

September 22, 2026

Grace R. Graham, Deputy Commissioner for Policy, Legislation, and International Affairs, FDA
Scott N. Goldie, General Health Scientist (Special Assistant), CDER, FDA
Philip Kurs, Regulatory Counsel/Senior Policy Advisor, CBER, FDA

RE: ACRO comment submission on FDA revised draft guidance on
Demonstrating Substantial Evidence of Effectiveness for Human Drug and Biological Products
[Docket No. FDA-2019-D-4964]

Dear Ms. Graham, Mr. Goldie, and Mr. Kurs,

Founded in 2001, the Association of Clinical Research Organizations (ACRO) is a non-profit trade association representing the world's leading clinical research and technology organizations, which provide specialized services that are integral to the development of drugs, biologics, and medical devices that enable patients to live longer, healthier, and more productive lives. ACRO members provide a wide range of specialized services across the entire spectrum of drug development—from preclinical, proof of concept, and first in human studies through post-approval, pharmacovigilance, and health data research. ACRO member companies employ more than 470,000 people worldwide and conduct research in every global region.

ACRO thanks the Agency for this opportunity to provide feedback on the draft guidance. In addition to Section One, which provides feedback in the form of general comments, and Section Two, which includes ACRO's line-specific comments, ACRO has also included a third section in our comment to FDA where we propose that the final guidance include a glossary of important terms as an addendum to the guidance.

The draft guidance introduces valuable, impactful streamlining and flexibility of clinical research by outlining and expanding upon the evidentiary conditions for utilizing *one adequate and well-controlled clinical investigation plus confirmatory evidence* to establish substantial evidence of effectiveness. However, some of the key terminology in the draft guidance is unclear, which could result in sponsor hesitance to pursue this option in their overall drug development programs. We believe an addendum glossary could provide useful precision without sacrificing the welcome flexibility emphasized throughout the draft guidance.

Section One: General Comments

ACRO recommends that the FDA Establish an Explicit Sliding-Scale/Totality-of-Evidence Framework

In the draft guidance, FDA states that the stronger the single trial, the less strong the confirmatory evidence may need to be. Specifically, the guidance presents two scenarios: “strong confirmatory evidence + adequate trial” versus “highly persuasive trial + early-phase confirmatory evidence.” In reality, development programs will frequently fall between those two scenarios. For example, moderately/highly persuasive pivotal trial + moderately strong mechanistic + Phase 2 + PK/PD evidence. The guidance is not clear whether that would be sufficient.

ACRO recommends that the FDA add a matrix illustrating acceptable combinations such as the following:

Pivotal Evidence	Confirmatory Evidence	Sufficiency for Substantial Evidence
Highly persuasive	Moderate/early phase	Potentially sufficient
Strong	Strong	Potentially sufficient
Moderate	Very strong	Potentially sufficient
Weak/inconsistent	Strong	Generally insufficient

ACRO recommends a prospective Mechanism for FDA Advice to Facilitate Sponsor Confidence in Pursuing a Single-Trial Strategy

In the draft guidance, the Agency recommends discussing a single-trial strategy before initiating the pivotal trial. ACRO recommends that sponsors be able to submit a “Confirmatory Evidence Plan” at the End-of-Phase 2 describing:

- Proposed pivotal evidence
- Sources of confirmatory evidence
- Rationale for independence/complementarity of the evidence
- Proposed statistical strength
- Biological plausibility
- Evidence integration
- Contingencies if the pivotal result is less persuasive than anticipated

Having the option for sponsors to submit a "Confirmatory Evidence Plan" in a structured process, as part of the End-of-Phase 2 meeting package, with formal timelines for FDA response, could materially reduce late-stage regulatory uncertainty for sponsors and the FDA. The FDA would have the opportunity to provide advice regarding whether the planned evidentiary package, if successful as proposed, could support substantial evidence of effectiveness.

We ask the FDA to consider incorporating Real-world Data (RWD) and Evidence (RWE) Concepts

FDA should consider adding a dedicated section on RWD/RWE that provides clearer guidance on the use of external comparator studies, natural-history studies, and other non-interventional evidence sources in supporting substantial evidence of effectiveness. Such a section could clarify expectations for data feasibility assessments, prospective evidentiary planning when leveraging pre-existing observational data, and the distinction between exploratory and confirmatory uses of RWD, which remains an area of regulatory uncertainty. Greater clarity in these areas would improve predictability, facilitate efficient development planning, and support consistent application of real-world evidence across development programs.

Section Two: Line-Specific Comments

Lines 47-52:

This section of the draft guidance currently reads:

Sponsors should discuss their anticipated approach to demonstrating substantial evidence of

effectiveness with FDA early in development, such as at a pre-investigational new drug application meeting, and no later than at an end-of-phase 2 meeting. When meeting with FDA, sponsors should provide a scientific justification for their proposed approach. Ultimately, whether substantial evidence of effectiveness has been demonstrated will depend on FDA's review of the marketing application.

This section of the draft guidance does not describe the degree of prospective regulatory feedback sponsors can expect. To provide this much-needed guidance ACRO recommends incorporating the following new language into the final guidance:

For programs proposing reliance on one adequate and well-controlled clinical investigation plus confirmatory evidence, sponsors are encouraged to present a prospective evidentiary plan describing the proposed investigation, sources and expected strength of confirmatory evidence, and how the evidence will be integrated.

FDA intends to provide advice regarding whether the proposed approach, if successfully executed as planned, would be reasonably capable of supporting substantial evidence of effectiveness. Such advice does not constitute a prospective approval determination.

In addition, it would also be helpful for the final guidance to address when sponsors should re-engage FDA if material elements of the evidentiary strategy change after the initial discussion, for example, the pivotal trial design, patient population, endpoint, statistical approach, or source of confirmatory evidence. This would make the early-engagement recommendation more actionable and help avoid late-stage misalignment.

Line 108:

The phrase “quality of trial conduct” is first used here:

The strength of evidence of effectiveness can be impacted by many factors, such as key clinical trial design elements (e.g., control, population selection, randomization, blinding, endpoint), quality of trial conduct (e.g., completeness of follow-up), aspects of the trial analysis (e.g., prespecification, analysis assumptions), the clinical and statistical persuasiveness of trial results, and aspects of the overall drug development program.

We ask the FDA to insert a footnote next to this first instance of the phrase “quality of trial conduct” to note that the Agency’s understanding of “quality” aligns with ICH GCP E6(R3) and its risk-proportionate/critical-to-quality framework.

Lines 185-188:

The draft guidance currently states:

There are also many other important design elements that can impact the strength of evidence generated in a trial. These include, for example, the degree of blinding to treatment assignment, choice of drug dosage, selection of clinical trial sites, use of concomitant therapies, trial duration, sample size, adaptive design elements, and approach to ascertain outcomes in participants.

Clarification would be valuable for trials designed to evaluate both an overall population and a pre-specified biomarker-defined subpopulation, particularly with respect to adaptation rules, multiplicity control, Type I

error control, and interpretation of treatment effects following population enrichment. Adaptive enrichment designs introduce statistical complexities that are not present in traditional fixed designs and would become particularly critical in a single pivotal trial.

Additional guidance regarding adaptive enrichment designs in the context of a single pivotal trial supporting substantial evidence of effectiveness would be valuable. Illustrative examples would help sponsors understand FDA's expectations for the use of such designs within a single-study development strategy.

Lines 192-200:

This portion of the draft guidance currently states:

The quality of trial conduct also impacts the strength of evidence generated by a trial. For example, this includes the maintenance of blinding of trial participants and investigators to treatment assignment in a double-blind design, the maintenance of confidentiality of interim results while the trial is ongoing, the quality of data collection, the extent of adherence to treatment and to the protocol, and the completeness of follow-up of participants, including the amount of and reasons for missing data on key endpoints. Clinical trials should be conducted in accordance with Good Clinical Practice (GCP) standards. Poor execution can render a trial of any design not adequate or not well-controlled and, therefore, unable to contribute to substantial evidence of effectiveness.

"Quality" and "poor execution" are currently insufficiently defined and could imply that routine operational deviations jeopardize an entire trial. To avoid such misinterpretations of FDA's intent, ACRO recommends incorporating the following new language into the final guidance:

The impact of trial-conduct issues should be evaluated based on their nature, magnitude, distribution across treatment groups, and potential to introduce bias or materially affect interpretation of the prespecified treatment effect. Routine operational deviations that do not meaningfully affect the integrity or interpretability of trial results would not ordinarily, by themselves, render an otherwise adequate and well-controlled trial inadequate.

Lines 209-222:

This section of the draft guidance currently states:

There are a variety of important statistical considerations in ensuring that trials are adequate and well-controlled and in evaluating the strength of evidence of effectiveness. An adequate and well-controlled trial should have a prespecified primary analysis that is aligned with a clinically meaningful primary estimand. The traditional approach to limiting erroneous conclusions of effectiveness is to select the design and analysis approach such that the type I error probability with respect to the primary estimand is controlled at a prespecified level. It is common to perform hypothesis testing at a one-sided significance level of 0.025 (or two-sided 0.05), where a p-value or confidence limit is used to determine whether the null hypothesis is rejected in favor of an alternative or is not rejected. In certain settings, such as when the prior evidence for the effectiveness of the drug is not strong, FDA may expect a more stringent prespecified significance level (see section IV.A.2). In settings where regulatory flexibility is warranted, FDA may consider flexibility in the significance level expected (see section V.B.1). Bayesian approaches may also be considered. There should be early communication and agreement with

FDA at the design stage on the specification of the prior distribution and the success criteria.

The draft guidance does not establish how prior evidence should be assessed or how sponsors can select alpha prospectively. To remedy this concern ACRO recommends incorporating the following new language into the final guidance:

Where FDA believes a significance level more stringent than the conventional one-sided 0.025 level may be warranted based on the strength of prior evidence, the scientific rationale and proposed success criterion should be discussed and, where feasible, agreed upon with the sponsor prospectively before initiation of the adequate and well-controlled investigation.

Lines 224-228:

These lines currently state:

The prespecified analysis plan should include appropriate statistical methods. For example, the primary analysis should produce valid estimates and inference for the primary estimand under plausible assumptions and should be prespecified with sufficient detail to allow reproduction of results. Any sources of multiplicity (e.g., multiple doses, multiple primary endpoints) should be handled appropriately.

FDA notes that sources of multiplicity, such as multiple doses and multiple primary endpoints, should be handled appropriately; however, the guidance does not specifically address multiplicity considerations for modern master protocol designs, including basket, umbrella, and platform trials, which are increasingly used in oncology and rare disease development.

ACRO would recommend that additional discussion is provided regarding appropriate strategies for controlling false positive conclusions when multiple cohorts, disease subtypes, or biomarker-defined populations are evaluated within a single trial.

Lines 289-303:

In a discussion of how one adequate and well-controlled clinical investigation plus confirmatory evidence can meet the substantial evidence standards, this section currently states:

FDA will consider many factors in assessing whether one adequate and well-controlled trial plus confirmatory evidence is sufficient for demonstrating substantial evidence of effectiveness. These factors are expected to include the design, conduct, analysis, and persuasiveness of results of the clinical trial; the source and strength of the confirmatory evidence; disease-specific considerations (e.g., seriousness, unmet need, prevalence); and whether it is ethical and practicable to conduct more than one adequate and well-controlled trial. The strength of the design, conduct, analysis, and results of the single trial will affect the strength of confirmatory evidence needed to establish substantial evidence of effectiveness. Given the lack of additional adequate and well-controlled trials to substantiate results, FDA generally expects that a demonstration of substantial evidence of effectiveness with this approach will involve a highly persuasive trial or a source of strong confirmatory evidence (see discussion in sections IV.A.1 and IV.A.2 below). There may also be specific circumstances where additional regulatory flexibility is considered in use of this approach (see section V). The approach of demonstrating substantial evidence with one adequate and well-controlled trial plus confirmatory evidence

should be discussed with FDA prior to initiating the single trial.

ACRO's concern here is that binary framing may obscure intermediate combinations. In order to mitigate this concern, ACRO recommends incorporating the following new text into the final guidance:

The strength of the investigation and confirmatory evidence should be considered together as part of the totality of evidence. Multiple complementary sources of confirmatory evidence may collectively provide sufficient confirmation even where no individual source would independently be considered strong.

Lines 308-312:

This section discusses substantial evidence being established via the scenario of evidence from an adequate and well-controlled trial plus a source of strong confirmatory evidence and currently reads:

In some circumstances, substantial evidence of effectiveness may be demonstrated by evidence from a single adequate and well-controlled trial in combination with a source of strong confirmatory evidence. FDA generally expects that such strong confirmatory evidence would come from related adequate and well-controlled trial data, such as trial data in related diseases or conditions or for related products.

ACRO is concerned that this section may unintentionally imply that strong confirmation ordinarily requires another controlled clinical trial. Therefore, we recommend additional clarification by incorporating the following new proposed language into the final guidance:

In appropriate circumstances, strong confirmatory evidence may also arise from multiple complementary sources whose collective relevance, reliability, consistency, biological coherence, and independence of bias provide strong confirmation.

Lines 346-350 and 373-375

Lines 346-350 of the draft guidance currently state that FDA expects that a single highly persuasive trial will typically:

Use a design that provides information that is generalizable and relevant to clinical practice in the United States, such as one that enrolls a broad and representative population across multiple sites, including a sufficient number of U.S. patients if the trial is multiregional, and uses a control arm and supportive therapies that reflect current standard of care.

And, lines 373-375 currently state that FDA expects that a single highly persuasive trial will typically:

Generate results that are largely supportive across important trial subsets (e.g., across subgroups defined by concomitant or prior therapy, disease stage, age, sex, race, or geographic region in a global trial)

This section discusses broad representative population across multiple sites, including sufficient U.S. patients. It also discusses trials in diseases or populations concentrated outside the United States. The agency calls for a broad, representative, multisite trial that is relevant to U.S. practice and includes a sufficient number of U.S.

patients if multiregional. The draft guidance at lines 373-375 also refers to geographic-subset consistency. Those provisions recognize generalizability and consistency of clinical trial results across diverse patient subsets and settings, but they do not a) define “sufficient,” b) address conditions whose disease burden or feasible enrollment lies mainly outside the United States, or c) reconcile the statement with the following FDA regulations and ICH guidelines:

- **21 CFR 312.120**, under which FDA accepts a well-designed and well-conducted foreign clinical study not conducted under an IND if it complies with GCP and FDA can validate the data through inspection if needed;
- **21 CFR 314.106**, under which even an application based solely on foreign clinical data may be approved when the data meet U.S. standards, apply to the U.S. population and medical practice, involve competent investigators, and are validatable; and
- **ICH E17**, which says regional allocation should be scientifically based and consider prevalence, accrual, intrinsic and extrinsic factors, and logistical feasibility. ICH E17 rejects arbitrary regional targets and recognizes that rare disease or infectious outbreaks may require one region to dominate enrollment.

This issue is relevant to conditions such as human immunodeficiency virus, tuberculosis, malaria, emerging infections, and globally dispersed rare diseases.

ACRO recommends the addition of language in bold for lines 346-350:

*Use a design that generates evidence applicable to the intended U.S. population and U.S. medical practice and that permits evaluation of clinically relevant intrinsic and extrinsic factors. A multiregional trial should enroll an appropriate, scientifically justified regional distribution of participants. Enrollment should include sufficient U.S. participants **or other scientifically adequate evidence** – including, as relevant, analyses of disease biology, diagnostic criteria, standard of care, intrinsic and extrinsic factors, PK/PD, exposure-response, endpoint ascertainment, and treatment-effect consistency—to support applicability to the United States. The appropriate regional allocation should account for disease prevalence and feasible accrual; FDA does not expect an arbitrary numerical U.S. quota. Sponsors should discuss the regional allocation, pooling strategy, transportability analysis, and any needed bridging evidence with FDA before initiating the trial.*

Moreover, because the terms “broad,” “representative,” and “sufficient” are undefined and could drive unnecessary and inappropriate U.S. enrollment, ACRO recommends incorporation of the following new text into the final guidance at the ends of lines 346-350:

Trials should enroll a population appropriately representative of the population for the intended-use population. For multiregional trials, U.S. participation should be sufficient to support applicability of the overall results to U.S. clinical practice, considering disease characteristics, standard of care, demographic and clinical factors, and consistency of treatment effect; a fixed percentage of U.S. enrollment is not generally expected.

Lines 352-362:

In the discussion of FDA expectations for what a single highly persuasive trial will typically do, this section of the draft guidance states:

Use a design and analysis plan intended to provide clinically and statistically highly persuasive results, in that the trial (1) uses a clinically meaningful primary endpoint, such as irreversible morbidity or mortality, and (2) has sufficient power to convincingly demonstrate an effect

– The appropriate significance level will depend on the prior (i.e., pretrial) probability of the drug being effective; in cases where the prior probability of effectiveness is low (e.g., because the disease pathophysiology and drug mechanism are not well understood), a single trial with hypothesis testing at the common one-sided 0.025 significance level may not adequately limit the probability of false positive effectiveness conclusions.

ACRO's concern is that "prior probability" is not operationally defined. Moreover, this section above should be read alongside three earlier sections which attach prior probability to potentially more stringent significance levels.

- Lines 266-268 (*"Such information is relevant because the probability of erroneous conclusions of effectiveness depends on the prior probability of effectiveness, the power, and the type I error probability"*)
- Lines 257-262 (*"The persuasiveness and interpretation of an individual trial's results should also be considered in the context of the overall drug development program. In many programs, various objectives such as finding an appropriate dose, studying different patient populations (e.g., with greater and lesser disease severity, pediatric and adult populations, pregnant women), or comparing the drug to other therapy will result in more than one trial providing data relevant to the evaluation of effectiveness"*)
- Lines 217 – 221 (*"In certain settings, such as when the prior evidence for the effectiveness of the drug is not strong, FDA may expect a more stringent prespecified significance level (see section IV.A.2). In settings where regulatory flexibility is warranted, FDA may consider flexibility in the significance level expected (see section V.B.1). Bayesian approaches may also be considered."*).

This approach (while not explicitly stated) appears to reference using a Bayesian approach, however, the January 2026 draft guidance, *Use of Bayesian Methodology in Clinical Trials of Drug and Biological Products*, does not define nor discuss "prior probability." The described sections here essentially show that prior probability is invoked **without** an adequate explanation of what this means, nor are sponsors told what evidence may establish "prior probability," including how not to include evidence that could be double counted or has overlap. Clarity on what prior probability is defined as and what the framework for weighing false-positive versus false-negative consequences should be is needed (i.e., safeguarding against the generation of evidence that has both the potential to inflate prior probability and then also serve as confirmation evidence).

To ameliorate ACRO's concerns, we ask the Agency to consider incorporating the following new language into the final guidance:

Assessment of prior evidence may consider, as applicable, understanding of disease biology, mechanism of action, pharmacodynamic evidence, dose-response and exposure-response information, prior clinical findings, evidence from related indications or products, and other

scientifically relevant evidence. Where a more stringent significance criterion is expected, it should generally be specified prospectively.

364-367:

This section – within the discussion of FDA expectations what a typical single highly persuasive trial should do – currently reads:

Generate primary analysis results that are clinically and statistically highly persuasive based on the clinical meaningfulness of the estimated magnitude of benefit, the associated uncertainty (e.g., confidence interval), and the p-value (or alternative measures such as the posterior probability of effectiveness with a Bayesian approach)

ACRO is concerned that this text could be interpreted as requiring a retrospective p-value threshold. To resolve this concern, we ask FDA to consider incorporating the following new text into the final guidance:

No single numerical p-value or confidence-interval criterion alone determines whether results are highly persuasive. Persuasiveness reflects the totality of the clinical and statistical evidence in the context of the prespecified design and success criterion.

Lines 373-375:

This section – within the discussion of FDA expectations on what a typical single highly persuasive trial should do – currently reads:

Generate results that are largely supportive across important trial subsets (e.g., across subgroups defined by concomitant or prior therapy, disease stage, age, sex, race, or geographic region in a global trial)

ACRO's concern here is that individual trials generally are not powered for subgroup consistency; random subgroup variation could be overinterpreted. To resolve this issue, we ask the Agency to consider inclusion of the following new language in the final guidance:

Generate results that do not demonstrate unexplained, clinically important heterogeneity across relevant prespecified subgroups. Absence of nominal statistical significance within individual subgroups should not, by itself, be interpreted as inconsistency where the trial is not designed or powered to demonstrate treatment effects within those subgroups.

In addition, we would recommend making geographic regions explicit here. For global multi-regional clinical trials (MRCTs), “largely supportive” could otherwise be interpreted as requiring each region, including the US, to independently show a positive treatment effect. It would be helpful to clarify that regional consistency should be assessed in the context of the overall trial and relevant intrinsic and extrinsic factors, and that nominal significance within each region is not expected unless the trial was designed for that purpose.

Lines 377-378:

This section – within the discussion of FDA expectations what a typical single highly persuasive trial should do – currently reads:

Have high-quality conduct, including comprehensive follow-up and minimal missing

data, and demonstrate robustness of results to plausible violations in assumptions

ACRO is concerned that "minimal" and "comprehensive" lack context. To resolve this concern, we recommend the following new text for the final guidance:

Have high-quality conduct, with follow-up and primary endpoint ascertainment sufficient to minimize risk of bias; missing data should be evaluated according to its amount, reasons, distribution between treatment groups, and potential impact on the estimand, with conclusions robust under scientifically plausible assumptions.

Lines 385-390:

The draft guidance currently explains:

If the single trial is adequately designed to provide highly persuasive results, it is expected that the available early-phase information used to support proceeding to such a trial may be able to provide sufficient confirmatory evidence. The final assessment of whether substantial evidence of effectiveness has been demonstrated will depend on the persuasiveness of the single trial results, including the characteristics discussed above, and the sufficiency of the confirmatory evidence.

The current language does not describe qualities of early-phase evidence that enhance its confirmatory value. To provide necessary clarity, ACRO recommends incorporating the following new language into the final guidance:

The confirmatory value of early-phase evidence may be strengthened by reproducibility across cohorts, evidence of dose-response or exposure-response, pharmacodynamic activity consistent with the proposed mechanism, effects on relevant clinical or biological measures, temporal coherence, and consistency with the pivotal trial.

Lines 486-496:

A key section of the draft guidance, on pages 13-14, discusses "Regulatory Flexibility" and "Clinical Circumstances Where Flexibility May Be Warranted." Lines 486-496 of the draft guidance currently state:

These clinical considerations should not be considered individually, but rather holistically, as more than one consideration will often be relevant. For example, many rare diseases are life-threatening and lack available therapy. In general, all relevant clinical considerations (e.g., disease severity, unmet need, and disease rarity) should inform the extent and types of flexibilities that may be appropriate. For example, certain flexibilities might be more acceptable for a rare disease that is life-threatening and lacks available therapies than another rare disease with less severe consequences with multiple therapeutic options already available. Both of these scenarios, however, present a very different clinical context than a common, nonserious disease with several available treatment options. However, in all scenarios, whether the disease is rare or common, fatal or less severe clinically, substantial evidence of effectiveness must be demonstrated.

We ask FDA to appropriately position the emphasis on rare/rapidly consequential disease. The draft guidance expressly applies to NDAs, BLAs, and supplements generally (lines 17-35), and lines 494-496 say the standard applies whether a disease is rare or common, fatal or less severe. Yet the most concrete flexibility examples

repeatedly involve mortality, irreversible morbidity, rare pediatric disease, single-gene defects, natural-history survival, and settings where repetition would be unethical. We suggest that FDA include in the draft guidance language that states rarity and rapid progression are not necessarily pre-requisites, though prior examples are more common among these diseases. We recommend including language that shows how the framework can apply to common serious diseases, as well as chronic disabling conditions, public-health threats, prevention, and regionally concentrated or globally common conditions.

Lines 542-549:

This section of the draft guidance currently states:

For example, mechanistic evidence of the drug's treatment effect in a particular disease may be appropriate to use as confirmatory evidence. Mechanistic evidence might be generated with a clinical trial of a pharmacodynamic endpoint or with nonclinical studies. The pathophysiology of the disease should be well-understood, and the drug's mechanism of action should be both clearly understood and shown to directly target the major driver(s) of the disease pathophysiology. An example might be when the disease is caused by a single gene and/or enzyme defect, and the drug's mechanism of action corrects the enzymatic or genetic defect or its sequelae.

The draft guidance is devoid of consideration of the current “new approach methodologies” (“NAM”) policy. This section addresses the mechanistic approach of generating evidence through pharmacodynamic endpoints or nonclinical studies. While the draft guidance does discuss in footnote 3 traditional animal-model evidence, it is discussed in the exceptional Animal Rule setting. FDA’s published draft guidance, “*General Considerations for Use of the New Approach Methodologies in Drug Development*” defines NAM’s broadly and states that fit-for-purpose (Section III.D, Fit-for-Purpose) nonclinical NAM data may be reviewed in a weight-of-evidence approach (lines 37-38). Exclusion or less than specific inclusion of the current NAM policy may lead sponsors to undervalue or remove fit-for-purpose NAM data when they are measuring/assigning confirmatory evidence or designing specific clinical monitoring. ACRO recommends that the FDA expressly include scientifically justified NAM evidence and cross-reference the current March 2026 draft guidance/NAM policy initiative.

Lines 551-562:

The draft guidance currently states:

In certain circumstances, natural history data may also provide confirmatory evidence. For example, a single randomized controlled trial showing marked improvement in survival could be supported by data from separate sources (e.g., a natural history study or registry) that demonstrate a very limited median survival time without treatment. The natural history data would provide reassurance that the outcomes observed in the control group of the trial accurately reflect those that would have been expected in the absence of the intervention. Natural history data being used as confirmatory evidence should be distinct from any data used as a control for the single adequate and well-controlled trial. Leveraging natural history data may involve use of real-world data (RWD) sources. Whether an RWD source may be appropriate to provide confirmatory evidence depends on several factors, such as the reliability and relevance of the RWD source for its proposed use and, when relevant, the appropriateness of the study design and the use of appropriate prespecified statistical methods and analyses.

ACRO's concern here is that "distinct" is not defined. It is unclear whether separate, non-overlapping cohorts from the same registry or RWD source could serve different purposes, or whether FDA expects entirely separate data source. To provide essential clarity, ACRO asks the Agency to consider incorporating the following new text into the final guidance:

Natural history data used as confirmatory evidence should be sufficiently independent from data used as a control for the single adequate and well-controlled trial to avoid overlap or double counting. Separate, non-overlapping cohorts from the same registry or RWD source may be acceptable when independence can be demonstrated, and the source is fit for each intended use.

Lines 575-587:

The final section of the draft guidance on "Safety Considerations" currently reads:

An FDA approval decision, among other things, requires a determination that a drug is safe for its intended use, which involves assessing whether the benefits of the drug outweigh its risks. Detailed considerations for the safety evaluation and benefit-risk assessment are beyond the scope of this guidance. However, it is important to note that sponsor development programs should be planned to both demonstrate effectiveness and support an appropriate safety and benefit-risk assessment. In some cases, adequate and well-controlled trial(s) that are sufficient to demonstrate effectiveness may not include a sufficient number of participants or a sufficient treatment duration to conclude that the drug is safe for its intended use. A larger or longer trial or an additional trial may be needed to ensure a safety database of adequate size and duration to support an appropriate benefit-risk assessment and drug approval. Sponsors should consult relevant guidances that discuss general expectations for the extent of drug exposure to assess safety, as well as additional factors that may impact the appropriate size of a safety database and other aspects of development program planning for the evaluation of drug safety.

As currently presented in the draft guidance, it is unclear whether the need for additional clinical data solely to support safety would affect FDA's acceptance of a single-trial strategy for demonstrating effectiveness. To provide greater clarity in the final guidance, ACRO asks FDA to consider including the following new text in the final guidance:

When additional clinical data are needed solely to support the safety database, those data would not necessarily need to constitute a second adequate and well-controlled efficacy trial, provided the overall development program supports an appropriate safety and benefit-risk assessment.

Conclusion

ACRO thanks FDA for the opportunity to provide this feedback. Please do not hesitate to contact ACRO (knoonan@acrohealth.org) if we can answer any questions or provide additional information.

Respectfully submitted,
Karen Noonan
Senior Vice President, Global Regulatory Policy

ADDENDUM: Proposed Glossary

Section Three: Proposed Glossary

Phrase	Definitions, considerations, and rationales
“highly persuasive”	In the draft guidance, there are currently no clear criteria for “highly persuasive” trial. The Agency identifies characteristics that may contribute to a highly persuasive trial, including clinically meaningful endpoint, sufficient power, compelling effect magnitude, supportive secondary endpoints, subgroup consistency, minimal missing data, and robustness. The draft guidance also states that the conventional one-sided 0.025 significance level may not be sufficient when the pretrial probability of effectiveness is low. However, sponsors still cannot readily determine prospectively whether their planned study has a reasonable likelihood of qualifying as “highly persuasive.” ACRO recommends that FDA provide a more structured, nonbinding framework, preferably with illustrative examples, describing how factors such as the magnitude and precision of the treatment effect, statistical evidence relative to the prespecified success criterion, clinical meaningfulness of the endpoint, supportive secondary endpoints, subgroup findings, missing data, robustness of results, trial conduct, and prior biological and clinical evidence are considered collectively in determining whether a trial is “highly persuasive.” Importantly, the Agency should clarify that “highly persuasive” does not require achievement of a specific numerical p-value threshold beyond the prespecified statistical success criterion. Otherwise, sponsors may interpret the guidance as creating an unstated heightened statistical standard for single-trial programs. ACRO recommends that FDA clarify that determination of whether a trial is “highly persuasive” is based on the totality of the trial

	evidence and the overall development context, for example, magnitude of treatment effect, reproducibility, confirmatory evidence, disease context, etc., and is not dependent on any single factor including, achieving a particular significance threshold beyond the prespecified statistical success criterion.
“sufficient number of U.S. patients”	U.S. participation adequate to support the applicability, interpretability, and regulatory relevance of the overall trial results to U.S. patients and U.S. clinical practice, based on scientific and clinical considerations rather than a predetermined numerical threshold or percentage. Relevant considerations may include disease characteristics and epidemiology, demographic and clinical representation, intrinsic and extrinsic factors that may affect treatment response, background therapy and standard of care, diagnostic and endpoint-assessment practices, pharmacologic factors, and the consistency of treatment effects across regions. The determination of sufficiency should be consistent with ICH E17 principles for multiregional clinical trials and should reflect the extent to which U.S. participation, together with the overall global data, permits assessment of whether the trial results are applicable to the U.S. population. A sufficient number of U.S. patients should not ordinarily require the trial to be powered to independently establish effectiveness within the U.S. subgroup or to meet a fixed minimum U.S. enrollment percentage.
“Largely supportive”	Results across relevant prespecified subgroups or geographic regions that do not show unexplained, clinically meaningful heterogeneity in treatment effect. “Largely supportive” should not require nominal statistical significance within each subgroup or region when the trial was not designed or powered to independently demonstrate treatment effect within those groups.
“a broad and representative population across multiple sites”	A trial population and site distribution that are appropriately representative of the population of A trial population and site distribution that are appropriately representative of the population for whom the proposed indication is intended and

	sufficiently diverse to support interpretation, applicability, and generalizability of the trial results. The appropriate breadth of the population, number of sites, and geographic distribution should be determined based on the disease or condition, proposed indication, prevalence, relevant demographic and clinical characteristics, intrinsic and extrinsic factors that may affect treatment response, standard of care, trial design, endpoint characteristics, and operational feasibility. A population may be considered broad and representative even when enrollment is intentionally enriched or restricted based on scientifically justified factors, such as biomarker status, disease stage, prior therapy, genotype, or other characteristics relevant to the proposed indication. For multiregional trials, representativeness should be considered in a manner consistent with ICH E17 and applicable FDA guidance addressing enrollment of populations representative of the intended-use population.
“strong confirmatory evidence”	Evidence independent of, or complementary to, the principal adequate and well-controlled investigation that materially increases confidence that the observed effectiveness finding reflects a true treatment effect. The strength of confirmatory evidence should be assessed based on factors such as relevance, reliability, consistency, biological coherence, complementarity, and independence of potential biases. Strong confirmatory evidence may consist of one highly informative source or multiple complementary sources whose collective evidentiary value provides strong confirmation.
“minimal missing data”	An amount and pattern of missing data that do not materially compromise interpretation of the treatment effect or introduce substantial risk of bias. Assessment should consider the amount of missing data, reasons for missingness, distribution across treatment groups, relationship to intercurrent events and the estimand, and robustness of conclusions under scientifically plausible sensitivity analyses, rather than relying on a fixed numerical threshold.
“sufficient confirmatory evidence”	

	Confirmatory evidence that, when considered together with the principal adequate and well-controlled investigation and the overall development program, provides sufficient independent or complementary support to materially reduce the likelihood that the observed effectiveness finding is attributable to chance, bias, confounding, or an anomalous study result. The amount and strength of confirmatory evidence necessary may vary according to the persuasiveness of the principal investigation and the totality of the available evidence. ACRO recommends inclusion of this definition because “strong confirmatory evidence” and “sufficient confirmatory evidence” are not necessarily the same things. The guidance contemplates that a highly persuasive trial may require less confirmatory evidence than an otherwise adequate trial.
--	---