# Variations Between Countries in Outcome After Stroke in the International Stroke Trial (IST)

## NU Weir, MBChB[*]; PAG Sandercock, DM[†]; SC Lewis, PhD[†]; DF Signorini, PhD[†]; CP Warlow, MD[†] on behalf of the IST Collaborative Group

[*]**Department of Neurology, Royal Hallamshire Hospital, Sheffield, UK**
[†]**Department of Clinical Neurosciences, Western General Hospital, Edinburgh, UK**

Whether a patient who has had an acute stroke dies, survives in a disabled state or recovers completely is determined by many factors. Outcome after stroke varies from country to country and recent studies have sought to explain the variation (1-5). Some of the variation in outcome may be due to differences between countries in the process of stroke care (3;4;6-10). The International Stroke Trial (IST) was a multinational randomized controlled trial that investigated the impact of aspirin, heparin, both or neither on the outcome of patients admitted to hospital with an acute ischemic stroke. Preliminary analyses (with all treatment groups combined) of the nearly 20,000 patients entered revealed that among the 37 participating countries, six month case fatality ranged from 0 to 33% and the proportion dead or dependent ranged from 32 to 89% (11). In itself, this observation was perhaps not surprising, given the inevitable differences between countries in the baseline characteristics (casemix) of the patients entered. The IST employed a randomization procedure which anticipated this possibility and treatment allocation was prospectively balanced *within* countries to prevent any between country differences in casemix from biasing the assessment of the effects of treatment. In this paper, we describe the large variations in outcome between countries in the IST in detail and use the trial data set to determine how far these differences might be explained by adjusting for important differences in casemix and consider how far other factors, in particular international differences in the treatment of acute stroke, might account for any variation in outcome that remains.

**METHODS**
**Exclusion of data from countries recruiting less than 500 patients**
   When comparing patients from different countries we wished to avoid the extreme variability in outcome that can occur with small sample sizes (12). We therefore arbitrarily chose to examine data from only those countries that had randomized more than 500 patients each. We combined data from the treatment and control patients in each country for all analyses and our findings therefore do not describe variation between countries in the efficacy of aspirin or heparin in patients with acute stroke.

**Baseline characteristics and measures of the process of stroke care**
   We extracted the following potential prognostic variables from the baseline dataset: age, gender, delay from symptom onset to randomization, conscious level, presence of atrial fibrillation (AF), systolic blood pressure on admission (sBP), presence of any infarction on CT scan, and presence or absence of eight different neurological deficits. We also extracted data on two potential markers of the quality of the process of care: whether patients had had a CT scan prior to randomization and whether patients who were discharged alive from the participating hospital were reported to be prescribed long-term aspirin.

**Crude and adjusted outcome at six months**
   We calculated the proportion who died from any cause (case fatality) and the proportion who were dead or needed help from another person for activities of daily living (death or dependency) at six months in each country (observed number). We used two previously described logistic regression models to adjust

these data for important differences in casemix (13). The prognostic models were constructed and validated on two separate subsets of the IST data set and take the following variables into account: age, sex, sBP, AF, conscious level and the total number of neurological deficits (see statistical appendix). We used the models to predict the probability of an outcome event for each patient and then calculated the total number of predicted events for each country as the sum of these probabilities (predicted number).

We expressed the adjusted outcome as a 'w-score', a method that measures the difference in the number of observed and predicted events per 1000 patients treated within each country (14;15). We calculated the w-score using the formula $w = 1000(o - p)/t$, where o is the observed number of events, p the predicted number of events and t the total number of patients per country. For example, if 500 patients are treated and a total of 100 deaths are predicted but 150 deaths are observed, then the w-score is 1000 x (150-100)/500 = +100, that is, 100 more deaths than predicted per 1000 patients treated. However, we were principally interested in the difference in adjusted outcome between countries and calculated these simply by subtracting w-scores. For example, if the w-score for case fatality in country A was +30 and that in country B was -60, we estimated that the absolute difference in adjusted case fatality between countries to be 90 deaths per 1000 patients treated. In order to directly compare outcomes before and after adjusting for casemix, we also calculated crude w-scores by defining p in each country as $tp_o$, where po is the proportion of patients that experienced the outcome in the entire study population (all countries combined).

### Correction for multiple comparisons

Because the study necessarily involved multiple comparisons between countries, we calculated 99% confidence intervals for the w-scores (according to Parry et al).(14) to reduce the possibility that any observed differences might be due to the play of chance.

### Relationship between measures of the process of care and adjusted outcome after stroke

We ranked the countries by the two adjusted outcomes and by the two measured items of the process of stroke care. Better outcome and better process of care are represented by lower numbers and vice-versa. We investigated the relationship between rankings with Spearman's rank correlation coefficient.

### The predictive properties of the prognostic models

We determined the predictive properties of our prognostic models by estimating their calibration and discrimination (16). Calibration refers to the degree of bias of model predictions for groups of patients and can be estimated by plotting a calibration curve. To derive the calibration curves, we ordered the data-set by ascending predictions of risk and then divided it into 10 equal sized groups (deciles). For each decile, we plotted the mean observed risk against the mean predicted risk. A model is well calibrated if, within each decile, the proportion of patients predicted to have an event and the proportion observed to have done so is the same i.e. if the calibration curve follows a 45° line. Discrimination refers to the ability of the model to differentiate between individuals that do and do not experience an event and may be estimated by calculating the area under the Receiver Operator Characteristic (ROC) curve (a plot of the sensitivity against (1 - specificity) of the model predictions). Thus, a model that predicts case fatality with an area under the ROC of, say, 80% will, in 80% of cases, correctly assign a higher risk of death to a randomly selected patient with a fatal outcome than to a randomly selected survivor.

Analyses were performed using SAS (version 6.12).

### RESULTS

### Patients included in the analysis

The IST recruited 19,435 patients from 37 countries. Nine countries recruited more than 500 patients: Argentina, Australia, Italy, the Netherlands, Norway, Poland, Sweden, Switzerland and the United Kingdom. These countries contributed a total of 15,116 patients (78% of the trial accrual); of these only 59 (0.4%) were lost to six month follow-up. The greatest losses to follow up occurred in Argentina and the

Netherlands (5% and 2%, respectively); in all other countries outcome data were available in over 99% of cases.

**Baseline characteristics and the process of care**

The proportion of patients with each prognostic variable at baseline varied highly significantly between countries (c2 tests, all p < 0.0001) (table 1), confirming our expectation that there would be marked variation between countries in the types of stroke patients considered eligible for the trial. Table 1 also shows substantial variation between countries in the process of stroke care, as judged by the proportion of patients who had a CT scan before randomization (range 42 to 98%) and the proportion of patients prescribed aspirin for long term secondary prevention at hospital discharge (range 53 to 73%).

**Table 1. Characteristics at baseline of the 15,116 patients randomized in the IST from the nine countries which recruited more than 500 patients, and their process of care in hospital.**

| Country | Baseline characteristics n (%) | | | | | | |
|---|---|---|---|---|---|---|---|
| | *n* | *Male* | *Age 70 years plus* | *Systolic BP = 160mmHg* | *Fully conscious* | *Atrial fibrill -ation\** | *= 4 Neuro -logical deficits* |
| AR | 545 | 310 (60) | 267 (51) | 280 (54) | 411 (79) | 58 (11) | 234 (45) |
| AU | 597 | 355 (60) | 352 (59) | 288 (48) | 438 (74) | 67 (12) | 296 (50) |
| IT | 3437 | 1932 (56) | 2190 (64) | 1891 (55) | 2747 (80) | 598 (19) | 1598 (46) |
| NL | 728 | 371 (52) | 394 (55) | 455 (64) | 574 (81) | 73 (10) | 361 (51) |
| NO | 526 | 281 (53) | 352 (67) | 316 (60) | 425 (81) | 33 (6) | 285 (54) |
| PL | 759 | 385 (51) | 382 (51) | 432 (57) | 540 (72) | 189 (25) | 425 (56) |
| SE | 636 | 344 (55) | 473 (75) | 407 (65) | 532 (84) | 95 (15) | 260 (41) |
| CH | 1631 | 842 (52) | 1130 (69) | 971 (60) | 1129 (69) | 263 (16) | 988 (61) |
| UK | 6257 | 3223 (52) | 4229 (68) | 3192 (51) | 4610 (74) | 1202 (21) | 3594 (57) |
| Total | 15116 | 8043 (53) | 9769 (65) | 8232 (55) | 11406 (76) | 2578 (18) | 8041 (53) |

| Country | Process of care n (%) | | | | |
|---|---|---|---|---|---|
| | *n* | *Entered within 24hrs of onset* | *Visible infarct on CT scan* Ý | *CT before random -ization* | *Dis- charge on aspirin* ý |
| AR | 545 | 235 (45) | 165 (18) | 399 (77) | 338 (68) |
| AU | 597 | 368 (62) | 232 (45) | 516 (87) | 377 (65) |
| IT | 3437 | 2242 (65) | 1045 (43) | 2421 (70) | 2104 (62) |
| NL | 728 | 505 (71) | 224 (36) | 615 (86) | 379 (53) |
| NO | 526 | 235 (45) | 222 (45) | 492 (94) | 379 (73) |
| PL | 759 | 474 (63) | 213 (56) | 383 (51) | 432 (58) |
| SE | 636 | 276 (44) | 288 (47) | 617 (98) | 427 (69) |
| CH | 1631 | 1119 (69) | 540 (45) | 1207 (74) | 908 (57) |
| UK | 6257 | 3380 (54) | 1681 (64) | 2607 (42) | 3287 (54) |
| Total | 15116 | 8834 (59) | 4610 (50) | 9257 (61) | 8631 (58) |

AR: Argentina, AU: Australia, IT: Italy, NL: Netherlands, NO: Norway, PL: Poland, SE: Sweden, CH: Switzerland, UK: UK

\* The presence of atrial fibrillation was not recorded in the 822 patients randomized from those countries in the pilot phase of the trial.

Ý Patients with a visible infarct are expressed as a percentage of those with a pre-randomization CT scan.

ý Expressed as a percentage of cases that were discharged alive. Data describing discharge on long term aspirin not collected in 292 cases.

**Variations in outcome at six months (observed and predicted)**

Table 2 shows the variation between countries in the proportion of patients who had died by six months (range 12 to 30%) and in the proportion who were dead or dependent at six months (range 42 to 79%). The proportion of patients predicted dead and the proportion predicted dead or dependent can be taken to summarize the casemix of each cohort, and not surprisingly, these also varied considerably between countries (table 2).

**Table 2. Variation in the observed and predicted number of patients who were dead, or 'dead or dependent', at six months between countries.**

| Country | Patients dead at six months | | Patients dead or dependent at six months | |
|---|---|---|---|---|
| | Observed n (%) | Predicted * n (%) | Observed n (%) | Predicted * n (%) |
| Argentina | 117 (23) | 95 (18) | 247 (48) | 297 (57) |
| Australia | 86 (15) | 133 (23) | 331 (56) | 364 (61) |
| Italy | 664 (20) | 744 (22) | 1969 (57) | 2093 (61) |
| Netherlands | 128 (18) | 141 (20) | 383 (54) | 422 (59) |
| Norway | 69 (13) | 114 (22) | 317 (60) | 332 (63) |
| Poland | 223 (30) | 172 (23) | 403 (53) | 480 (64) |
| Sweden | 78 (12) | 136 (22) | 263 (42) | 381 (61) |
| Switzerland | 366 (22) | 464 (28) | 1091 (67) | 1128 (69) |
| UK | 1711 (27) | 1643 (26) | 4933 (79) | 4166 (67) |
| Total | 3442 (23) | 3641 (24) | 9937 (66) | 9663 (64) |

\* We derived the total number of predicted events per country by summing the individual predicted probabilities generated by our prognostic models.

**Crude and adjusted case fatality at six months**

Figure 1 shows the w-scores for case fatality for each country both before and after adjusting for casemix. The figure clearly illustrates the substantial variation in crude case fatality between countries, the most extreme difference being between Sweden and Poland (171 more deaths per 1000 patients treated in Poland). After adjusting for casemix, some of the differences in case fatality between pairs of countries altered considerably. For example, adjustment for casemix reduced the differences in case fatality between Switzerland and Sweden and between the UK and the Netherlands by about 70%. However, the most striking finding shown in figure 1 is that, on the whole, adjusting for important differences in casemix had little influence on the variation in case fatality between countries, which remained very substantial. For example, at six months, the difference between Sweden and Italy was 69 deaths per 1000 patients and between the UK and Poland 57 deaths per 1000 patients; at its most extreme, 160 more patients were dead at six months per 1000 treated in Poland than in Sweden despite adjusting for casemix.
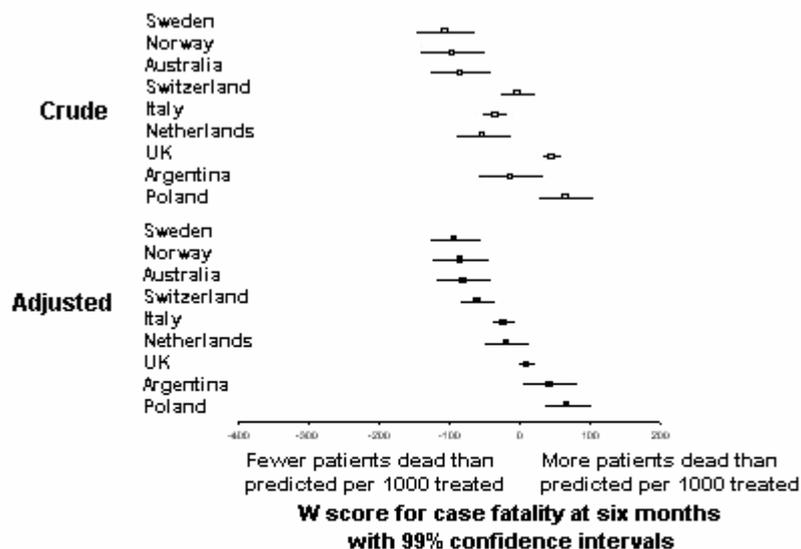
Figure 1. Comparison between countries of crude and adjusted case fatality at six months (presented as w-scores with 99% confidence limits). Point estimates indicate the absolute difference between the observed and the predicted number of patients dead at six months within each country, expressed as the number of excess deaths per thousand patients treated.

## Crude and adjusted 'death or dependency' at six months

Figure 2 shows the w-scores for death or dependency at six months before and after adjusting for casemix. As with case fatality, the crude number of patients dead or dependent at six months varied considerably between countries; the largest difference being between Sweden and the UK (375 more patients were dead or dependent per 1000 treated in the UK). Adjusting for casemix markedly reduced some of the between country differences, for example, that between Switzerland and Italy reduced by 86% and that between Argentina and Australia reduced by 60%. Again, however, despite adjusting for casemix, the overall variation in death or dependency between countries remained very substantial. For example, at six months, the difference between Sweden and Poland was 86 patients dead or dependent per 1000 treated and between Switzerland and the UK 146 patients dead or dependent per 1000 treated; at its most extreme, 311 more patients were dead or dependent per 1000 treated in the UK than in Sweden even after adjusting for casemix.
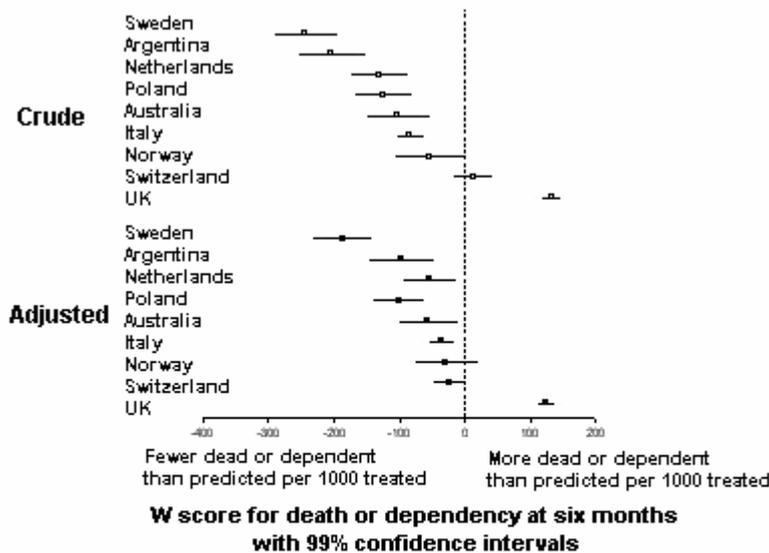
Figure 2. Comparison between countries of crude and adjusted 'death or dependency' at six months (presented as w-scores with 99% confidence limits).
Point estimates indicate the absolute difference in the observed and the predicted number of patients dead or dependent within each country, expressed as the number of excess cases dead or dependent per thousand patients treated.

**The association of the process of care with adjusted outcome**

The ranking of countries by adjusted outcome and by the proportion of patients that received each item of care is shown in table 3. Higher adjusted case fatality was strongly correlated with a lower rate of CT scanning before randomization (r 0.78) and also with a lower rate of prescription of long term aspirin (r 0.53). The correlation between adjusted death or dependency and the two process of care variables was weaker (r 0.40 in both cases).

Table 3. Rankings of countries according to outcome adjusted for casemix and by the provision of the two items of the process of care.
Better outcomes and better process of care are assigned lower numbers.

| Country | Ranked by outcome adjusted for casemix | | Ranked by the process of care | |
|---|---|---|---|---|
| | Case fatality at 6 months | Death or dependency at 6 months | CT before randomization | Discharged on long term aspirin |
| Sweden | 1 | 1 | 1 | 2 |
| Norway | 2 | 7 | 2 | 1 |
| Australia | 3 | 4 | 3 | 4 |
| Switzerland | 4 | 8 | 6 | 7 |
| Italy | 5 | 6 | 7 | 5 |
| Netherlands | 6 | 5 | 4 | 9 |
| UK | 7 | 9 | 9 | 8 |
| Argentina | 8 | 3 | 5 | 3 |
| Poland | 9 | 2 | 8 | 6 |

**The performance of the predictive models**

Figure 3 shows that both prognostic models were well calibrated (both calibration plots follow a 45° line). Both models also showed moderately good discrimination with an area under the ROC of 0.79 in each case. For both outcome states, however, a substantial proportion of the predicted probabilities (33% for death, 57% for death or dependency) lay between 25 and 75% i.e. the models failed to place substantial numbers of individuals into groups either very likely or very unlikely to have an outcome event.
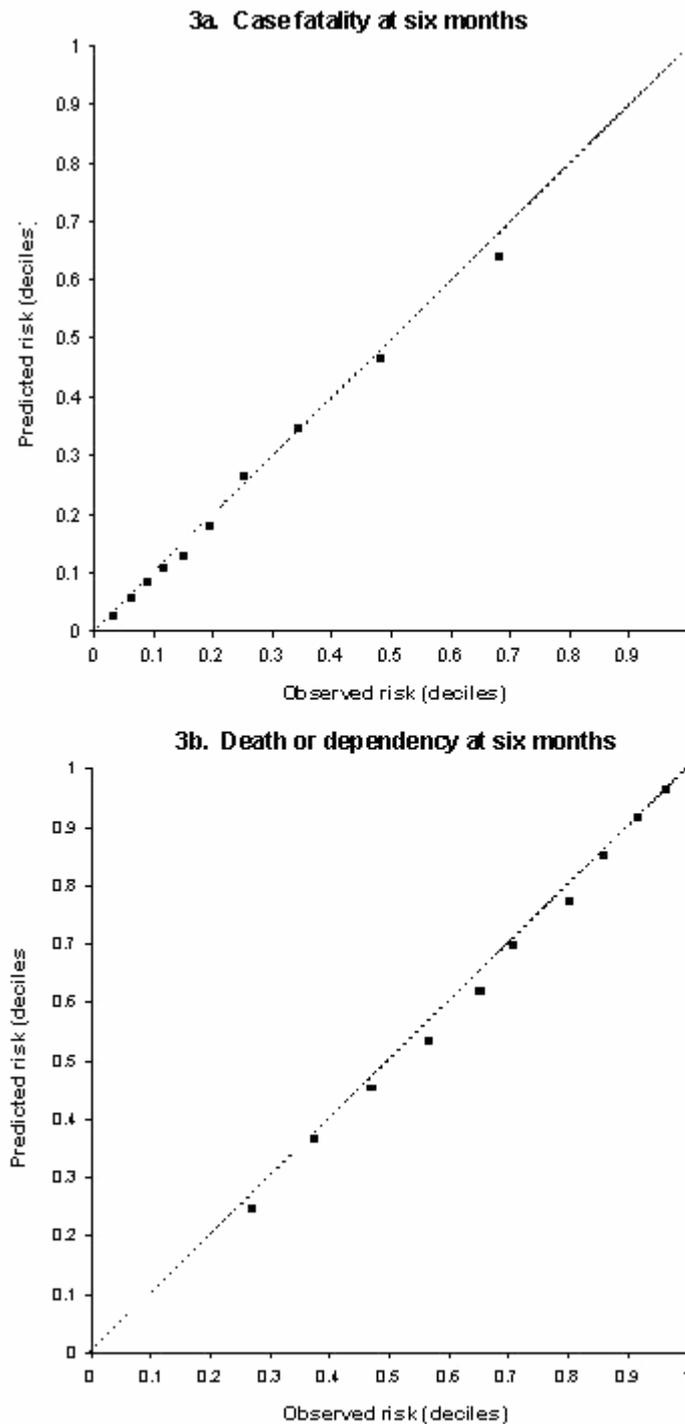


Figures 3a and 3b.
Calibration plots for the models predicting case fatality and 'death or dependency' at six months. The diagonal line represents perfect calibration.

## DISCUSSION

The principal aim of this study was to determine how far the large differences in crude outcome between countries participating in the IST might be explained by differences in the baseline characteristics of the

patients entered into the trial. The answer to this question is clearly complex. For certain pairs of countries, it appears that the differences in crude outcome could be explained very largely by differences in casemix, in particular, by differences in the age, sex, proportion with atrial fibrillation and baseline stroke severity (measured by conscious level, number of neurological deficits and systolic blood pressure) of the patients entered into the trial. For example, most of the large differences in case fatality between Sweden and Switzerland and in death or dependency between Switzerland and Italy were explicable in this way. However, for the majority of countries, adjustment for the factors in our prognostic models led only to modest changes in outcome and, as a result, the overall variation in case fatality and in death or dependency between countries remained substantial. This finding was quite a surprise. We had anticipated that, with a prospective design, the availability of detailed and important prognostic data, virtually complete follow up and the use of a validated prognostic model, we would have been able to account for most of the overall variation in outcome. Clearly, this was not the case and factors other than those considered in our analysis must account for the majority of the variation in outcome between countries.

Given the well described international differences in the treatment of acute stroke and speculation that these differences may contribute to differences in stroke outcome between countries (3;4;6-10), we considered whether differences in care between countries in the IST might explain their residual differences in outcome. Our observation that lower case fatality was strongly correlated with a higher proportion of patients with a CT head scan and with a higher proportion prescribed aspirin on discharge lends some support to this possibility. Similarly, it is notable that the Swedish and Norwegian cohorts had the lowest case fatality of all. The Scandinavian countries were early advocates of the organization and specialization of stroke care and, at the time of the IST, the majority of Swedish and Norwegian patients were likely to have been treated in a stroke unit (10). This is in sharp contrast to the UK, Argentina and Poland, the countries with the highest case fatality, where the majority of patients would have received conventional ward care (17). Such differences in care may well underlie at least some of the difference in outcome between the two groups of countries (18).

On the whole, however, the evidence from these data that differences in care might explain some of the residual differences in outcome is weak. Stroke interventions, especially those relating to secondary prevention, may quite properly be withheld from patients who survive in a very poor functional state. Our simplistic measurements of the quality of care do not take this into account and portray all cases where the intervention is withheld as an error. Our finding that higher case fatality is strongly correlated with apparently worse provision of care was therefore perhaps inevitable. The weaker correlation between death or dependency and our measures of the process of care also argues against a significant impact of quality of care on the differences in outcome between countries. In particular, if a difference in the proportion of patients treated on a stroke unit were an important reason for the difference in outcome, how is one to reconcile Norway's apparently excellent 'performance' as measured by case fatality but its comparatively poor 'performance' when measured by the combined outcome of death or dependency? Similarly, how does one explain the opposite findings for Argentina and Poland? Perhaps, one might argue, where there is better care, patients otherwise destined to die are more likely to survive but do so in a dependent state e.g. Norway; and where there is inferior care patients with a poor prognosis die and so cannot be counted as dependent e.g. Argentina and Poland.

However, if this were so, how would one explain the ranking of Sweden as the country with the best performance on both measures of outcome or the considerably worse performance of the UK when measured by death or dependency than when measured by case fatality alone?

A stronger argument that differences in stroke care are not the major cause of the residual differences in outcome is the sheer size of the absolute differences between countries. The differences in the proportion of patients dead or dependent between the UK and the other eight countries lay between 150 to 300 events per 1000 patients treated. These absolute differences in outcome are two to four times larger than the treatment effect of stroke unit care (18) and twice as large as the benefit of giving thrombolysis within 3 hours of the onset of stroke (19). That is to say, the differences in outcome are much larger than might plausibly be explained by the differential use of even the most efficacious of known interventions. Indeed,

at the time of the IST, thrombolysis was not in routine use. Thus, while it remains plausible that differences in medical treatment may account for some of the residual differences in outcome, other factors are likely to account for much more.

As with all observational studies, these other factors are the play of chance, bias and residual variation in casemix (confounding). Given our use of large samples and the narrow 99% confidence intervals, chance is highly unlikely to explain much of the residual variation in outcome between countries. Similarly, given the near complete follow up and the unambiguous nature of death, biased measurement cannot explain the residual variation in case fatality. However, measurement error may underlie some of the residual variation in death or dependency. Functional status is difficult to define and measure, and it is known that international comparisons are prone to bias (20;21). The fact that the range of differences between countries was greater for death or dependency than for death alone suggests that such bias might have operated here. In non-English speaking countries, the follow-up questions were translated into the local language without back-translation to check for alterations, and subtle but important differences in interpretation may have been introduced (22;23). The method used to collect outcome varied between countries. In some it was by postal questionnaire, in others by telephone, and in others still by a combination of the two methods. These differences might also have influenced response. Cultural differences in the perception of disability and dependence may also have lead patients in different countries to report dependency differently despite similar degrees of impairment in function (21;22;24;25). In general, however, the impact of these second order biases is modest. They would also be unlikely to explain the marked difference in death or dependency between patients in the UK and Australia, countries who used exactly the same outcome questionnaire and method of follow up and who experience reasonably similar cultures. Further, empirical research suggests that people in Sweden and the UK, the countries with the greatest difference in the proportions dead or dependent, value health states very similarly (26). It seems likely, therefore, that most of the residual variation in outcome between countries in the IST must be due to unmeasured variation in casemix.

The reason for the marked variation in casemix between countries in the IST is that the fundamental entry criterion for the trial was simply that the clinician had to be substantially uncertain whether or not to treat a given patient with aspirin, heparin, both or neither. Variation between countries in 'uncertainty', in the types of physician participating in the trial and in the types of stroke patients routinely admitted to hospital (9) are therefore all likely to have played their part. Although detailed, our prognostic models might have better accounted for the marked differences in casemix if they had adjusted for other recognized prognostic variables, such as pre-stroke functional status and living arrangements, co-morbid conditions, such as heart failure and diabetes mellitus, and post stroke urinary incontinence, hyperglycemia, and the size of the stroke lesion on brain imaging.(27) However, they certainly could not have accounted for the (probably many) important differences in casemix that are not currently understood. The limitations of our models are illustrated by the fact that for large numbers of patients the predictions of risk are in the middle, non-confident range i.e. they are not particularly good at separating patients into high and low risk groups. This in turn implies that they explain only a small part of the total variability in outcome (analogous to having a low $r^2$ statistic in linear regression). This is the case despite the fact that both models show excellent calibration and moderately good discrimination, two widely quoted measures of model 'performance'. This observation highlights the important difference between a prognostic model having a good 'fit' and it providing clinically useful predictions (28).

Regardless of their explanation, our findings have a number of implications. First, they emphasize the potential limitations of adjusting non-randomized comparisons of stroke outcomes for differences in casemix, even with high quality and complete data from large numbers of patients and especially when the differences in unadjusted outcome are very large. These observations may be of particular relevance to those attempting to draw inferences about the quality of care from non-randomized comparisons of stroke outcome, whether between hospitals, regions or countries, and especially if their data are retrospective, incomplete or less completely adjusted than our data. Second, investigators need to be aware that simply because casemix adjustment has been performed using models that show excellent calibration and moderately good discrimination, a considerable amount of variation in outcome may remain to be

explained. Therefore, and thirdly, before reaching conclusions about the quality of stroke care, the plausibility of ascribing any residual differences in outcome to variation in the use of currently understood stroke interventions should be carefully considered. Fourthly, those wishing to measure between country differences in functional outcome after stroke should recognize that this task is prone to various subtle forms of bias.

Lastly, our findings also have implications for the design of *multinational* randomized controlled trials. The variations in outcome between groups of patients in different countries in the IST do not affect the validity of the overall trial results because the trial used a method of allocation that ensured balance *within* countries (minimization). This design enabled the trial to detect the effect of a treatment (a reduction in death or dependency of 10 cases per 1000 treated) that was 30 times smaller than the largest difference in outcome between countries. If, however, a multinational study did not ensure balanced allocation within countries, then treatment effects might be spuriously generated or obscured. For instance, in an imaginary trial of a truly ineffective treatment using the same study population as the IST, if by chance the proportion of patients randomized to drug or to placebo happened to be 2:1 in Sweden and, for the same number of patients, 1:2 in the UK, then an apparent benefit of the truly ineffective drug would be observed even though the trial would have allocated equal numbers to each intervention.

In summary, this study demonstrates the potential limitations of using analyses of observational data to explain international differences in outcome after stroke. Differences in the quality of care, chance, measurement error and cultural bias may account for some of the residual variation in outcome between countries in the IST, but most of the unexplained variation is likely to reflect the difficulty of achieving perfect casemix adjustment. To avoid these biases, multinational randomized controlled trials in stroke must ensure a balance of treatment and control within each country as well as in the overall trial. Furthermore, those wishing to draw inferences from non-randomized comparisons of outcome after stroke should consider the issues raised in this paper carefully.

**ABSTRACT**
Background and purpose: This study describes the large variations in outcome after stroke between countries that participated in the International Stroke Trial and sought to define whether they could be explained by variations in casemix or by other factors.
Methods: We analyzed data from the 15,116 patients recruited in Argentina, Australia, Italy, the Netherlands, Norway, Poland, Sweden, Switzerland and United Kingdom. We compared crude case fatality and the proportion of patients 'dead or dependent' at six months; we used logistic regression to adjust for age, sex, AF, systolic BP, conscious level and number of neurological deficits. We used the frequency of pre-randomization CT head scan and prescription of aspirin at discharge to indicate quality of care.
Results: The differences in outcome (all treatment groups combined) between the 'best' and 'worst' countries were very large for death (171 cases per 1000 patients) and for death or dependency (375 cases per1000 patients). The differences were somewhat smaller after adjusting for casemix (160 and 311 cases per 1000 patients, respectively). Process of care may have accounted for some but not all of the residual variation in outcome.
Conclusion: Adjusting for casemix explained only some of the variation in outcome between countries. The residual differences in outcome were too large to be explained by variations in care, and are most likely to reflect differences in unmeasured baseline factors. These findings demonstrate the need to achieve balance of treatment and control within each country in multinational randomized controlled stroke trials; and for caution in the interpretation of non-randomized comparisons of outcome after stroke between countries.

**STATISTICAL APPENDIX**

We calculated the predicted risk of an outcome event (P) for each patient using the logistic regression models as follows: $P = e^y / (1 + e^y)$ where y is the linear predictor of the model and e is the exponential constant. We derived the linear predictor y of each model as follows:

*Case fatality at six months:*
y = -7.3529 + (0.0603 x age) - (0.1637 x sex) + (0.5130 x AF) + (0.9533 x conscious level) + (0.3272 x number of neurological deficits).
Dead or dependent at six months:
$y = -1.5288 + (1.059 \times age) + (0.2988 \times age^2) + (0.3066 \times sex) + (0.1963 \times AF) + (1.0634 \times conscious level) - (0.1012 \times sBP) + (0.2291 \times sBP^2) + (0.4249 \times number of neurological deficits)$
Coding of variables

Age = (age - 70)/20; $age^2$ = (age variable)$^2$

Sex: female = 1, male = 0

Atrial fibrillation (AF): present = 1, absent = 0

Conscious level: drowsy/unconscious = 1, fully conscious = 0

Systolic BP (sBP) = (sBP - 160)/60; $sBP^2$ = (sBP variable)$^2$

Number of neurological deficits = observed number except 0 coded 1; and 8 coded 7 (co-variate treated as a continuous variable because it has a linear relationship with each outcome)

## ACKNOWLEDGMENTS

## SOURCES OF FUNDING

## POTENTIAL CONFLICT OF INTEREST

None.

## REFERENCES

1. Thorvaldsen P, Asplund K, Kuulasmaa K, Rajakangas A, Schroll M, for the WHO MONICA Project. Stroke incidence, case fatality, and mortality in the WHO MONICA Project. Stroke 1995;26(3):361-7.

2. Stegmayr B, Harmsen P, Rajakangas A, Rastenyté D, Sarti C, Thorvaldsen P, Tuomilehto J. Stroke around the Baltic Sea: incidence, case fatality and population risk factors in Denmark, Finland, Sweden and Lithuania. Cerebrovasc.Dis. 1996;6:80-8.

3. Wolfe CD, Tilling K, Beech R, Rudd AG, for the European BIOMED study of stroke care group. Variations in case fatality and dependency from stroke in western and central Europe. Stroke 1999;30(2):350-6.

4. Wolfe CDA, Giroud M, Kolominsky-Rabas P, Dundas R, Lemesle M, Heuschmann P, Rudd A, for the European Registries of Stroke (EROS). Variations in stroke incidence and survival. Stroke 2000;31(9):2074-9.

5. Brainin M, Bornstein N, Boysen G, Demarin V. Acute neurological stroke care in Europe: results of the European Stroke Care Inventory. European Journal of Neurology 2000;7(1):5-10.

6. Caro JJ, Huybrechts KF, Duchesne I, for the Stroke Economic Analysis Group. Management patterns and costs of acute ischemic stroke. An international study. Stroke 2000;31(3):582-90.

7. McKevitt CJ, Beech R, Pound P, Rudd AG, Wolfe CDA. Putting stroke outcomes into context. Assessment of variations in the processes of care. The European Journal of Public Health 2000;10(2):120-6.

8. Beech R, Ratcliffe M, Tilling K, Wolfe C, on behalf of the participants of the European Study of Stroke Care. Hospital services for stroke care. A European perspective. Stroke 1996;27(11):1958-64.

9. Norris JW, Bogousslavsky J, Asplund K, Wester PO, Davis SM, Yamaguchi T, Oita J. Stroke management around the world. Cerebrovasc.Dis. 1994;4:430-40.

10. Asplund K, Rajakangas A, Kuulasmaa K, Thorvaldsen P, Bonita R, Stegmayr B, Suzuki K, Eisenblätter D. Multinational comparisons of diagnostic procedures and management of acute stroke: the WHO MONICA study. Cerebrovasc.Dis. 1996;6:66-74.

11. International Stroke Trial Collaborative Group. The International Stroke Trial (IST); a randomized trial of aspirin, subcutaneous heparin, both or neither among 19435 patients with acute ischemic stroke. Lancet 1997;349:1569-81.

12. Counsell CE, Clarke MJ, Slattery J, Sandercock PA. The miracle of DICE therapy for acute stroke: fact or fictional product of subgroup analysis? BMJ 1994;309(6970):1677-81.

13. Slattery J, Sandercock P. Prediction of death or dependency at six months in International Stroke Trial patients using data collected at randomization. Cerebrovasc.Dis. 1996;6(S2):8 (abstract).

14. Parry GJ, Gould CR, McCabe CJ, Tarnow-Mordi WO. Annual league tables of mortality in neonatal intensive care units: longitudinal study. International Neonatal Network and the Scottish Neonatal Consultants and Nurses Collaborative Study Group. BMJ 1998;316(7149):1931-5.

15. Smith EJ, Ward AJ, Smith D. Trauma scoring methods. Br.J.Hosp.Med. 1990;44(2):114-8.

16. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Annals of Internal Medicine 1999;130:515-24.

17. Lindley RI, Amayo EO, Marshall J, Sandercock PA, Dennis M, Warlow CP. Hospital services for patients with acute stroke in the United Kingdom: the Stroke Association survey of consultant opinion. Age Aging 1995;24 (6):525-32.

18. Stroke Unit Trialists' Collaboration. Collaborative systematic review of the randomized trials of organized (stroke unit) care after stroke. BMJ 1997;314:1151-8.

19. Wardlaw JM, Warlow CP, Counsell C. Systematic review of evidence on thrombolytic therapy for acute ischemic stroke. Lancet 1997;350(9078):607-14.

20. Suris JC, Blum RW. Disability rates among adolescents: an international comparison. Journal of Adolescent Health 1993;14:548-52.

21. Chamie M. Survey design strategies for the study of disability. World Health Organization Quarterly 1989;42:122-40.

22. Hunt SM. Cross-cultural issues in the use of socio-medical indicators. Health Policy 1986;6(2):149-58.

23. Picavet HSJ, van den Bos GAM. Comparing survey data on functional disability: the impact of some methodological differences. Journal of Epidemiology and Community Health 1996;50:86-93.

24. Chino N. Efficacy of Barthel Index in evaluating activities of daily living in Japan, the United States, and United Kingdom. Stroke 1990;21(SII):II-64-II-65.

25. Groce NE. Disability in cross-cultural perspective: rethinking disability. Lancet 1999;354:756-57.

26. Hunt SM, Wiklund I. Cross-cultural variation in the weighting of health statements: a comparison of English and Swedish valuations. Health Policy 1987;8:227-35.

27. Counsell CE. The prediction of outcome in patients with acute stroke (MD thesis).
Cambridge University, UK, 1998.

28. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat.Med. 2000;19(4):453-73.

**Top**

Your questions, contributions and commentaries will be answered
by the lecturer or experts on the subject in the Stroke list.
Please fill in the form (in Spanish, Portuguese or English) and press the "Send" button.

Question, contribution or commentary:

Name and Surname:

Country: Argentina

E-Mail address: @

**Send**    **Erase**

**Top**

---

**2nd Virtual Congress of Cardiology**

---

| **Dr. Florencio Garófalo** | **Dr. Raúl Bretal** | **Dr. Armando Pacher** |
|---|---|---|
| Steering Committee President | Scientific Committee President | Technical Committee - CETIFAC President |
| **fgaro@fac.org.ar** | **rbretal@fac.org.ar** | **apacher@fac.org.ar** |

**fgaro@satlink.com**    **rbretal@netverk.com.ar**    **apacher@satlink.com**

This company contributed to the Congress

Gador