rt-PA and Stroke: Does IST-3 Make It All Clear or Muddy the Waters?

Answers to the November 2012 Journal Club Questions

Ryan P. Radecki, MD; Yashwant G. Chathampally, MD, MS; Gregory M. Press, MD, RDMS

Editor’s Note: You are reading the 30th installment of Annals of Emergency Medicine Journal Club. This Journal Club refers to the Third International Stroke Trial (IST-3) that was published in Lancet. Information about the journal club can be found at http://www.annemergmed.com/content/journalclub. Readers should recognize that these are suggested answers. We hope they are accurate; we know that they are not comprehensive. There are many other points that could be made about these questions or about the article in general. Questions are rated novice “Nov,” intermediate “Int,” and advanced “Adv” so that individuals planning a journal club can assign the right question to the right student. The “novice” rating does not imply that a novice should be able to spontaneously answer the question. “Novice” means we expect that someone with little background should be able to do a bit of reading, formulate an answer, and teach the material to others. Intermediate and advanced questions also will likely require some reading and research, and that reading will be sufficiently difficult that some background in clinical epidemiology will be helpful in understanding the reading and concepts. We are interested in receiving feedback about this feature. Please e-mail journalclub@acep.org with your comments.

DISCUSSION POINTS

1. The Third International Stroke Trial (IST-3) is described as a randomized, controlled, open-treatment trial to address the efficacy of intravenous recombinant tissue plasminogen activator (rt-PA) outside the exclusionary criteria of the European license for treatment of acute stroke. Enrollment began in 2000 and concluded in 2011.1

Nov: A. The authors summarize the eligibility criteria in terms of the "uncertainty principle."1 Describe this principle as applied in the context of IST-3 and compare it to the principle of clinical equipoise in general practice treatment decisions. Is this an ethical enrollment criterion? Will enrollment based on this principle be consistent across the entire study time frame?

Int: B. The investigators defined an “experienced” stroke center as one that treated at least 3 patients with rt-PA in the preceding 12 months before joining the trial.1 What do you think about this definition? What percentage of patients were enrolled at experienced centers? Consider also the temporal distribution of patient enrollment in the study period. How might these factors affect internal and external validity?

Adv: C. After enrolling 276 patients, IST-3 switched from a double-blind trial to an open-treatment investigation. Control group patients in the open-treatment arm did not receive a corresponding bolus or infusion. Describe the placebo and nocebo effects and how they might impact study results. What outcomes are more likely to be exaggerated by the lack of a placebo?

2. The authors conclude “for the types of patients recruited in IST-3, despite the early hazards, thrombolysis within 6 h improved functional outcome.”1

Nov: A. Is this statement consistent with the results of the primary outcome measured in this study? On what evidence do the authors base this conclusion?

Int: B. The authors allude to “early hazards” after treatment with rt-PA. There was a 1.59 (95% confidence interval 1.23 to 2.07) unadjusted increased odds of death within the first 7 days after treatment with rt-PA, but this effect disappeared at 6 months. Hypothesize possible underlying mechanisms for the disappearance of this difference over time. What data acquired in IST-3 might allow limited retrospective exploration of these hypotheses?

Adv: C. The authors include an “ordinal shift analysis” as a secondary outcome in IST-3. Describe the advantages of shift analysis over a dichotomous endpoint in terms of information theory. What limitations might shift analysis have compared with a dichotomous endpoint? Should shift analysis increase or decrease power to detect a difference between treatment arms?

3. The editorial accompanying the study publication in the Lancet concludes with the statement “. . . the default situation for the first health-care professional who identifies the stroke patient should be to treat, and the role of stroke and emergency physicians is now not to identify patients who will be given rt-PA, but to identify the few who will not.”2

Volume 61, No. 4 : April 2013 Annals of Emergency Medicine 489
A. This statement is based on the results from IST-3 and the accompanying updated meta-analysis of rt-PA for acute stroke. Does IST-3 provide evidence that only a “few” patients will not benefit from rt-PA? Concisely summarize the findings of this study in the context of evidence from previous stroke trials.

B. Extensive conflict of interest disclosures are reported for study authors. What effect might these disclosures have on clinicians when they interpret study results? What biases have been demonstrated as occurring in studies whose authors have substantial conflicts of interest?

C. A patient presents to your emergency department with an acute stroke that was noted nearly 5 hours before arrival. A noncontrast computed tomography scan of the head shows no hemorrhage or signs of recent ischemic change, and the patient has no important vital sign or laboratory abnormalities. The family at bedside asks whether the patient is eligible for “clot-busting treatment.” Describe the informed consent you might offer the patient and family for rt-PA in the context of the results of IST-3.

ANSWER 1

Q1. The Third International Stroke Trial (IST-3) is described as a randomized, controlled, open-treatment trial to address the efficacy of intravenous recombinant tissue plasminogen activator (rt-PA) outside the exclusionary criteria of the European license for treatment of acute stroke. Enrollment began in 2000 and concluded in 2011.1

Q1.a The authors summarize the eligibility criteria in terms of the “uncertainty principle.” Describe this principle as applied in the context of IST-3 and compare it to the principle of clinical equipoise in general practice treatment decisions. Is this an ethical enrollment criterion? Will enrollment based on this principle be consistent across the entire study time frame?

The authors state that if a patient had either a clear indication or contraindication for treatment with rt-PA, he or she was not eligible for enrollment in the trial. They go on to state, after informed consent, that only if both the clinician and patient (or relevant proxy) believed that the treatment was “promising but unproven” could the patient be enrolled as a study subject.1 The trial included the oversight of an independent data monitoring committee that, according to the predefined measures, would have recommended protocol changes including termination of the trial if a state of clinical equipoise no longer existed. The authors state that no such recommendations were made.

Djulbegovic, in his 2007 article, reviewed the following work regarding the concept of equipoise and the uncertainty principle. In 1963, Bradford Hill wrote that randomization should be accepted “only in our state of ignorance, the treatment given being a matter of indifference.” In 1974, Friedman introduced the term “equipoise,” stipulating that a physician may only enroll a patient in a trial if he or she is genuinely uncertain which treatment is preferred. As defined, equipoise is based on uncertainty at the level of the individual physician rather than at the level of the patient or community of physicians. In 1987, Friedman suggested that the state of equipoise be shifted from the individual physician to a community of expert practitioners. In 1998, Peto coined the term “uncertainty principle,” requiring that a patient be enrolled in a trial only if both the patient and the physician met an unquantified threshold of uncertainty with regard to the relative merits of the treatment options. In 2001, Lilford et al restated the unresolved question as, “How much uncertainty can one accept before we enroll a patient into a trial and by whom (patients, physicians, and community)?”

Specific to the IST-3 trial, a state of clinical equipoise exists about rt-PA’s benefit for patients who do not meet criteria for treatment per the European license. For example, the authors state that fewer than 100 patients older than 80 years have been enrolled in randomized trials, resulting in insufficient data to conclude whether these patients should be treated with rt-PA. The authors state that the trial was intended from its outset to be “pragmatic,” focused on improving the external validity. This implies that the authors are willing to compromise internal validity to gain external validity and execute a trial that supports not just efficacy but also effectiveness.1 An “efficacy” trial is one that attempts to reject the null hypothesis by finding a difference in the dependent variable by testing the treatment in question under ideal conditions, prioritizing rigorous internal validity over generalizability across diverse practice settings. If a treatment fails under ideal conditions, then it is not likely to be additionally tested for “effectiveness,” or generalizability, in diverse practice conditions. The design consideration is that a treatment that is efficacious when applied in experienced centers by experts under ideal conditions may be less so when application is limited in diverse practice settings by variable experience or expertise.

Enrollment was not likely to be uniform, given the sources of variability within the trial, including subjectivity in the application of the uncertainty principle, institutional experience, provider experience, and geography. Additionally, the 11-year duration of the trial is likely to create enrollment variability because of changes in practice standards regarding thrombolytics in stroke during that period. All these sources of variation likely altered the enrollment procedure and the type of patient enrolled, contributing to the underenrollment that required a recalculation in sample size. The lack of uniformity in the sampling of the population of interest represents a potentially important confounder because it generates a heterogeneous sample, and that heterogeneity will be translated to the results, limiting interpretation.

As with many other dilemmas in clinical research, the tension between ethical considerations and scientific rigor is ever present. In their letter to the editor, the authors state that according to the trial’s pragmatic nature performed...
predominantly on patients traditionally regarded as not meeting criteria for rt-PA, they expected modest results.

Q1.b The investigators defined an “experienced” stroke center as one that treated at least 3 patients with rt-PA in the preceding 12 months before joining the trial. What do you think about this definition? What percentage of patients were enrolled at experienced centers? Consider also the temporal distribution of patient enrollment in the study period. How might these factors affect internal and external validity?

Given the conventional perceptions of the term “experienced stroke center,” the criterion of having treated 3 patients in the previous year with a protocol for open-label treatment for rt-PA is likely to be surprising to most readers. Can a center with an organized system for stroke care be considered truly experienced if they treat only 3 patients per year with rt-PA? The reader must decide whether such limited experience in a given site and large variations in experience between sites might result in differential outcomes between institutions and regions. To inform this decision, it would be useful to know what percentage of sites met the minimum criteria. The article does not report the individual sites’ experience classification. Inexperience can be a threat to internal validity by introducing systematic bias in several ways, including protocol violations before, during, and after the administration of rt-PA. In a trial this diverse and expansive, it would be difficult to identify all protocol violations that may have occurred at the various sites. If inexperienced sites contribute disproportionately to poor outcomes for the treatment group, then experience (or lack thereof) may have confounded the results toward the null hypothesis (favoring the control).

In the answer to question 1a, we discussed the potential bias that may result from the inherent subjectivity in the application of the uncertainty principle. For sites with experience close to the 3-patient minimum criterion, it is likely that the majority of acute care physicians who enrolled patients had little to no consistent experience with the pretreatment patient interaction, including consent. Their inexperience will also increase the variability in the application of the uncertainty principle in that they are more likely to be uncertain in comparison to more experienced physicians. Inexperience in medicine usually translates to discomfort rather than enthusiasm. A notable exception is the physician experienced in stroke thrombolysis who encounters a fatal hemorrhage and loses confidence. However, among those who encounter fatal bleeding, those with experience are more likely to put the bleeding in perspective compared with those who have little experience. These imbalances are likely to manifest as sampling bias in the prerandomization eligibility phase, in which limited experience may be an important threat to recruitment.

The trial is described as international and multicenter. At face value, the expansive general design lends itself well to the intention of improving external validity and increasing treatment eligibility. The implication of such a design is that the results, at the very least, can be applied to the involved centers and additionally to other similar centers internationally. According to the recruitment, are the various regions equally represented? In their 2011 update, the authors display the recruitment by country (save the designation of “United Kingdom”) in Table 1. Almost 50% of all subjects were enrolled in the United Kingdom at 75 centers. Approximately 90% of all subjects were enrolled in Europe and nearly 96% were enrolled in Europe and Australia. The regional list in Table 1 in IST-3 groups countries by region, including “Australasia,” though no Asian countries are listed in Table 1 of the 2011 update, in which each country is listed individually. For readers to interpret this international distribution accurately, they must consider the relative representation of different regions of the world. Regarding external validity and generalizability of the results, the reader should consider differences in stroke care and acute treatment protocols in the site of interest, with standards in Europe and Australia.

The trial was conducted during an 11-year period, resulting in temporal variations. During that time, the practice standards of a novel therapy should be expected to evolve because additional data and experience alter such practice. Official government approval and licensing for rt-PA dramatically affected the trial when the design was changed from a double-blind, placebo-controlled trial to an open-label trial. The previously discussed subjectivity in the enrollment criteria is also subject to temporal variation. The physician and patient decisionmaking and experience at the outset of the trial are likely to change during the 11 years as individual and institutional experience and opinions about stroke thrombolysis evolve. For physicians, new data from positive-outcome trials such as the European Cooperative Acute Stroke Study (ECASS)-3, which extended the treatment window to 4.5 hours, and other observational studies would serve to bolster confidence during the decade and alter enrollment. Thus, the characteristics of the patients enrolled at the beginning of the trial are likely to vary in important ways from those enrolled toward trial completion, creating a heterogeneous sample.

Q1.c After enrolling 276 patients, IST-3 switched from a double-blind trial to an open-treatment investigation. Control group patients in the open-treatment arm did not receive a corresponding bolus or infusion. Describe the placebo and nocebo effects and how they might impact study results. What outcomes are more likely to be exaggerated by the lack of a placebo?

The nocebo effect is the causation of sickness or death by the expectation of sickness or death. The placebo effect is the causation of improvement or wellness by the expectation of improvement or wellness. The placebo and nocebo effect have 3 distinct components, the expectation of the patient, the expectation of the physician or healer, and the expectation that results from the relationship between both patient and healer. The expectation can be relative to the patient’s beliefs about the disease in the absence of a treatment or when a treatment is administered.
Speigel described how the placebo effect has been demoted in modern medicine to the role of a contaminant in research. For ethical reasons, less is known about the nocebo effect because its study would require generating negative expectations, which are expected to generate worse outcomes. Placebo and nocebo effects have been found to be greater when the outcomes are continuous and subjective, as in the case of pain treatment. Several recent meta-analyses have demonstrated conflicting results, with the placebo effect ranging from 7% in a meta-analysis by Hróbjartsson and Götzsche compared with 20% in a review by Wampold et al. Variability in the effects described in these reviews likely results from the types of diseases and outcomes studied. Studies with patient-reported outcomes had the largest effect.

The initial pilot phase of IST-3 was a double-blind trial, but the design was modified to an open label after Boehringer Ingelheim’s wish to stop supplying drug and placebo. Recruitment for the expansion phase began in August 2003, the same year that rt-PA was approved for use in Europe. The use of a placebo control in a trial is to blind the subject and investigator, if double blinded, of treatment assignment and balance the systematic bias introduced when outcomes are affected by subject expectation. Without blinding and with full knowledge of the rt-PA administration (versus no treatment), the subject’s positive and negative expectations with regard to rt-PA may result in a larger corresponding positive (placebo) or negative (nocebo) effect on the outcome, which is unrelated to the drug itself. With the loss of a true placebo arm, we are unable to estimate the true treatment effect. This is a substantial threat to the trial’s internal validity. The subject’s expectations are unprotected from the expectations of all the various providers involved in his or her care, as well as family members. Their expectations are subject to myriad cues given by all nonblinded parties whether they are subtle, obvious, subconscious, or conscious. If the subject, either because of an individual bias or a cultural bias, has a positive expectation of rt-PA, the treatment effect is likely to be exaggerated by placebo effect and outcomes expected from the natural history of stroke in the no-treatment arm will be worse because of the nocebo effect. Some may argue that subjects with negative opinions of rt-PA may balance these effects; however, eligibility for enrollment was based on the explicit patient impression that rt-PA represented a “promising, yet unproven” treatment and therefore those with negative opinions and expectations are unlikely to be recruited. This likely leaves the described placebo/nocebo-related bias, as well as reporting bias based on treatment.

Also, the selection of very enthusiastic patients may introduce a sampling bias before randomization, and when considered in the context of an open-label trial, this may result in underreporting of neurologic deficits by patients who received rt-PA and overreporting by those who received placebo. This will result in overestimation of effect by creating a positive bias in the treatment group and a nocebo effect in the no-treatment group.

Q1.d Describe potential sources of bias present in the IST-3. Do these sources of bias favor the treatment arm or the control arm?

1. Selection bias favors treatment: The uncertainty principle applied in IST-3 enrollment necessitates that treatment with rt-PA not be clearly indicated or contraindicated according to the European license, as well as the subjective belief in potential benefit (“promising but unproven”) on the part of the physician and patient. Many patients who were eligible for rt-PA were excluded from the trial on this basis; however, the authors do not report the individual reasons that the patients were not enrolled. Had they done so, the reader would be able to determine how many were excluded because of an unfavorable view of the drug by the physician or patient. Such subjectivity is a major potential source of bias because physicians among institutions, countries, cultures, and continents are likely to have variable opinions of the efficacy of rt-PA. A strong bias (unconscious or not) may result in a selection bias, were the study to enroll only those patients who view the drug most favorably.

2. Nonblinding favors treatment: More than 90% of enrollment was in the nonblinded phase and lacked a placebo control. When a study omits a placebo control and is not blinded, the potential for overreporting is significant. The positive bias in favor of the treatment on the part of the physician and patient is likely to be transferred to the reporting of outcomes by patients receiving rt-PA. For those in the control group, the nocebo effect may result in patients overreporting poor outcomes, further threatening internal validity by increasing the difference between reported treatment and control outcomes.

3. Subjective outcomes favors treatment: Changing the outcome of interest from objective (mortality) to subjective (self-reported function) threatens internal validity in that outcome classification can vary with individual application and interpretation. This is further compounded by the bias introduced by the design feature that included mail-in follow-up, in which patients mailed in their self-assessments. Subjective outcomes combined with no placebo and nonblinding results in a significant potential for bias that favors the treatment group, given the selection bias described above. Follow-up using self-reported “mail-in” further magnifies the reporting bias for treatment and against control because of the lack of trained data collectors who may have otherwise been able to limit biased reporting.
4. Inexperience bias could favor either treatment or control: We discussed earlier the bias introduced by the potential for relative inexperience in the definition of “experienced centers.” This relative inexperience and lack of consistent opportunities for individual providers to enroll patients within the expectations of the uncertainty principle would increase the variability in both enrollment and treatment. If the inexperience translates to discomfort and insecurity with rt-PA, the physician will tend to enroll only the most enthusiastic patients, favoring the treatment group. If physician inexperience translates to favorable opinions about rt-PA because no bleeding events were observed, then the physician may enroll a more diverse, higher-risk sample, favoring the control. Institutional and provider inexperience may also lead to delays and other protocol violations that favor the control group. The important concept here is that such inexperience may lead to a lack of uniformity in enrollment and treatment between individual physicians and institutions.

5. Conflict of interest favors treatment: Most reputable journals have made important progress in disclosure of relationships that may result in a conflict of interest. The article footnote details the financial and other conflicts of interests for the authors. However, simply listing the conflicts does not eliminate the potential for bias. The reader must appreciate this potential. There is no simple adjustment that can correct for such bias. Even reviewers without financial conflict may be conflicted by other personal and professional bias that cannot be classified easily.

The additive or, worse, synergistic effect of the subjective outcome bias, selection bias in favor of treatment inexperience bias, reporting bias, and nonblinding bias could represent a very strong cumulative bias toward rejecting the null hypothesis favoring the treatment. The results of IST-3 are neutral before taking into account the biases favoring treatment. The reader must assess the effect of these biases for overestimation of treatment effect. If the exaggeration of effect is great enough, the seemingly neutral results may actually be masking a negative trial result.

ANSWER 2

2. The authors conclude "for the types of patients recruited in IST-3, despite the early hazards, thrombolysis within 6 h improved functional outcome." 11

Q2.a Is this statement consistent with the results of the primary outcome measured in this study? On what evidence do the authors base this conclusion?

The authors’ primary conclusion ignores the primary outcome. The prespecified primary outcome was the proportion of patients alive and independent at 6 months, and no significant difference was found between the rt-PA and control group (37% versus 35%, respectively; P = .18). IST-3 is by far the largest randomized controlled trial of thrombolytics for stroke to date. Very large trials are at risk of revealing minor benefit that may reach statistical significance but be of no clinical meaningfulness. 10 The inability of rt-PA to demonstrate even statistically significant benefit for the primary outcome would seem to warrant top billing as the authors’ conclusion. Rather, the authors chose to focus their conclusion on the results of a secondary outcome, the ordinal analysis. This analysis revealed a favorable shift in the Oxford Handicap Score (OHS) for the rt-PA group and is the evidence on which IST-3 can be seen as a positive-result study. Further discussion of the ordinal analysis can be found in answer 2c.

A basic tenet of clinical research is to establish and respect a primary outcome. 11 A primary outcome should be fundamental to both the development of methods and the drawing of conclusions. In the conception of a trial, the primary outcome is what the investigators deem to be the most important answer worth seeking. Sample size is calculated according to the primary outcome, and studies are uniquely powered to investigate the primary outcome. Focusing on a primary outcome diminishes the potential for data snooping bias, or dredging and presenting data that suit the investigators’ partiality. 12 Evaluating multiple secondary outcomes increases the risk of reaching inappropriate conclusions. For example, in a trial with an intervention that is truly no different from the control, if there were 20 different outcome measures, there is a 64% chance that one will be statistically significant if the α level for each outcome test is set at .05.

Q2.b The authors allude to "early hazards" after treatment with rt-PA. There was a 1.59 (95% confidence interval 1.23 to 2.07) unadjusted increased odds of death within the first 7 days after treatment with rt-PA, but this effect disappeared at 6 months. Hypothesize possible underlying mechanisms for the disappearance of this difference over time. What data acquired in IST-3 might allow limited retrospective exploration of these hypotheses?

Within the first 7 days of treatment, the absolute risk of death was 3.8% higher for patients in the rt-PA group compared with control, for a number needed to harm of 1 in 26. However, by 6 months, 27% of patients had died in both groups. The transience of the mortality discrepancy may mitigate safety concerns. Early death is compensated for by less death during the next several months. The authors comment on the importance of longer-term follow-up and report that further study in the United Kingdom, Norway, and Sweden continues. These data may help provide clarity on the hypothesis that improved functional outcome results in improved survival beyond 6 months. However, the difference in early mortality raises serious concerns. Although death at 6 months was equivalent, other statistical measures are likely more revealing than percentage dead at a single point. The authors do not provide survival curves (such as in the National Institute of Neurological Disorders and Stroke Study Group Part 2) that would likely graphically demonstrate harm for rt-PA across most of the 6-month interval. 13 Other analyses such as the Cox proportional hazards model estimate temporal benefit or harm and deliver a more insightful evaluation of the rate and timing...
of events than a single-point measure.\textsuperscript{14} Quite simply, high early mortality suggests a significant loss of days alive for patients in the rt-PA group during the study period.

One should also consider that some approximation of mortality curves might be expected as patients age, regardless of the intervention. When an intervention has an immediate effect on mortality and then no further influence, the discrepancy will eventually wane because of the inevitability of death. After the immediate mortality effect, the absolute difference in number of survivors does not remain constant; rather, a similar percentage of patients in each group will die during each measured period, and the mortality discrepancy will progressively decrease. This effect is most pronounced with older populations and high natural mortality rates. The IST-3 population was old, with 57% older than 80 years, and the 6-month mortality rate was considerable, at 27%.

We can make a rough estimate of the anticipated approximation of mortality curves if rt-PA had no mortality effect after 7 days. Hypothetically, let us assign a mortality rate of 20% in both groups for the period 7 days to 6 months.

The control group began with 1,520 patients, but within 7 days 107 (7.0%) had died, leaving 1,413. With a subsequent mortality rate of 20%, 1,131 survivors would remain, or 74.4% of the original control group.

The rt-PA group began with 1,515 patients, but within 7 days 163 (10.8%) had died, leaving 1,352. A mortality rate of 20% would leave 1,081 surviving patients, or 71.4% of the original rt-PA group.

The 3.8% mortality discrepancy at 7 days would shrink to 3% by 6 months simply because the same percentage of patients was subtracted from different-sized groups. This does not discount that the mortality rate from 7 days to 6 months was lower in the rt-PA group, but some reduction of the mortality discrepancy would have been observed even if it were not.

Why was the mortality rate higher in the control group after 7 days? Two simple theories are as follows: (1) rt-PA provided benefit (eg, rt-PA resulted in improved functional status and hence fewer deaths between 7 days and 6 months); and (2) early intracerebral hemorrhage and death occurred in patients with more severe strokes, leaving a healthier rt-PA group after 7 days.

It is well described that the risk for early intracerebral hemorrhage is higher in patients with more severe strokes.\textsuperscript{15} Unfortunately, data such as functional status at 7 days were not available in IST-3, limiting our ability to draw conclusions about longer-term mortality effects.

As we further analyze associations in the IST-3 data set, the thoughtful reader must contemplate causal links. The authors provided data on fatal and nonfatal events within 7 days in Table 3. Intracerebral hemorrhage accounted for most of the mortality discrepancy between groups. The abundance of previous evidence describing this complication, our understanding of the pharmacodynamics of rt-PA, and the proximal relationship in time strongly suggest a causal mechanism of death.\textsuperscript{16}

Is the causal link as strong between rt-PA and the nullification of the mortality discrepancy at 6 months? In defense, the strength of association is powerful, meaning it is unlikely the results were due to chance. Improved functionality resulting in improved survival is certainly biologically plausible. The temporal relationship is rather consistent with timeframes of benefit observed in previous studies.\textsuperscript{12} However, it is reasonable to argue that the causal link is weakened with greater temporal distance between the intervention and effect. Ambiguity is suggested by concepts described above (anticipated approximation of mortality curves in elderly patients and the likelihood of a healthier rt-PA group after 7 days). Moreover, data are not available on cause of death during the 6-month period as they are for the 7-day period. We are left to wonder who the patients were who died. Were all excess deaths in the control group a result of neurologic disability because of untreated (with rt-PA) stroke? More than half of the patients in IST-3 were older than 80 years, a population at risk of dying for a host of reasons independent of whether they received a certain medication 5 months before. Understanding the nature of death (pneumonia, myocardial infarction, hit by bus, etc) for the 6-month period as plainly as for the 7-day period would help refine causal suppositions. Although appropriate randomization and sample sizing should have resulted in equal distribution of unrelated deaths, these data are not available for our consideration.

Q2.c The authors include an "ordinal shift analysis" as a secondary outcome in IST-3. Describe the advantages of shift analysis over a dichotomous endpoint in terms of information theory. What limitations might shift analysis have compared with a dichotomous endpoint? Should shift analysis increase or decrease power to detect a difference between treatment arms?

Despite being a secondary outcome, the ordinal analysis is the centerpiece of IST-3. Shift analysis does not refer to the shifting of focus from the primary outcome to a secondary outcome; rather, it is an analysis of the overall distribution of ordinal scores in the intervention and control groups and the generation of an odds ratio to describe the direction of shift toward or away from better functional health. The conclusion that rt-PA is beneficial is based on this analysis. However, in a previous article, the authors stated that an ordinal analysis was not universally applicable and that underlying assumptions necessary for validity would not be met with the IST-3 data set.\textsuperscript{5} They concluded that the ordinal analysis was not appropriate for the primary outcome but would be included as a secondary analysis.

Mortality and disability are the most important outcomes for stroke. Mortality is plainly a dichotomous measure (dead/alive). Disability can also be described in a binary manner, such as good/bad, but much of the detail is lost with this breakdown. The OHS is a variant of the modified Rankin Scale, both of which are ordinal, meaning they have greater than 2 categories of a natural ordering. For the primary outcome, independence was defined dichotomously (OHS=0 to 2), but for the ordinal
analysis, a finer assessment of disability was made, with OHS levels 0, 1, 2, 3, and 4 to 6 registering as discrete outcomes. Commonly, outcomes of persistent vegetative state and death are collapsed together because some physicians and patients consider them equally devastating, but it is unclear why in IST-3 the OHS level of 4 (defined as unable to live independently but does not require constant attention) was grouped with 5 and 6 in the shift analysis.12

Stroke is a disease that affects cardinal aspects of behavior and function, and it is reasonable for disability measures to recognize this complexity. The purported advantage of shift analysis is that point-to-point benefits can be appreciated. For example, shift analysis can discern differences between an OHS level of 2 and 1 or 1 and 0, whereas a dichotomous measure with a cutoff between 2 and 3 could not. Shift analyses are said to be more efficient, that is, they can detect signals across all ordinal transitions rather than just at a single dichotomous point.18 Others have argued that the error inherent in classifying patients at a specific ordinal level decreases the efficiency of the metric and necessitates even larger sample sizes to demonstrate effect.19 Numerous studies have demonstrated that the modified Rankin Scale (and presumably OHS) has poor reliability, particularly in differentiating patients with middle-range scores.19,20 These studies involved neurologists scoring patients by direct examination, whereas in IST-3 the scores were mostly garnered by telephone or postal questionnaire from nonblinded laypersons, threatening the reliability and validity of the outcome measurement.

Another purported advantage of shift analysis is greater interpretability.3 Odds of favorability can be calculated and converted to patient-oriented outcomes such as number needed to treat. However, ordinal shift analysis is complicated, and making sense of it for practicing physicians and patients can be troublesome. In IST-3, the ordinal shift demonstrated benefit for rt-PA, with an adjusted odds ratio of 1.27. How do we translate this into valuable information that can be communicated to our patients? It is much simpler to describe dichotomous results, such as the risk of death or the chance of a good outcome. With ordinal analysis, the difficulty is not necessarily in communicating the odds of favorable shift but in quantifying favorability. That is, patients deserve to know not just the chances of an outcome but also what exactly the outcome is, what “shift” means in tangible terms. Shift does not refer to the chances of transition from one ordinal measure to another (eg, improvement of OHS 3 to 2) or the chances of one group achieving one ordinal measure better than the other. Rather, the outcome is a ratio of the odds of “favorability” along the ordinal scale for the intervention group compared with the control. With IST-3, although the odds of favorable shift may be quantified, the benefit itself is an ambiguous concept of unclear clinical significance.

ANSWER 3

3. The editorial accompanying the study publication in the Lancet concludes with the statement “. . . the default situation for the first health-care professional who identifies the stroke patient should be to treat, and the role of stroke and emergency physicians is now not to identify patients who will be given rt-PA, but to identify the few who will not.”

Q3.a This statement is based on the results from IST-3 and the accompanying updated meta-analysis of rt-PA for acute stroke.3 Does IST-3 provide evidence that only a “few” patients will not benefit from rt-PA? Concisely summarize the findings of this study in the context of evidence from previous stroke trials.

To correctly conclude that only a “few” patients will not benefit from rt-PA, the implication is that the remaining majority of patients universally benefit from treatment. For this assertion to be valid, a reasonable expectation of the treatment in question is that it otherwise results in a clinically important improvement in its intended population.

To address the universality aspect, it is important to recognize that the treatment population presenting with acute stroke is heterogeneous, with varying comorbidities and underlying pathophysiologic mechanisms. Reflecting this heterogeneity, it is generally accepted that certain constellations of underlying clinical features result in a decreased likelihood of benefit or a preponderance of harm. Several clinical decision instruments have been developed in an attempt to prognosticate the outcomes of thrombolysis in acute stroke, with a goal of better identifying these patients who will not benefit.21,22 The findings in IST-3 do not contradict this previous evidence because a cursory review of Figure 3 from the original article suggests multiple clinical factors associated with trends favoring the control arm.

Contemporary treatment of acute stroke with rt-PA is based on evidence accumulated from stroke trials during the last 20 years, the most prominent of which include the National Institute of Neurological Disorders and Stroke and the ECASS-3.13,23 To examine the findings of IST-3, pooled into the greater context of previous stroke trials, an updated meta-analysis including 12 trials of rt-PA is in same issue of Lancet.3 When IST-3 is added to this meta-analysis, the general effect is a regression toward neutrality. Although most outcomes remained unchanged from a statistical significance standpoint, the previously significant treatment effect favoring thrombolysis for patients treated beyond 3 hours is no longer present.

Considering that IST-3 enrolled nearly as many patients as all previous stroke trials combined, this regression casts further doubt on the universality of benefit for thrombolysis in acute stroke.

A counterargument proposed by stroke neurologists to support a claim for universality of benefit is based on the ordinal shift analysis. Their interpretation of shift analysis is that the results better retain information otherwise discarded by a dichotomous endpoint (eg, modified Rankin Scale or OHS 0 to 2) and better demonstrate consistent benefit across the entire spectrum of outcomes.24 However, the small absolute differences in ordinal outcomes indicate that these benefits, even if universally present, are not clinically important.
At face value, therefore, it is inaccurate to state that thrombolysis universally results in a clinically important benefit. The evidence is lacking from IST-3 and the meta-analysis to support the quoted statement. A more reasonable conclusion might rather suggest only a “few” may conclusively benefit, and the “default situation” should be continued vigilance to prevent avoidable harms.

Q3.b Extensive conflict of interest disclosures are reported for study authors. What effect might these disclosures have on clinicians when they interpret study results? What biases have been demonstrated as occurring in studies whose authors have substantial conflicts of interest?

Conflict of interest in medical literature is pervasive but has yet to demonstrate conclusive, consistent effects. Several small survey studies have evaluated the effect of conflicts of interest on physician perception. A survey among readers of the British Medical Journal found that respondents were significantly more likely to downgrade the importance, relevance, validity, and believability of a clinical trial when conflicts of interest were reported. A survey of American College of Gynecology physicians rated conflicts of interest as an important factor in evaluating clinical trials but also found that the same physicians did not consistently account for conflicts of interest when rating a sample article. Finally, a survey of American Board of Internal Medicine–certified physicians indicated that conflicts of interest resulted in a decrease in perception of trial rigor and in the confidence of the reported results. These data suggest that physicians understand the importance of recognizing the influence of conflicts of interest in published literature, but its resultant downstream effect is less certain.

The survey of American Board of Internal Medicine physicians was accompanied by an editorial titled “Believe the Data.” This editorial, written by the editor-in-chief of the journal, suggested that the lack of trust in industry-funded results was unwarranted. This standpoint may be surprising in light of findings from 2 systematic reviews indicating that industry-funded studies had increased odds ratios of 3.60 and 4.05 for results favoring the funding source. These data suggest that physicians understand the importance of recognizing the influence of conflicts of interest in published literature, but its resultant downstream effect is less certain.

Q3.c A patient presents to your emergency department with an acute stroke that was noted nearly 5 hours before arrival. A noncontrast computed tomography scan of the head shows no hemorrhage or signs of recent ischemic change, and the patient has no important vital sign or laboratory abnormalities. The family at bedside asks whether the patient is eligible for “clot-busting treatment.” Describe the informed consent you might offer the patient and family for rt-PA in the context of the results of IST-3.

The generally accepted components of informed consent for a patient or surrogate with capacity to make medical decisions include:

- the nature of the intervention;
- the reasonable alternatives to the intervention; and
- the relevant risks, benefits, and the uncertainties related to each alternative.

The nature of the intervention for rt-PA is systemic thrombolysis with the goal of reperfusion of the ischemic tissue. The proposed mechanism by which this plausibly results in physiologic benefits is salvage of a still-viable ischemic penumbra around any irreversible damage. In most current settings, the only alternative to rt-PA for acute stroke is to withhold that specific treatment. Therefore, the discussion of relevant risks and benefits typically contrasts only these 2 options.

The benefit from rt-PA is generally reported in clinical trials as improvements in long-term disability compared with placebo, whereas the harms are associated with increased intracranial hemorrhage and death. According to data from thousands of patients in clinical trial and registry data, several prognostic instruments have been produced. An analysis from the Virtual International Stroke Trials Archive registry produced a nomogram with an area under the curve of 0.808 for prediction of functional recovery and an area under the curve of 0.706 for prediction of survival at 3 months. For patients receiving rt-PA, the DRAGON score (dense middle cerebral artery sign or early infarct signs on admission CT head scan, prestroke modified Rankin Scale score > 1, Age, Glucose level on admission, onset-to-treatment time, and NIHSS) has been developed to predict favorable and unfavorable outcomes at 3 months. Another decision instrument, the iScore, risk-stratifies patients receiving rt-PA and provides a model estimating the likelihood of a favorable outcome for treated patients relative to placebo. These tools, as well as future developments and refinements, should be considered for use at the bedside as
decision instruments to augment clinical judgment and to provide a numeric context with which to educate patients and families.

A sample template for a discussion of the relative risks and benefit for rt-PA in this patient might take the form of the following:

“According to our clinical assessment, we believe your family member has experienced a stroke. A stroke occurs when a blood vessel supplying a portion of the brain is blocked. This results in a portion of the brain dying because of a lack of blood supplying oxygen.”

“A medication for use in treating acute stroke is approved in the United States and Europe. This medication works by attempting to break up the blockage, restoring blood flow and oxygen to the brain. However, the approval for the use of this medication does not extend up to 5 hours. We do, however, use this medication off license under certain circumstances.”

“If we do not use this medication, using the best data available, we roughly estimate your long-term chance of requiring assistance with daily living to be [estimate #1] and your chance of being alive in 3 months to be [estimate #2].”

“The benefit on which approval of this medication is based is that patients who receive it are less likely to be severely disabled. However, not all patients are likely to benefit from its use. According to what we know from previous research trials, your [age, computed tomography findings, glucose level, time to treatment, stroke severity, preexisting disability, heart disease, kidney disease] makes you [more/less] likely to benefit. The risk of this medication is that its use may result in additional bleeding complications, approximately half of which are fatal. Overall, our best estimate for your long-term chance of requiring assistance with daily living is [estimate #1a], and your chance of being alive at 3 months, with the additional bleeding complications, is [estimate #2a].”

“At this institution, we provide excellent stroke care, including therapy and rehabilitation, regardless of whether we use this medication or not.” Based generally on the available evidence, this patient would not be likely to benefit from rt-PA. Although the IST-3 subgroup analysis indicates benefit in the 4.5- to 6-hour subgroup, this is inconsistent with the time-to-treatment effect proposed by other literature and the basis of modern stroke care using rt-PA. The accompanying meta-analysis incorporating the IST-3 results reports no statistically significant benefit associated with rt-PA use beyond the 3-hour mark. The IST-3 results suggest several baseline factors associated with notable trends favoring rt-PA. However, further prospective investigation of these associations, controlling for confounding factors, is necessary before their inclusion in discussions pertaining to informed consent.

Section editors: Tyler W. Barrett, MD, MSCI; David L. Schriger, MD, MPH

REFERENCES


Volunteer for a Committee and Lend Your Experience and Expertise

ACEP and Emergency Medicine Need Your Assistance

The process to select members to serve on national ACEP committees is beginning and all ACEP members are encouraged to apply.

EMRA members who are interested in serving as that organization’s representative on an ACEP committee should also apply. The process is the same for resident and active members and you can expedite the process by using the online application.

If you are not currently serving on a national ACEP committee you must submit a current CV to volunteer for a committee. You can either attach the file to the online form or mail it to ACEP headquarters. You may also want to submit a letter of support from your chapter. Members who do not know how to contact their state chapters should call Dawn Scrofano, Chapter Services Manager, at 800-798-1822, ext. 3227, or send an e-mail to dsfano@acep.org. The online committee interest form is available on ACEP’s Web site at http://webapps.acep.org/Membership/committeeinterest.aspx

Although most committee work will be accomplished through email and conference calls, committee members are expected to attend the organizational meetings at the Scientific Assembly in Seattle October 14-17, 2013.

Committee interest must be submitted by May 17, 2013. If you have any questions, please contact Mary Ellen Fletcher, CPC, CEDC, at 800-798-1822, ext. 3227, or send an e-mail to mlfletcher@acep.org. Alex M. Rosenau, DO, CPE, FACEP, ACEP’s President-Elect, will finalize the committee appointments in June.

Remember, your participation will make a difference. Please consider volunteering. ACEP and emergency medicine need your experience and expertise.