



Published in final edited form as:

Crit Care Med. 2011 January ; 39(1): 158–162. doi:10.1097/CCM.0b013e3181fb7b49.

Do-Not-Attempt-Resuscitation Orders and Prognostic Models for Intraparenchymal Hemorrhage

Claire J Creutzfeldt, MD, Kyra J Becker, MD, Jonathan R Weinstein, MD PhD, Sandeep P Khot, MD, Thomas O McPharlin, RPh, Thanh G Ton, PhD, WT Longstreth Jr, MD, and David L Tirschwell, MD

Department of Neurology, School of Medicine, University of Washington, Seattle, WA The work was performed at University of Washington Harborview Medical Center in Seattle, WA

Abstract

Objective—Statistical models predicting outcome after intraparenchymal hemorrhage (IPH) include patients irrespective of do-not-attempt-resuscitation (DNAR) orders. We built a model to explore how the inclusion of patients with DNAR orders affects IPH prognostic models.

Design—Retrospective, observational cohort study from May 2001 until September 2003

Setting—University-affiliated tertiary referral hospital in Seattle, Washington

Patients—424 consecutive patients with spontaneous intraparenchymal hemorrhage

Measurements—We retrospectively abstracted information from medical records of IPH patients admitted to a single hospital. Using multivariate logistic regression of presenting clinical characteristics, but not DNAR status, we generated a prognostic score for favorable outcome (FO, defined as moderate disability or better at discharge). We compared observed probability of FO with that predicted, stratified by DNAR-status. We then generated a modified prognostic score using only non-DNAR patients.

Main Results—Records of 424 patients were reviewed: 44% had FO, 43% had a DNAR-order and 38% died in hospital. Observed and predicted probability of FO agreed well with all patients taken together. Observed probability of FO was significantly higher than predicted in non-DNAR patients and significantly lower in DNAR patients. Results were similar when applying a previously published and validated prognostic score. Our modified prognostic score was no longer pessimistic in non-DNAR patients, but remained overly optimistic in DNAR patients.

Conclusions—Although our prognostic model was well calibrated when assessing all IPH patients, predictions were significantly pessimistic in patients without, and optimistic in those with DNAR orders. Such pessimism may drive decisions to make patients DNAR in whom a FO may have been possible, thereby creating a self-fulfilling prophecy. To be most useful in clinical decision-making, IPH prognostic models should be calibrated to large IPH cohorts in whom DNAR orders were not used.

Correspondence to Dr. Claire Creutzfeldt at Department of Neurology, Box 359775, Harborview Medical Center, 325 Ninth Avenue, Seattle, WA 98020-2420. Phone: 206 744 3251 Fax: 206 744 8787 clairejc@uw.edu.

No reprints will be ordered.

There was no financial support for this study.

Dr. Tirschwell received a grant from NIH. All other authors have no potential conflicts of interest to disclose.

Ethics approval: was obtained from the University of Washington Human Subjects Review Committee.

Keywords

Intraparenchymal Hemorrhage; Do-Not-Resuscitate Orders; Prognosis; Decision Analysis; Ethics

Outcome prediction models are an attempt to provide reliable information about a patient's chances of survival or recovery and to facilitate decisions about certain aspects of care. Primary intraparenchymal hemorrhage (IPH) accounts for about 15% of all strokes and has the highest mortality with 20-30% of IPH patients dying within 3 months. Many statistical models predicting outcome after IPH have been developed (1) but have included patients irrespective of do-not-attempt-resuscitation (DNAR) orders (2). Even though DNAR orders do not indicate the withholding of any treatment other than cardiopulmonary resuscitation (CPR), they are associated with a decreased willingness to treat (3) and with an increased risk of death after IPH (4, 5). Clinicians and investigators applying such prognostic models have acknowledged the risk of a self-fulfilling prophecy (6), which seriously threatens the validity of such models. In this study, we sought not the promotion of yet another IPH prognostic model but to demonstrate quantitatively how inclusion of DNAR patients in the derivation of a prognostic model leads to overly pessimistic predictions of outcome for individual patients not made DNAR.

Methods

Setting

We reviewed records of 424 consecutive patients admitted to a single hospital for spontaneous (primary) intraparenchymal hemorrhage (IPH) between May 2001 and September 2003, excluding patients whose IPH was not arterial or was secondary to a vascular malformation, neoplasm, or some other known source. Harborview Medical Center is a Joint Commission certified Primary Stroke Center that serves as a tertiary referral center for patients with severe neurological disease. Approximately 20% of all Seattle and surrounding King county residents with IPH are admitted to this hospital. The Human Subjects Review Committee reviewed and approved the study.

Do-not-attempt-resuscitation (DNAR) orders

At this hospital, CPR is performed by default on any patient whose heart stops suddenly, unless a physician signs an order not to attempt CPR, a so-called do-not-attempt-resuscitation (DNAR) order. Such an order is entered on a special form that includes the justification for the DNAR order, such as the preference of the patient or healthcare surrogate, a futility judgment by the treating physicians, or both. A DNAR order is not meant to affect any treatment other than CPR. In this study, the presence or absence of a DNAR order was documented but not included in the development of the IPH prognostic model.

Variables for the Prognostic model

The variables used for prognostic model development were abstracted retrospectively from medical records. Demographic data included age, gender, race and pre-hospital functional status. Clinical data included initial blood pressure, heart rate and rhythm, temperature, Glasgow Coma Scale (GCS) score, serum glucose level, and brain imaging characteristics. Based on initial imaging reports, hemorrhage location was classified as supratentorial or infratentorial, and the following findings were documented as present or not: mass effect, midline shift, intraventricular extension, hydrocephalus, and herniation. The medical chart was carefully reviewed for functional status at the time of discharge and then translated into the modified Rankin scale (mRS). A favorable outcome (FO) was defined as moderate

disability or better at time of hospital discharge, corresponding to a mRS score of three or less.

Prognostic score development

Clinical characteristics described in the previous section were entered into a multivariate logistic regression model to identify characteristics that were independently and significantly ($p < 0.05$) associated with FO. The prognostic score for each patient was calculated as the linear sum of the beta coefficients derived from the logistic regression model and was based on each patient's values for the variables in the model. The predicted probability of FO is a mathematical manipulation of the prognostic score to confine the results to the range of 0.0 to 1.0. Thus, each patient had an individualized prognostic score and a corresponding predicted probability of FO. We then divided the patient cohort into five groups of equal size (quintiles) based on their predicted probability of FO, with quintile 1 representing a low probability of FO (worst prognosis) and quintile 5 a high probability of FO (best prognosis). Finally, observed and predicted proportions of FO were stratified by DNAR status. The ratio of observed (O) to predicted (P) proportion of FO was used to define pessimism; if a higher proportion of patients was observed to have favorable outcome than had been predicted ($O/P > 1$), the model was considered pessimistic.

To support our findings, we used another validated prognostic model, the ICH Score (7), with each ICH score value as an indicator variable in a logistic regression to generate predicted probabilities of FO for each ICH score value, as we had done for the prognostic score described above. The score could be calculated only in the subset of our cohort where all the needed data elements were available, or 300 of the 424 patients. As in the original report (7), scores ranged from 0 to 5 and no one scored 6. Both non-DNAR and DNAR patients were included in the model that generated probability of FO. We then compared the O and P proportions of FO by ICH Score stratified by DNAR status. Details of the models are contained in supplemental Table I.

Formal statistical testing of goodness-of-fit of the prognostic models in all patients and in the non-DNAR and DNAR patients separately was performed using the Hosmer-Lemeshow goodness-of-fit (HL-GOF) test. This approach compares observed to predicted outcomes in groups divided by percentiles of predicted outcome using the chi-square test statistic. A non-significant p-value (> 0.05) for this HL-GOF test indicates a model with good fit, a significant p-value indicates important differences exist between the observed and predicted values and a model with a poor fit (8).

All analyses were performed using STATA (version 10, StataCorp, College Station, TX).

Results

We identified 424 patients with a mean age of 65 years: 49% were women and 43% had a DNAR order (Table 1). Overall, 38% died in hospital and 44% had a favorable outcome (FO) at discharge. The DNAR order was written on the day of admission in about one third (64/183, 35%), by hospital day 2 or 3 in another third (61/183, 33%) and on days 4-6 (22/183, 12%) or thereafter (36/183, 20%) in the remaining third. According to our multivariate logistic regression including all patients, irrespective of DNAR orders, factors significantly and independently associated with outcome were age, GCS score, heart rate, IPH mass effect, intraventricular extension, premorbid level of function, and systolic blood pressure (see Table 1). Even though significantly more DNAR patients were on anticoagulation at the time of admission than non-DNAR patients, after multivariate logistic regression Warfarin use was not an independent predictor of outcome. Excluding patients who were on Warfarin did not substantively change the results (data not shown). The overall

performance of our prognostic score was excellent, with the area under the receiver operator curve (ROC) curve being 0.93 (an area under ROC curve of 1.0 indicates perfect performance).

When all patients were taken together, irrespective of DNAR orders (Figure 1a), observed and predicted proportion of FO agreed well indicating a model with a good fit (HL-GOF $P = 0.97$). When accounting for DNAR status (Figure 1 b, c), this agreement is lost, indicating models with a poor fit (both with HL-GOF $P < 0.001$). Favorable outcome was underestimated in non-DNAR patients and overestimated in DNAR patients. The ratio of observed to predicted proportion of FO (O/P) in any given quintile can be used to demonstrate the pessimism of such a prognostic model: For example, in quintile 3 in the non-DNAR patients, observed FO was 55% compared to a predicted FO of 31%. The predicted FO was underestimated, or pessimistic ($O/P = 55/33 = 1.8$, greater than 1). On the other hand, in the DNAR patients in quintile 3, observed FO was 3% compared to a predicted FO of 29%. The predicted FO was overestimated, or optimistic ($O/P = 3/29 = 0.10$, smaller than 1).

Using the ICH Score as an alternative prognostic score (7) to generate the probability of FO for both non-DNAR and DNAR patients yielded similar results. The predictions were pessimistic in the non-DNAR group and optimistic in the DNAR group (Figure 2; again both HL-GOF $p < 0.001$).

Finally, when including only non-DNAR patients to create another prognostic model, the independent prognostic factors were the same as for the entire cohort, except that age, ventricular extension and heart rate were no longer significantly associated (see supplemental Table 1). Predicted probabilities of FO were no longer pessimistic in the non-DNAR patients and the model had a good fit (HL-GOF $p = 0.94$), though they remained overly optimistic when applied to the DNAR group (Figure 3).

Discussion

To demonstrate quantitatively the pessimism intrinsic to IPH prognostic models, we created our own model that was derived from both non-DNAR and DNAR patients. Although our model was well calibrated when assessing the entire cohort (Figure 1a), it overestimated the probability of favorable outcome (FO) for DNAR patients (Figure 1b) while underestimating that for non-DNAR patients (Figure 1c). Using the validated ICH Score (7) instead of our own prognostic scoring system led to a similar result (Figure 2). This pessimism in non-DNAR patients could be eliminated by developing the prognostic model using only non-DNAR patients (Figure 3).

Prognostication of patient outcome by relying on clinician experience and judgment represents the art of medicine. Prognostic models are an attempt to make this art a science. In order to create an objective prognostic model, important prognostic factors need to be identified and considered in the analyses. Most IPH prognostic models include patients with DNAR orders in the creation of the models without accounting for them in their analyses (2). Only three published studies specifically examined the role of DNAR orders or withdrawal of life-sustaining measures in the prediction of outcome after IPH (5, 10, 11). The few studies that accounted for any limitation of care in the development of an IPH outcome prediction model did this by excluding particularly severe patients (11-13). It is just this group of patients, however, in whom the issue of prognostic pessimism is important. Our model tended to be particularly pessimistic for those patients with moderate to moderately severe IPH, represented by quintiles 2 and 3 in Figure 1 and in whom a DNAR order has not been signed. This pessimism may drive the decision to assign a DNAR order

to an individual in whom a favorable outcome may have been possible. The unintended result may be a self-fulfilling prophecy (9, 14), in which DNAR orders themselves lead to worse patient outcome and increased mortality. The findings from our models may be used to estimate the number of patients who would lose their chance of a favorable outcome due to IPH prognostic pessimism (Figure 4). The data for this figure is derived from a tertiary referral center, where a higher percentage may be expected of patients with moderate to moderately severe IPH and, within this group, of patients with DNAR orders. To illustrate the uncertainty around our estimate, we computed a 95% confidence interval on the observed risk difference: Still, at least one hundred thousand IPH patients worldwide and several thousands just in the US might be denied aggressive care because of an overly pessimistic prediction of poor outcome.

Although our study has strengths in the large number of patients, assessment of functional outcomes rather than death, and careful analyses with supplemental tables that enable the interested reader to follow the development of our prognostic model, it also has its weaknesses. The retrospective abstraction of information when the outcome is “function at hospital discharge” risks the reviewer’s knowledge of the outcome affecting the classification of prognostic factors. The DNAR orders were easily identified but were not initiated in a standard fashion or at a standard time. Our analysis did not incorporate the reasons for why a DNAR order was written, and it is conceivable that some patients were made DNAR based on a medical complication developing later in their hospital course. To assess the possible impact of such a bias, we re-analyzed the data excluding patients with a late DNAR order, defined as one written at or after 72 hours of hospitalization. Even with this new definition, favorable outcome was overestimated for DNAR patients and underestimated for non-DNAR patients, with the pessimism remaining statistically significant. We also did not know if patients had DNAR orders before their bleeds. Given the excellent pre-morbid functional status, regardless of subsequent DNAR status, few if any patients likely had DNAR orders before their bleeds.

We did not collect information on withdrawal of life-sustaining measures and so-called comfort care. We may have neglected to collect information on some other major prognostic factor or determinant of DNAR status but our model contained variables similar to the many other prognostic models previously reported (1). In comparison to patients who were not made DNAR, DNAR patients were more likely to have medical comorbidities, a worse clinical status on admission, and a poor outcome (table 1). The potential for unintended consequences of DNAR orders is well known (3-5). In other words, patients with a poor-prognosis IPH are more likely to have a DNAR order, and patients with a DNAR order have a worse prognosis. However, this potential causal relationship does not lessen the main interpretation of our study, namely that the presence of DNAR patients in cohorts used for prognostic model derivation leads to pessimistic outcome predictions for those patients who are not DNAR. This model was developed from the experience of a single tertiary referral center and has not been externally validated. It should be emphasized that this model was developed solely for demonstration purposes and is not intended for clinical use. It should also be kept in mind that the ICH Score was developed to assess 30-day mortality, and more recently has been validated for the prediction of functional outcome at one year (15), whereas our analyses assessed functional outcome prediction at discharge. Although the lack of long-term follow-up may weaken the clinical usefulness of our model, it does not change the observed effect of including DNAR patients in the derivation of prognostic models for IPH. Creating a prognostic model using data from non-DNAR patients only eliminated the problems introduced in models by ignoring DNAR status, but also raises the question as to whether this non-DNAR subset is representative of all IPH patients.

Conclusion

In order to make an informed decision about their loved-one's medical care, families ask what the most likely outcome will be. Implicit in this question is the assumption that this is the outcome that would be achieved if aggressive care were to be continued. To be most objective and useful in clinical decision-making, IPH prognostic models should be calibrated to large IPH cohorts with the full spectrum of disease severity, in which patients are not made DNAR and in whom aggressive care is reasonably pursued. While the limitation of aggressive care is likely appropriate in individual cases with severe IPH, medical providers need to acknowledge and account not only for their own values and biases but also for those inherent in the prognostic models they use, when communicating with families, interpreting patient's wishes, and helping them make a decision about life or death.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Ariesen MJ, Algra A, van der Worp HB, et al. Applicability and relevance of models that predict short term outcome after intracerebral haemorrhage. *J Neurol Neurosurg Psychiatry*. 2005; 76(6): 839–44. [PubMed: 15897508]
2. Ton T, Creutzfeldt C, Becker KJ, et al. Accounting for do-not-resuscitate orders and withdrawal of care in prognostic models for intraparenchymal hemorrhage [abstract]. *Neurology*. 2009; 72(11, suppl. 3):A274.
3. Beach MC, Morrison RS. The effect of do-not-resuscitate orders on physician decision-making. *J Am Geriatr Soc*. 2002; 50(12):2057–61. [PubMed: 12473020]
4. Hemphill JC III, Newman J, Zhao S, et al. Hospital usage of early do-not-resuscitate orders and outcome after intracerebral hemorrhage. *Stroke*. 2004; 35(5):1130–1134. [PubMed: 15044768]
5. Zahuranec DB, Brown DL, Lisabeth LD, et al. Early care limitations independently predict mortality after intracerebral hemorrhage. *Neurology*. 2007; 68:1651–1657. [PubMed: 17502545]
6. Holloway RG, Benesch CG, Burgin WS, et al. Prognosis and decision making in severe stroke. *JAMA*. 2005; 294(6):725–33. [PubMed: 16091575]
7. Hemphill JC 3rd, Bonovich DC, Besmertis L, et al. The ICH score: a simple, reliable grading scale for intracerebral hemorrhage. *Stroke*. 2001; 32(4):891–7. [PubMed: 11283388]
8. Hosmer, DW.; Lemeshow, S. *Applied logistic regression*. 2nd ed. Wiley; New York: 2000.
9. Becker KJ, Baxter AB, Cohen WA, et al. Withdrawal of support in intracerebral hemorrhage may lead to self-fulfilling prophecies. *Neurology*. 2001; 56(6):766–72. [PubMed: 11274312]
10. Matchett SC, Castaldo J, Wasser TE, et al. Predicting mortality after intracerebral hemorrhage: comparison of scoring systems and influence of withdrawal of care. *J Stroke Cerebrovasc Dis*. 2006; 15(4):144–50. [PubMed: 17904067]
11. Weimar C, Benemann J, Diener HC. Development and validation of the Essen Intracerebral Hemorrhage Score. *J Neurol Neurosurg Psychiatry*. 2006; 77(5):601–605. [PubMed: 16354736]
12. Weimar C, Roth M, Willig V, et al. Development and validation of a prognostic model to predict recovery following intracerebral hemorrhage. *J Neurol*. 2006; 253(6):788–793. [PubMed: 16525882]
13. Rost NS, Smith EE, Chang Y, et al. Prediction of functional outcome in patients with primary intracerebral hemorrhage. *Stroke*. 2008; 39:2304–2309. [PubMed: 18556582]
14. Hemphill JC 3rd. Do-not-resuscitate orders, unintended consequences, and the ripple effect. *Crit Care*. 2007; 11(2):121. [PubMed: 17338835]
15. Hemphill JC 3rd, Farrant M, Neill TA Jr. Prospective validation of the ICH Score for 12-month functional outcome. *Neurology*. 2009; 73(14):1088–94. [PubMed: 19726752]

16. Johnston SC, Mendis S, Mathers CD. Global variation in stroke burden and mortality: estimates from monitoring, surveillance, and modelling. *Lancet Neurol.* 2009; 8(4):345–54. [PubMed: 19233730]
17. Lloyd-Jones D, Adams RJ, Brown TM, et al. on behalf of the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart Disease and Stroke Statistics--2010 Update. A Report From the American Heart Association. *Circulation.* 2010; 121:e1–e170. [PubMed: 20048228]

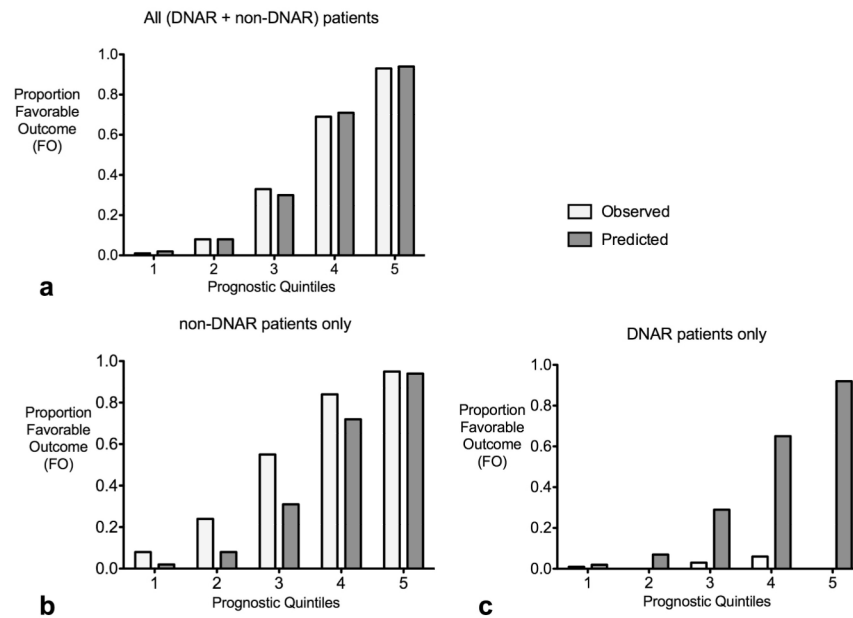


Figure 1. Observed (white columns) versus predicted (grey columns) probabilities of favorable outcome (FO) by prognostic quintile

a. When all patients (DNAR and non-DNAR patients) are taken together, observed and predicted probabilities of FO agree well indicating a model with a good fit (goodness-of-fit p -value = 0.97). Number of patients is 85 per quintile (84 in quintile 5).

b, c. When patients are stratified by DNAR status, observed and predicted probabilities of FO no longer agree well indicating models with poor fit (for both, goodness-of-fit p -value < 0.001). Numbers of patients for quintiles 1-5 respectively are 12, 29, 49, 69, 82 for non-DNAR patients (b), and 73, 56, 36, 16, 2 for DNAR patients (c).

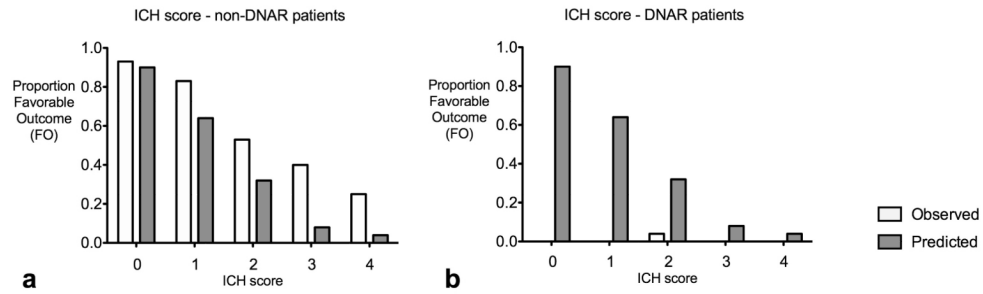


Figure 2. Model performance by ICH Score

This model was generated in 300 patients in whom the ICH Score (1) could be calculated, irrespective of DNAR orders. Model performance is shown, stratified by DNAR status, and shows disparities between observed (white columns) and predicted (grey columns) probability of favorable outcome (FO) indicating models with poor fit (for both, goodness-of-fit p-value < 0.001). Numbers of patients for ICH scores 0-4 respectively are 58, 60, 32, 15, 4 for non-DNAR patients (a), and 2, 18, 25, 58, 24 for DNAR patients (b). Because only 4 patients had ICH Score 5, all of whom had a poor outcome, these were not included in logistic model or figures.

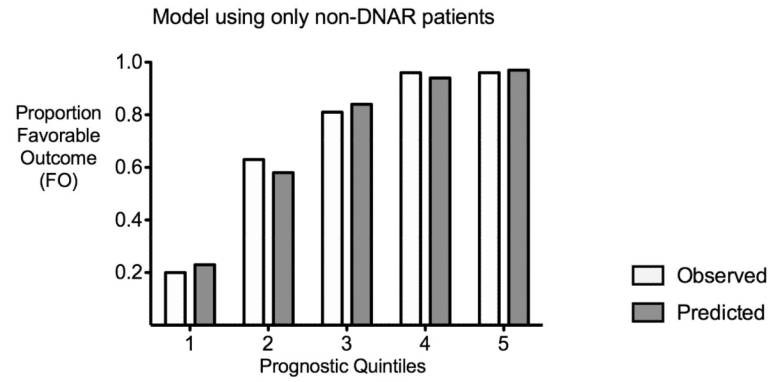


Figure 3. Prognostic model using only non-DNAR patients
Observed (white columns) and predicted (grey columns) probabilities of favorable outcome (FO) agree well, the pessimism in the non-DNAR patients is eliminated.

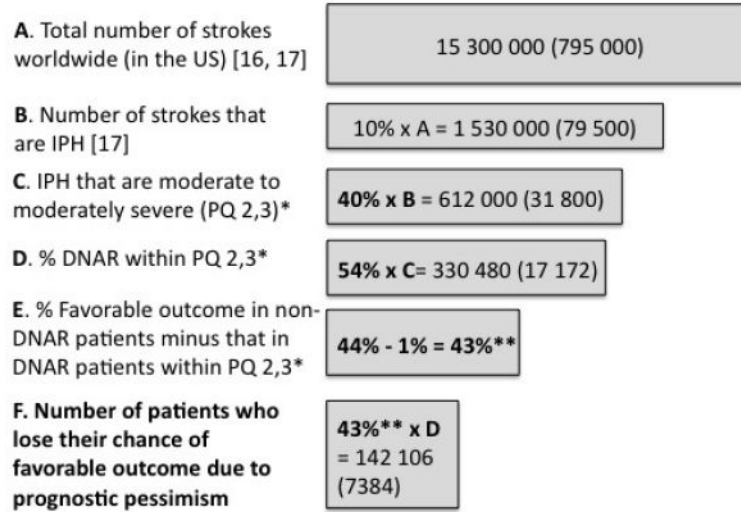


Figure 4. Estimated number of lives affected by pessimism in IPH models, worldwide (in the US)

*PQ 2, 3= prognostic quintiles 2 and 3 from Figure 1, i.e. moderate and moderately severe IPH.

**95% confidence interval (CI) on difference of 43% is 31 – 54%, therefore 95% CIs on the final estimates for worldwide (US) are 102 449 – 178 459 (5 323 – 9 273).

Table 1

Baseline characteristics of all intraparenchymal hemorrhage patients, comparing do-not-attempt-resuscitate patients with non- do-not-attempt-resuscitate patients

	Overall (n = 424)	Non-DNAR (n = 241)	DNAR (n = 183)	<i>p</i> value
Baseline characteristics				
Age (SD)	65	63 (17)	69 (14)	< .001
Gender, female	49	48	50	.66
Hypertension	67	68	65	.59
Diabetes	20	21	18	.42
CAD	12	13	10	.43
Warfarin use	13	8	20	<.001
Premorbid independence	90	85	94	.004
Admission				
GCS (IQR)	10.5	14 (9-15)	6 (4-9)	<.001
SBP (sd)	170	169 (37)	171 (38)	.54
HR (sd)	85	82 (18)	90 (26)	<.001
mass effect	59	40	85	<.001
IVH	51	34	72	<.001
Interventions				
Intubated	56	34	84	<.001
Ventriculostomy	15	12	17	.19
Craniotomy	12	12	11	.96
Outcome at discharge				
FO	44	71	8	<.001
Died in hospital	38	5	83	<.001

DNAR, do-not-attempt-resuscitation (order); CAD, coronary artery disease; GCS, Glasgow Coma Scale score; SBP, systolic blood pressure; DBP, diastolic blood pressure; HR, heart rate; IVH, intraventricular hemorrhage; FO, favorable outcome (moderate disability or better).

All numbers are percentages or means (standard deviations).