detail in Table 3.1. Comparing these clusters with the demographic data, we see that the LHS cluster contains patients with an average age of 26, compared to 66 in the rest of the dataset. Within the RHS cluster, the two most dominant patients, who are also clustered more compactly, namely patients 28 and 37, are the ones who suffer both from diabetes and cardio-vascular disease. And the third cluster corresponds to the remaining patients.

K-means clustering was performed with two, three, and nine clusters. The advantage of the k-means method is that it gives us a comparative quantitative evaluation of the clusters discovered with hierarchical clustering. The disadvantage is that we have to manually experiment with different numbers of clusters. The clustering was evaluated by matching clusters to the patient IDs, patient age (discretised in two groups), occurrence of diabetes, occurrence of cardio-vascular disease (CVD), and occurrence of both diabetes and CVD. The error rates (incorrectly clustered instances divided by

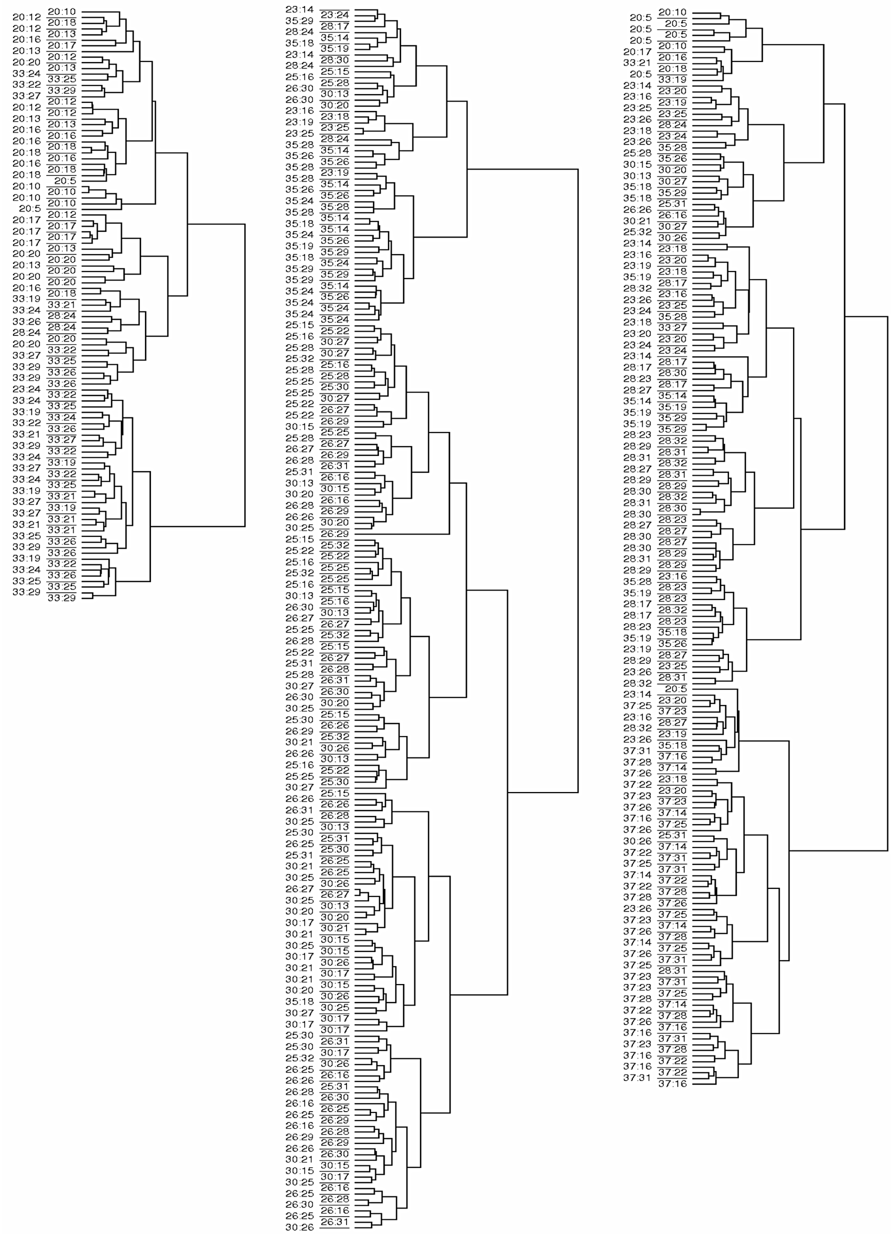

Fig. 3.1. The three main clusters derived using hierarchical clustering in detail. LHS cluster is dominated by patients 20 and 33, and the RHS cluster by patients 28 and 37.

Table 3.2. Error rates for different classes and numbers of clusters

# of clusters	Class	Error
9	Patient ID	27.7%
2	Age	2.9%
2	Cardio-vascular	33.7%
2	Diabetes	41.1%
3	CVD and Diabetes	19.5%

total number of instances) are summarised in Table 3.2. We see that for k=9 the different patients are clustered fairly well (28% error is not very high for a 9-class problem); binary clustering (for k=2) essentially separates the two age groups; finally, while CVD and diabetes independently do not form significant sub-groups, their conjunction (when k = 3) stands out as an important cluster.

Conclusions: Both types of clustering identify essentially the same 3 clusters, which is encouraging. Further, the clinician, using his considerable experience, had selected 3 distinct groups of patients (see the introductory paragraph of this section), and in fact he has confirmed that the 3 clusters identified by the algorithm correspond exactly to those 3 clusters. These 3 clusters are clinically significant as they would be expected to react to dialysis in distinctly different ways; in particular, the third group (diabetes and cardio-vascular disease) are likely to present more challenges to dialyse them successfully. So clinically it is likely that the latter group would be monitored more closely during dialysis. These results are encouraging because it appears to be possible for a machine learning algorithm on the basis of a small amount of demographic data and the information collected during the dialyses sessions, to identify these clinically significant groups. This also suggests that other patients who perhaps can not be clearly identified by their existing medical history, can be detected by the algorithm on the basis of their dialysis data sets, as patients who are likely to be "unstable" / complex patients to dialyse. Further, this suggests that more sophisticated analyses should be applied to the data sets produced by this group of "complex" patients, probing for example for early signs of instability etc.

3.2 Determining If It Is Possible to Predict Changes in the Patient's BP in Terms of Other Parameters

The aim of this analysis is to predict the occurrence of hypotension during a dialysis session. We were particularly interested in investigating for each patient whether the following are correlated:

1. Rate of change of BP and Rate of change of heart rate
2. Rate of change of BP and rate of Fluid removal
3. Rate of change of BP and rate of toxin removal

We consider a hypo-tensive event to be a drop greater than or equal to 10 mmHg in systolic pressure between two successive measurements (at 30 minute intervals). We have considered a dataset obtained from the same 9 patients, for each of which eight complete sessions of approximately four hours were available. Blood pressure (systolic and diastolic) and heart rate were measured at 30 minute intervals during these sessions. Toxin removal and fluid removal rates were available, and these were approximately constant within each individual session. Blood volume was measured every minute; we have converted that to 30 minute averages during pre-processing. The demographic data was also included in the analysis. We have incorporated some of the time-series information by creating the additional features of the changes in heart rate, blood volume, systolic and diastolic pressure since the previous measurement, and appending these to the data vector of each instance. The target attribute in each of these instances was derived from the change in systolic pressure in the next half hour. Each dialysis session yielded about 8 measurements; of these the first and the last were used to derive the changes of values compared to the second and last but one, respectively. We therefore had roughly 6 different instances labelled as hypotensive or non-hypotensive for each dialysis session. In fact since a few of the sessions were slightly longer, we obtained a total of 442 instances, of which 120 were labelled positive (hypotension event occurring) and 322 negative.

Table 3.3. Confusion matrix for hypotension prediction

	Predicted positives	Predicted negatives	Total
Actual positives	32	88	120
Actual negatives	11	311	322
Total	43	399	442

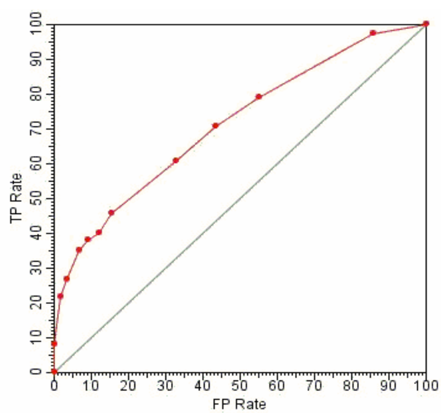

Fig. 3.2. ROC curve for hypotension prediction

Experiments were run using the WEKA data mining software tool [Witten and Frank 2005]. We used a Bayesian network classifier combined with a hill-climbing algorithm for structure learning, and evaluated the classifier using 10-fold cross validation. The ROC curve obtained by the algorithm (shown in Figure 3.2) was analysed to determine the optimal decision threshold for the Bayesian classifier. The confusion matrix for the optimal accuracy point is given in Table 3.3. The maximum accuracy reached was 77.6% (32+311/442), with a true positive rate of 26.7% (32/120) and false positive rate of 3.4% (11/322). This is an encouraging result, since predicting blood pressure trends over a half-hour period is a hard task; the algorithm predicts correctly more than a quarter of the hypo-tensive events, with only 11 false positive cases.

4 Conclusions and Further Work

These pilot studies have indicated that it may be possible to relate a patient's physiological behaviour on dialysis with their underlying pathological status. This offers an intriguing possibility that further analysis may allow more precise characterisation of patients and enable clinicians to tailor therapy more appropriately. Planned further work includes:

- Using the clustering approaches on a much wider range of patients to see if it is still possible for the algorithms to identify a number of clinically significant groups (e.g. those that are "unstable" / complex under dialysis)
- Investigating whether it is possible to improve the blood pressure predictions. As noted in section 3.2 some of the false positives (i.e. avoidance of hypotension) could well be due to nursing interventions (e.g. infusing of saline fluid). So in future we aim for better access to the patients' nursing notes.
- Pursuing some of the other issue list in section 1; addressing this agenda should enable us to progressively customize dialysis sessions to individual patients.
- Collecting opinions from a group of renal experts on a range of dialysis sessions, compare their analyses, and then hold face-to-face session(s) where these differences are discussed / resolved.
- Evaluating these Aides with a number of renal physicians and perhaps nurse practitioners.

Acknowledgements

- Ricardo & Roberto Bellazzi & colleagues (Pavia University & the Northern Italian Renal/ Dialysis Centre) gave us invaluable help in exporting the data collected by the EXALIS system in forms which could be used by our several analysis tools (e.g. the JAB system)
- This work has been supported under the EPSRC's grant number GR/N15764 and the Advanced Knowledge Technologies Interdisciplinary Research Collaboration, which comprises the Universities of Aberdeen, Edinburgh, Sheffield, Southampton and the Open University.
- The staff and patients of the Dialysis Unit, Dr Gray's Hospital, Elgin, Grampian.
- Bill Mackie & colleagues (Technical Staff, The Dialysis Unit, Aberdeen Royal Infirmary).

References

1. Craw, S., Sleeman, D.: Automating the Refinement of Knowledge-Based Systems. In: Aiello, L. (ed.) Proceedings of ECCAI-90, pp. 167–172. Pitman, London (1990)
2. Daugirdas, J.T., Blake, P.G., Ing, T.S.: Handbook of Dialysis, 3rd edn. Lippincott, Williams, & Wilkins, Baltimore (2001)
3. Hand, D.J., Mannila, H., Smyth, P.: Principles of Data Mining. MIT Press, Cambridge (2001)
4. Howells, T.P.: Edinburgh Monitor-Browser© Software, [v1.0.0] Computer Program (1994) tim.howells@nc.uas.lul.se
5. McQuatt, A., Andrews, P.J.D., Sleeman, D., Corruble, V., Jones, P.A.: The analyses of Head Injury data using Decision Tree techniques. Artificial Intelligence in Medicine. In: Horn, W., et al. (eds.) Proceedings of AIMDM'99 Conference, Aalborg, Denmark, June 1999, pp. 336–345. Springer, Heidelberg (1999)
6. MacQueen, J.B.: Some Methods for classification and Analysis of Multivariate Observations. In: Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, vol. 1, pp. 281–297. University of California Press (1967)
7. Sleeman, D., Luo, Z., Christie, G., Coghill, G.: Analysing Time Series Medical Data-sets. In: Proceedings of Knowledge Based Systems & Services for Health Care, Bonn, May 2004, p. 1–4 (2004)
8. Ward, J.H.: Hierarchical Grouping to optimize an objective function. Journal of American Statistical Association 58(301), 236–244 (1963)
9. Witten, I.H., Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, San Francisco (2005)