

FOOD AND DRUG LAW JOURNAL

*Analyzing the Laws, Regulations, and Policies
Affecting FDA-Regulated Products*

The Chicken and the Egg—Animal Welfare, Food Safety and
Federalism

Rita-Marie Cain

Genetically Modified Food Labeling in China: In Pursuit of a
Rational Path

Xiao Zhu, Michael T. Roberts, and Kajie Wu

An Artless Tale: Challenges Faced in Clinical Research

Stephen Raper

FDA's Troubling Failures to Use Its Authority to Regulate
Genetically Modified Foods

Leslie Francis, Robin Kundis Craig, and Erika George

FDA's Flexibility in Subpart H Approvals: Assessing Quantum of
Effectiveness Evidence

Frank J. Sasinowski and Alexander J. Varond



GEORGETOWN LAW



VOLUME 71 NUMBER 1 2016

FDA's Flexibility in Subpart H Approvals: Assessing Quantum of Effectiveness Evidence

FRANK J. SASINOWSKI^{*}
ALEXANDER J. VAROND^{**}

Cataloging FDA's Flexibility in Subpart H Approval¹

Abstract

This article examines the strength of scientific and clinical evidence for FDA's nineteen non-AIDS, non-cancer Subpart H approval determinations over the Accelerated Approval program's twenty-four year existence. The authors researched the bases for FDA's determinations when an unvalidated surrogate or intermediate clinical endpoint is "reasonably likely to predict clinical benefit." The four key factors set forth in FDA's "Guidance for Industry, Expedited Programs for Serious Conditions – Drugs and Biologics" were applied to past Subpart H approvals. For the nineteen precedents, the authors found wide variances between the quantum and quality of evidence on each of the four factors, indicating that a lack of evidence on any single factor was not disqualifying in and of itself. The results of this study, therefore, show that FDA exercises extraordinarily more regulatory flexibility than either FDA's foundational statutes or even FDA's most recent 2014 Expedited Programs Guidance explicitly express. Given recent legislative exhortations and the increasing promise of personalized medicine and translational sciences, the authors conclude that Subpart H should be further explored and utilized. The authors provide a detailed analysis of the precedents established in the nineteen approvals.

I. INTRODUCTION

Born out of FDA's own ingenuity to address the AIDS crisis in the mid-1980s, Subpart H authority has existed for well over twenty years. This FDA-created program has aided in the development of many critical therapies for persons with AIDS and cancer. However, despite this success in these two areas, very few therapies have been approved via the Subpart H pathway for serious and life-threatening diseases other than AIDS and cancer. By cataloging FDA's Subpart H approval of the non-AIDS and

^{*} Frank J. Sasinowski, M.S., M.P.H., J.D., Director, Hyman, Phelps & McNamara, P.C.; Adjunct Professor of Neurology, University of Rochester School of Medicine; Board of Directors, National Organization for Rare Disorders (NORD).

^{**} Alexander J. Varond, J.D., Associate, Hyman, Phelps & McNamara, P.C.

¹ The initial version of the analysis contained in this article was submitted in August 2013 to FDA in response to a request for comments to FDA's draft guidance entitled, "Expedited Programs for Serious Conditions—Drugs and Biologics" (June 2013) ("Draft Guidance"). While much of that initial analysis remains, this article revises that analysis and extends the discussion in several key areas. This additional, revised analysis is the result of feedback we received from key opinion leaders and senior FDA officials at the Center for Drug Evaluation and Research, including Robert Temple, M.D. (who many regard as one of the key architects of Subpart H).

non-cancer therapies, this article illuminates the basis for FDA's determinations of when an unvalidated surrogate is "reasonably likely to predict clinical benefit," as well as on FDA's exercise of "flexibility" in Subpart H approvals.

Recently, Subpart H² has taken on significant added importance as an innovative regulatory vehicle for providing therapies to patients suffering from serious and often rare diseases where there is inadequate available therapy. Four recent milestone events illustrate this: the passage of FDA Safety and Innovation Act ("FDASIA") in July 2012; the publication of the President's Council of Advisors on Science and Technology ("PCAST") Report to the President on Propelling Innovation in Drug Discovery, Development, and Evaluation ("PCAST Report") in September 2012; FDA's release of "Guidance for Industry: Expedited Programs for Serious Conditions—Drugs and Biologics" ("Expedited Programs Guidance") in May 2014; and the ongoing debates in the Senate over the U.S. House of Representatives' passing of the 21st Century Cures Act.

A. FDASIA

In FDASIA, Congress and President Obama revised the statutory provisions of Subpart H to "facilitate somewhat broader use of [Subpart H] to expedite patients' access to important treatments for serious conditions[,] . . . provide additional flexibility[,] . . . provide clarification concerning the use of clinical endpoints[,] . . . [and] make clear that FDA has the authority to consider pharmacologic or other evidence . . . in determining whether an endpoint is reasonably likely to predict clinical benefit."³ While these were added in July 2012 by statute, Congress was essentially codifying the practices and policies that FDA had already put into place and acted upon previously. In Section 901 of that statute, FDASIA did, however, direct FDA to *expand FDA's use* of this Subpart H authority:

The FDA should be encouraged to implement more broadly effective processes for the expedited development and review of innovative new medicines intended to address unmet medical needs for serious or life-threatening diseases or conditions, including those for rare diseases or conditions, using a broad range of [unvalidated] surrogate or clinical endpoints. . . . This may result in fewer, smaller, or shorter clinical trials for the intended patient population or targeted subpopulation without compromising or altering the high standards of the FDA for the approval of drugs. Patients benefit from expedited access to safe and effective innovative therapies to treat unmet medical needs for serious or life-threatening diseases or conditions. For these reasons, the statutory authority in effect on the day before the date of enactment of this Act governing expedited approval of drugs for serious or life-threatening diseases or conditions should be amended in order to enhance the authority of the FDA to consider appropriate scientific data, methods, and tools, and to expedite development and access to novel treatments for

² For simplicity, this article exclusively employs the term "Subpart H" even though the authors recognize that occasionally another term, either "Accelerated Approval" or "Fast Track" would be correct.

³ FOOD & DRUG ADMIN., GUIDANCE FOR INDUSTRY: EXPEDITED PROGRAMS FOR SERIOUS CONDITIONS—DRUGS AND BIOLOGICS 15 (May 2014) [hereinafter EXPEDITED PROGRAMS GUIDANCE].

patients with a broad range of serious or life-threatening diseases or conditions.

SENSE OF CONGRESS.—It is the sense of Congress that the Food and Drug Administration should apply the accelerated approval and fast track provisions set forth in section 506 of the Federal Food, Drug, and Cosmetic Act (21 U.S.C. 356), as amended by this section, to help expedite the development and availability to patients of treatments for serious or life-threatening diseases or conditions while maintaining safety and effectiveness standards for such treatments.⁴

B. Propelling Innovation in Drug Discovery, Development, and Evaluation

In September 2012, President Obama became the first president to comprehensively address the complexities of developing new medicines for Americans when he released his report, “Propelling Innovation in Drug Discovery, Development, and Evaluation.”⁵ The PCAST Report instructed FDA to expand the use of its Subpart H authority.⁶ Specifically, the PCAST Report recommended that:

The FDA should expand the scope of acceptable endpoints used to approve drugs for serious or life-threatening diseases with unmet needs. Under current law, the FDA has considerable discretion in deciding whether [an unvalidated] surrogate or intermediate clinical endpoint is “reasonably likely, based on epidemiologic, therapeutic, pathophysiologic, or other evidence, to predict clinical benefit.” At one extreme, the FDA might be highly risk-averse, requiring near-certainty that the [unvalidated] surrogate or intermediate endpoint will translate to clinical benefit. At the other extreme, the Agency might accept endpoints that are simply correlated with disease outcome or plausibly related to disease outcome based on current scientific understanding. Neither extreme would serve the public well. The FDA’s interpretation of “reasonably likely . . . to predict” can have a major impact on the pace of medical innovation and on patient safety. . . . Historically, the use of [Subpart H] has been primarily used in a limited number of therapeutic areas—principally, HIV/AIDS, cancer, and inhalation anthrax (87 percent of cases). . . . We believe that the Nation would benefit if the FDA were to expand the use in practice of acceptable indicators to other serious or life-threatening diseases.⁷

The FDA should make fuller use of authorities previously granted by legislation and not yet fully utilized. The FDA should expand the use in practice of its existing authority for [Subpart H]. The FDA should direct

⁴ Food and Drug Administration Safety and Innovation Act, Pub. L. No. 112-144, § 901, 126 Stat. 993, 1082–83.

⁵ One of the authors of this article was involved in the process and was a key contributor to this report.

⁶ See PRESIDENT’S COUNCIL OF ADVISORS ON SCI. & TECH., REPORT TO THE PRESIDENT ON PROPELLING INNOVATION IN DRUG DISCOVERY, DEVELOPMENT, AND EVALUATION 59–68 (Sept. 2012) [hereinafter PCAST REPORT].

⁷ *Id.* at 59.

its staff, across all divisions, to make full use of the [Subpart H] track for all drugs meeting the statutory standard of addressing an unmet medical need for a serious or life threatening illness and demonstrating an effect on a clinical endpoint . . . or on [an unvalidated] surrogate⁸ endpoint that is reasonably likely to predict clinical benefit.⁹

C. *Guidance for Industry: Expedited Programs for Serious Conditions—Drugs and Biologics*

In May 2014, FDA released its Expedited Programs Guidance. The Expedited Programs Guidance lists and describes factors that FDA views as critical to Subpart H approvals.¹⁰

D. *21st Century Cures Initiative*

On May 1, 2014, the House Energy and Commerce Committee's Subcommittee on Health launched the 21st Century Cures Initiative with the goal of "accelerat[ing] the pace of discovery, development, and delivery cycle so we can get innovative new cures and treatments to patients more quickly."¹¹ The effort represents the first time that Congress has set itself towards taking a comprehensive look at what steps it can take to accelerate the pace of cures in America. The Subpart H pathway can play an important role in fulfilling this focused, bipartisan effort by lawmakers. During testimony at the first hearing of the 21st Century Cures Initiative, held on May 20, 2014, one of the authors of this article called on Congress to increase Subpart H's visibility and encourage Sponsors and FDA to use intermediate clinical endpoints ("ICE"). Ongoing drafts of the 21st Century Cures Act, including drafts released in January 2015 and April 2015, and the version approved in the House of Representatives on July 10, 2015 by a vote of 344 to 77 promote the expanded development and use of unvalidated surrogate endpoints and the Subpart H pathway.¹²

E. *Propelling Subpart H Forward*

All four of these activities highlight the renewed recognition of the promise of FDA's Subpart H authority to address the needs of those suffering from serious

⁸ The 1992 FDA Subpart H regulations and the Congressional codification of those regulations in the Fast Track provisions of the 1992 FDA Modernization Act ("FDAMA") both authorize FDA to rely upon either a surrogate endpoint or a clinical endpoint so long as either of these two is "reasonably likely to predict" a more meaningful clinical (i.e., patient) benefit (i.e., the type of clinical benefit that would qualify for a traditional approval). In its Expedited Programs Guidance, FDA calls this kind of clinical evidence an "intermediate clinical" endpoint. See EXPEDITED PROGRAMS GUIDANCE, *supra* note 3, at 15. In this article, the authors simplify this by using the term "surrogate" to mean both types of prognostic evidence. The original draft proposed that Subpart H regulation only include "surrogate" but the Office of Management and Budget ("OMB") asked FDA to add an additional alternate path forward and FDA then inserted what FDA now calls "the intermediate clinical endpoint" into FDA's notice of proposed rulemaking ("NPRM") on Subpart H on April 15, 1992.

⁹ PCAST REPORT, *supra* note 6, at 61.

¹⁰ See EXPEDITED PROGRAMS GUIDANCE, *supra* note 3, at 16–21.

¹¹ *21st Century Cures: Hearing Before Subcomm. on Health, H. Energy & Commerce Comm.*, 113th Cong. (2014) (statement of Fred Upton, Chairman, H. Energy & Commerce Comm.), <http://energycommerce.house.gov/sites/republicans.energycommerce.house.gov/files/Hearings/Health/20140520/HHRG-113-IF14-MState-U000031-20140520.pdf>.

¹² See, e.g., H. ENERGY & COMMERCE COMM., 114TH CONG., DISCUSSION DRAFT OF 21ST CENTURY CURES ACT §§ 2021, 2022 (Apr. 2015).

diseases without adequate available therapy. Moreover, FDA's issuance of its Expedited Programs Guidance addresses the Agency's Subpart H authority. Consequently, the following analysis of FDA precedents is timely in that it should promote a better understanding of the circumstances under which Subpart H may be employed. In this way, it is hoped that the regulatory ingenuity of FDA in creating Subpart H and the recent Congressional and Executive exhortations to better mobilize this Subpart H power may find fuller expression.¹³

II. PURPOSE OF THIS STUDY

This study was designed to examine closely how much flexibility FDA provides in reviewing Subpart H applications — that is, what particular factors the Agency considers in approving a drug under Subpart H and how much weight it assigns to each of these factors.

This study examines the nineteen Subpart H¹⁴ approvals identified by FDA in its public database that are for conditions other than AIDS and cancer. AIDS and cancer therapies were excluded from the analysis because there is comparatively greater regulatory certainty associated with Subpart H approvals and unvalidated surrogate endpoints for these two therapeutic areas.¹⁵ The intent here is to determine, based on an examination of the publicly available information used to support approval, the amount (or quantum) and quality of evidence for each factor that was necessary for approval of non-AIDS and non-cancer therapies using the Subpart H approval pathway.

This study aims to determine the evidentiary foundation for FDA's findings that an unvalidated surrogate or clinical endpoint was "reasonably likely to predict" patient benefit sufficient to meet the statutory standard of "substantial evidence of effectiveness." Recall that the Federal Food, Drug, and Cosmetic ("FD&C") Act provides that for FDA to grant approval for a new drug, there must be "substantial evidence of effectiveness" derived from "adequate and well-controlled investigations."¹⁶ This language, which dates from 1962, provides leeway for FDA medical reviewers to make judgments as to what constitutes "substantial evidence" of a drug's effectiveness, that is, of a drug's therapeutic benefit to patients.

¹³ One of the two authors here has participated in two analyses of FDA orphan drug precedents that has, to some, proved of some utility. Frank Sasinowski, *Quantum of Effectiveness Evidence in FDA's Approval of Orphan Drugs*, 46(2) DRUG INFO. J. 238–263 (Mar. 2012); Frank Sasinowski et al., *Quantum of Effectiveness Evidence in FDA's Approval of Orphan Drugs: Update, July 2010 to June 2014*, THERAPEUTIC INNOVATION & REG. SCI. (Apr. 27, 2015). The authors hope that this analysis of Subpart H precedents may also prove to be useful.

¹⁴ Subpart H refers to FDA regulation of 21 C.F.R. pt. 314 Subpart H which is the drug regulation; the parallel regulation for biologics is 21 C.F.R. pt. 601 Subpart E. Some of these nineteen approvals which the authors refer to as Subpart H Approvals were approvals of biologics under 21 C.F.R. pt. 601 Subpart E, including the very first "Subpart H" approval that was not for AIDS or cancer: Betaseron for multiple sclerosis in 1993.

¹⁵ In addition, drugs approved via Subpart H on the basis of FDA's so-called "Animal Efficacy Rule," 21 C.F.R. §§ 314.610 (drugs), 601.91 (biologics), under which FDA can rely on evidence from animal studies to provide substantial evidence of effectiveness when it is unethical or unfeasible to conduct human efficacy studies to obtain approval of a countermeasure product (e.g., Levaquin (levofloxacin) for inhalational anthrax, Cipro (ciprofloxacin) for inhalation anthrax, Avelox (moxifloxacin) for septicemic and pneumonic plague), were excluded.

¹⁶ Federal Food, Drug, and Cosmetic Act § 505(d), 21 U.S.C. § 355 (2015).

There is a persistent and important misconception among many, including industry Sponsors, that the Subpart H approval pathway can somehow be employed when evidence of effectiveness is inadequate to support full or traditional or conventional approval. It is critical for all to understand that the statutory requirement of establishing statistical evidence of effectiveness applies with equal force to Subpart H approvals.

What, then, is the difference? The difference is in the endpoint for which there is “substantial evidence.” In a Subpart H approval, the endpoint is an unvalidated surrogate or intermediate clinical endpoint, whereas in a full approval, the endpoint is a measure of how the patient meaningfully “feels . . . , functions . . . , or survives.”¹⁷

Through this analysis, the authors hope to promote a better understanding of the circumstances under which Subpart H may be employed in order to facilitate the development and expedited review of new drugs with the potential to address unmet medical needs for serious and life-threatening illnesses and to mobilize expanded FDA use of Subpart H.

III. METHODS

To identify the non-cancer and non-AIDS drugs approved as new chemical entities via Subpart H, the authors relied upon FDA’s publicly available list of such drugs approved through August 26, 2014.^{18, 19}

For each of the nineteen Subpart H approved drugs (“approval precedents”), the authors sought access to the FDA approval letter, the labeling at the time of that approval (in order to exclude subsequent supplemental information that later added new clinical data), the decision memoranda of FDA officials who approved the product, and the reviews of the medical and statistical officers (collectively “approval documents”). While such documents were retrievable in most cases, only subsets of these approval documents were recoverable for some drugs, especially for several of the earliest Subpart H approvals.

The analysis of each of the nineteen approval precedents is organized according to the order of factors cited in FDA’s Expedited Programs Guidance. Organizing the analyses in this way maximizes the usefulness of the article when comparing it to the Expedited Programs Guidance and also provides a logical structure.

An analysis of each of the approval precedents is presented in the section entitled Case Study Analyses, which describes the most relevant information pertinent to each factor in FDA Expedited Programs Guidance. A numerical scorecard was developed to evaluate each precedent. The scoring system was set up to provide enough granularity to account for differences in the level of support for each factor, while at the same time keeping it simple enough to avoid unnecessary complication. As a result, a 20-point scale was devised, with each factor assigned varying weights

¹⁷ EXPEDITED PROGRAMS GUIDANCE, *supra* note 3, at 17.

¹⁸ See, e.g., U.S. FOOD & DRUG ADMIN., ACCELERATED AND RESTRICTED APPROVALS UNDER SUBPART H (DRUGS) AND SUBPART E (BIOLOGICS), <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/DrugandBiologicApprovalReports/ucm121597.htm> (last updated Aug. 26, 2014).

¹⁹ The authors acknowledge that only FDA-approved products are included in this analysis. This is because FDA does not make publicly available the kinds of information that would be necessary to conduct a similar analysis on those therapies that FDA reviewed but never approved. Indeed, FDA does not even publish a list of therapies that are not approved.

according to the relative importance of the factor. For example, a 0–7 point scale indicated that the factor was very important. Thus, a factor scored on a scale of 0–7 points was 3.5 times more important than a factor scored on a scale of 0–2 points. A theoretically “perfect” Subpart H therapy for a rare disease would receive a score of 20 points.

This 20-point composite scoring system offers a mechanism to roughly mimic the weighing of evidence FDA reviewers must conduct in their reviews of the evidence of treatment benefit. The ultimate approval decision is much more complex in that FDA then must evaluate the strength of the clinical evidence of benefit against the risks, that is, the safety considerations associated with the investigative therapy and do this while also giving due consideration to the severity of the disease and the lack of any other alternative therapy for the patients with the particular condition. FDA’s draft version of the Expedited Programs Guidance from June 2013 (“Draft Guidance”) captures the essence of this succinctly: “FDA considers all relevant evidence and weighs the uncertainty against the severity of the disease to be treated and the lack of available therapy.”²⁰

A. Part 1: Rarity of the Condition

FDASIA amended the FD&C Act to include the “rarity of the condition” as an additional factor to consider when determining whether a therapy qualifies for Subpart H. Specifically, FDASIA provided FDA with the authority to approve a therapy under Subpart H when the Agency determines “that the product has an effect on [an unvalidated] surrogate . . . that is reasonably likely to predict clinical benefit . . . taking into account the . . . *rarity, or prevalence of the condition* and the availability . . . of alternative treatments.”²¹

In addition, under Section VII. C. “Evidentiary Criteria for Accelerated Approval,” FDA acknowledges that “whether an endpoint is reasonably likely to predict clinical benefit is a matter of judgment.”²²

Given the flexibility FDA affords orphan drugs and the new FDASIA statutory requirement for considering the rarity of the condition, a range of 0–2 points, or 10% of the total score, is assigned to the “rarity” assessment of each of the nineteen Subpart H approval precedents.²³

B. Part 2: Understanding of the Disease²⁴

The second element of the analysis of each approval precedent is the degree to which the underlying disease is understood. As FDA explains in the Expedited

²⁰ FOOD & DRUG ADMIN., DRAFT GUIDANCE FOR INDUSTRY: EXPEDITED PROGRAMS FOR SERIOUS CONDITIONS—DRUGS AND BIOLOGICS 18, lines 610–611 (June 2013).

²¹ FD&C Act § 506(b)(1)(A), 21 U.S.C. § 356 (2012) as amended by the Food and Drug Administration Safety and Innovation Act, Pub. L. No. 112–144, § 901, 126 Stat. 993, 1083–84 (emphasis added).

²² EXPEDITED PROGRAMS GUIDANCE, *supra* note 3, at 19.

²³ While it is not a requirement that a therapy be for a rare condition to qualify for Subpart H consideration, FDASIA’s addition of rarity to the factors that must be considered means that, to be accurate, the authors’ scoring system has to give it some weight. It also means that FDA should be more willing to give a rare disease Subpart H consideration than a prevalent disease were everything else the same. While there is merit to this, that discussion is beyond the scope of this article.

²⁴ Parts 2–4 not only track the order in the Expedited Programs Guidance, but also are tied to the language quoted in FDASIA: “The evidence to support that an endpoint is reasonably likely to predict

Programs Guidance, understanding the disease process is fundamental to understanding the “biological plausibility” of the unvalidated surrogate.²⁵

This factor is assigned a value of 0–4 because FDA regards this as “an important factor” indicating the relative importance in understanding the disease process: “the extent to which the pathophysiology of a disease is understood is an important factor in determining whether an endpoint is reasonably likely to predict clinical benefit.”²⁶ If a disease process is well-understood, identifying an adequate unvalidated surrogate endpoint is more likely.

C. Part 3: Understanding of the Relationship Between Drug Effect and Disease Process

Under Section VII. C. 2. “Understanding of the Relationship Between the Drug’s Effect and the Disease Process,” FDA notes that “[t]he extent to which a drug’s effect on the unvalidated surrogate endpoint is known to predict an effect on the disease because the effect is on the causal pathway or correlates with clinical outcome is critical.”^{27, 28}

FDA, in its Expedited Programs Guidance, then lists several factors to consider in identifying and assessing an unvalidated surrogate endpoint, including, “[w]hether there is reliable and consistent epidemiologic evidence supporting the relationship between the endpoint and the intended clinical benefit” and “[w]hether the effect on the [unvalidated] surrogate endpoint has been shown to predict a clinical benefit with another drug or drugs.”²⁹ Therefore, the third part of the analysis assesses the evidence for these factors, noting that, for the purposes of this analysis, epidemiological evidence is interpreted more broadly to include all observational studies, including long-term longitudinal studies and “natural history” studies. This part of the analysis of each approval precedent essentially assesses the predictive or prognostic potential of the unvalidated surrogate.

Given that this is an important factor in FDA’s evaluation of Subpart H candidates, this factor is assigned a value of 0–4.

D. Part 4: Clinical Evidence for the Unvalidated Surrogate and for the Clinical Benefit

In its Expedited Programs Guidance, FDA acknowledges the primacy of clinical evidence for the drug itself, both on the unvalidated surrogate and on the clinical

clinical benefit . . . may include epidemiological, pathophysiological, therapeutic, pharmacologic, or other evidence . . .” Food and Drug Administration Safety and Innovation Act, Pub. L. No.112–144, § 901(a), 126 Stat. 993, 1082–83. Specifically, Part 2 relates to “pathophysiological” evidence; Part 3 relates to “epidemiological . . . pharmacologic, or other evidence;” and Part 4 relates to “therapeutic evidence.”

²⁵ EXPEDITED PROGRAMS GUIDANCE, *supra* note 3, at 19.

²⁶ *Id.* at 20.

²⁷ *Id.* at 21.

²⁸ The Expedited Programs Guidance continues (and in so doing helps to explain the relative value of what in this analysis we have divided into Parts 3 and 4): “Sometimes this relationship can be assessed epidemiologically but it is most persuasively established by knowing that a drug that affects the surrogate also affects a clinical outcome.” EXPEDITED PROGRAMS GUIDANCE, *supra* note 3, at 21. In this analysis, we, accordingly, weight more heavily Part 4 (Clinical Evidence) as compared to Part 3 (which is, in part, epidemiology).

²⁹ EXPEDITED PROGRAMS GUIDANCE, *supra* note 3, at 21.

benefit. FDA further explains that the guidance “does not, however, address the specific clinical evidence needed to support a conclusion that a particular [unvalidated] surrogate endpoint or intermediate clinical endpoint is reasonably likely to predict clinical benefit or [irreversible morbidity or mortality (“IMM”)] because such evidence is case-specific and is not readily generalizable.”³⁰ In a similar way, this analysis does not need, nor does it try, to distill or deduce generalized requirements from the many precedents. The authors in this analysis, instead, assess the strength of the clinical evidence in *each* case, for *each* drug’s effect, both on the specific unvalidated surrogate and on the particular clinical benefit. In this way, the authors describe the conditions of each approval around a scoring system so that the specifics of each case stand, both on their own and together, across the totality of the Subpart H approval landscape. Accordingly, the fourth and final part of each analysis is the strength of clinical evidence on the unvalidated surrogate itself, as well as on the clinical benefit. Here, the clinical evidence is broken into two separate categories: clinical evidence for the unvalidated surrogate and/or ICE endpoints (0–7 point scale) and clinical evidence for the clinical benefit (0–3 point scale).

Clinical evidence for the unvalidated surrogate is assigned a weight of 7-points because it is *by far* the single most important factor in Subpart H approvals. This is because Subpart H directs FDA to apply the same legal/regulatory standard of “substantial evidence of effectiveness” on the unvalidated surrogate endpoints or ICE that it does for traditional approvals.

Any clinical evidence for the clinical benefit itself is given 3-point weighting because, for Subpart H approvals, evidence of clinical benefit is not required. The authors have, however, given it modest weight because, although there is no requirement that a Subpart H approval demonstrate any evidence of clinical benefit, the demonstration of any such benefit (beyond the intermediate clinical endpoint, of course) would anchor and strengthen the prognostic value of the evidence on the unvalidated surrogate endpoints and/or ICE. Given that such evidence of clinical benefit is not required, it is not at all surprising that ten of the nineteen approval precedents include no evidence of clinical benefit. Where it is present, preliminary evidence of clinical benefit in Subpart H approvals strengthens the “signal” that the findings on the unvalidated surrogate endpoints and/or ICE are predictive of the clinical benefit. A full score of 3 on this would not mean that the approval documents contained evidence that would support a traditional approval, but it would indicate how strongly the clinical benefit evidence reinforced the evidence for the unvalidated surrogate and/or ICE endpoints.

E. Summary of Scoring System

The authors’ scoring system is presented in Table 1. As a caveat, these weights are a matter of judgment, as are each of the assessments or “scores.” Other individuals may prefer either greater or lesser weights for any of these factors, and may even decide that some of these factors should not be included at all or still others be added.

³⁰ *Id.* at 20.

Table 1: Subpart H Scoring Rubric

Factor	Element within Factor	Range
Part 1: Rarity of the Disease	Rarity of the disease	0–2
Part 2: Understanding of the Disease Process	Understanding of the pathophysiology of the disease	0–4
Part 3: Understanding of Relationship Between the Drug’s Effect on the Unvalidated Surrogate and the Disease	Understanding from epidemiological evidence, animal models, other drugs in similar pharmacologic class or other sources	0–4
Part 4: Strength of Clinical Evidence	Clinical evidence on surrogate or ICE	0–7
	Clinical evidence of benefit	0–3

The available approval documents for each of the nineteen non-cancer, non-AIDS approval precedents were evaluated and each approval precedent was scored according to the system presented. Results are summarized in Figure 1 and Table 2.³¹

³¹ Individuals with direct knowledge of any disease precedents, especially Sponsors and FDA reviewers and supervisory officials as well as experts in the medical or patient communities, may disagree with the scores given to any factor for any of these nineteen Subpart H precedent approvals. These alternative views would be understandable, and even welcome, especially when based on a more thorough understanding of the disease, of the science, or of confidential evidence available only to the Sponsors and FDA.

IV. RESULTS

Figure 1 summarizes the “scores” for each of the nineteen Subpart H approval precedents according to the factors laid out in the Expedited Programs Guidance. Table 2 summarizes the cumulative totals for each precedent.

Figure 1: Cumulative Scores for Non-AIDS/Non-Cancer Subpart H Precedents (0–20)[†]

Drug	Part 1: ² Rarity of the Disease	Part 2: ² Understanding of the Disease Process (Section VII.C.1) ¹	Part 3: ² Understanding of the Relationship Between the Drug's Effect on Surrogate and the Disease (Section VII.C.2) ¹	Part 4: ² Strength of Clinical Evidence (Section VII.C.) ¹		Total
	Rarity (0-2)	Understanding of the Disease (0-4)	Epidemiological, Pharmacologic and "Other Evidence" on the Surrogate or Disease (0-4)	Surrogate Endpoint (0-7)	Clinical Benefit (0-3)	
1. Northera	2	1	1	4	1	9
2. Sirturo	2	3	3	5	-1	12
3. Ferriprox	2	4	2	6	0	14
4. Makena	2	1	3	7	1	14
5. Soliris	2	2	3	3	1	11
6. Promacta	2	3	1	7	1	14
7. Exjade	2	4	3	3	0	12
8. Tysabri	2	2	3	7	1	15
9. Luveris	2	1	3	3	2	11
10. Fabrazyme	2	4	2	7	0	15
11. Remodulin	2	4	1	5	1	13
12. Celebrex	2	4	4	6	0	16
13. Synercid	1	4	2	1	0	8
14. Remicade	2	1	1	6	0	10
15. Priftin	2	3	3	4	0	12
16. Sulfamylon	2	4	2	1	0	9
17. ProAmatine	2	4	2	5	0	13
18. Biaxin	1	3	2	3	1	10
19. Betaseron	2	2	3	7	2	16
	Mean: 1.9±0.3 Range: 1-2 (out of 2)	Mean: 2.8±1.2 Range: 1-4 (out of 4)	Mean: 2.3±0.9 Range: 1-4 (out of 4)	Mean: 5.3±2.2 Range: 1-9 (out of 10)		

¹These citations are to the FDA Guidance

² Parts 2, 3, and 4 are also from FDC Act Sec. 506, as amended by FDASIA Sec. 901, specifically, FDASIA uses the following terms for each of these parts

- Part 2 relates to "pathophysiological" evidence
- Part 3 relates to "epidemiological, . . . Pharmacologic, or other evidence . . ."
- Part 4 relates to "therapeutic" evidence

Statistics	Score
Average Score	12
Min	8
Max	16
Median	12
SD	2.4

In the Case Study Analyses at the end of this article, there is narrative text that highlights relevant information pertinent to each of the Expedited Programs Guidance factors for each of these Subpart H approvals. Table 2 summarizes the scoring and scores for the nineteen precedents.

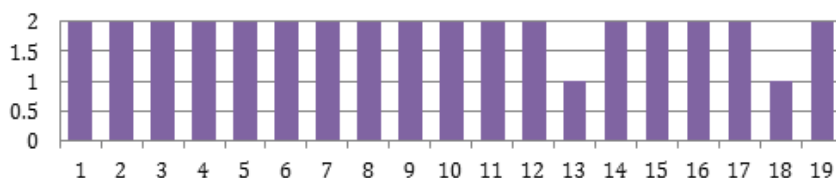
[†] Precedents are listed in reverse chronological order.

Table 2: Subpart H Score Card – Cumulative Results

Factor	Max Possible Score	Actual Score Ranges
Part 1 <i>Rarity of the Disease</i>	2 points (10% of total)	1–2 points
Part 2 <i>Understanding of the Disease Process</i>	4 points (20% of total)	1–4 points
Part 3 <i>Understanding the Relationship Between the Drug’s Effect on Unvalidated Surrogate and the Disease</i>	4 points (20% of total)	1–4 points
Part 4 <i>Strength of Clinical Evidence</i>	10 points (50% of total)	1–9 points
Cumulative Score <i>Parts 1–4</i>	20 points (100% of total)	8–16 points

A. Part 1: Rarity

Part 1 of the analysis of each precedent assesses the first factor, rarity. For each precedent, the assessment of rarity is included under the heading of “Part 1” in Figure 1 and in the narratives for each precedent in the case studies below. The consistency of findings across the nineteen precedents with respect to rarity is robust. *See* Figure 2.

Figure 2: Rarity of the Disease (02)[‡]

This analysis catalogues for the first time (to the authors’ knowledge) that, for those therapies approved by FDA under Subpart H for conditions other than AIDS or cancer, the therapies have nearly always been overtly designated by FDA as orphan drug therapies.³²

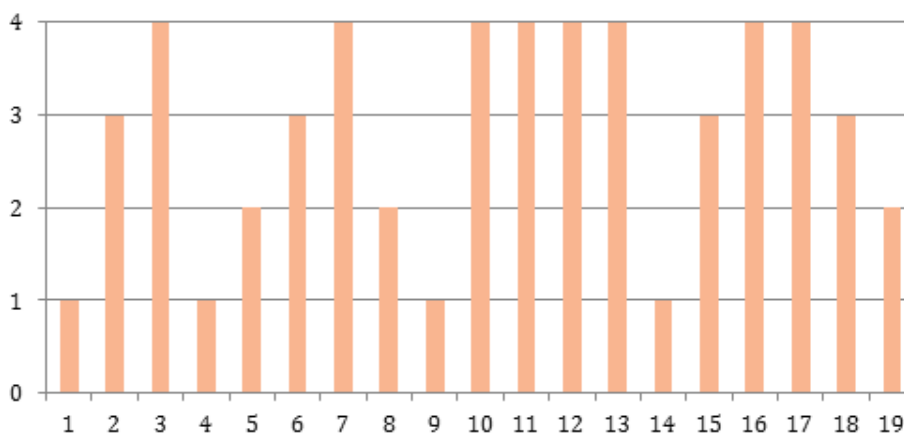
[‡] 1. Northera; 2. Sirturo; 3. Ferriprox; 4. Makena; 5. Soliris; 6. Promacta; 7. Exjade; 8. Tysabri; 9. Luveris; 10. Fabrazyme; 11. Remodulin; 12. Celebrex; 13. Synercid; 14. Remicade; 15. Priftin; 16. Sulfamylon; 17. ProAmatine; 18. Biaxin; 19. Betaseron.

³² While multiple sclerosis (MS) today is not an orphan condition, at the time of Betaseron’s approval in August 1993, Betaseron had been designated as an orphan drug therapy for all of MS, not just for one stage of this disease.

B. Part 2: Understanding of the Disease Process

Understanding of the disease process is the next key factor listed by FDA in the Expedited Programs Guidance. For eight of the nineteen precedents, a maximum score of four was achieved. See Figure 3. This is consistent with FDA’s view that this can be “an important factor in determining whether an endpoint is reasonably likely to predict clinical benefit.”³³

Figure 3: Understanding of the Disease Process (04)^{||}



However, four precedents (Northera for neurogenic orthostatic hypotension, Makena for pre-term birth, Luveris for pregnancy, and Remicade for Crohn’s Disease) received scores of “1” on a scale of 0 to 4 because in each case the pathophysiology of the underlying disease is complex and not so clearly understood. A key takeaway from this observation is that, while a clear understanding of the pathophysiology of the disease process facilitates reliance upon an unvalidated surrogate and/or ICE, the existence of a relatively weak understanding of the disease process is not, in and of itself, incompatible with Subpart H.

C. Part 3: Relationship Between Unvalidated Surrogate and Disease

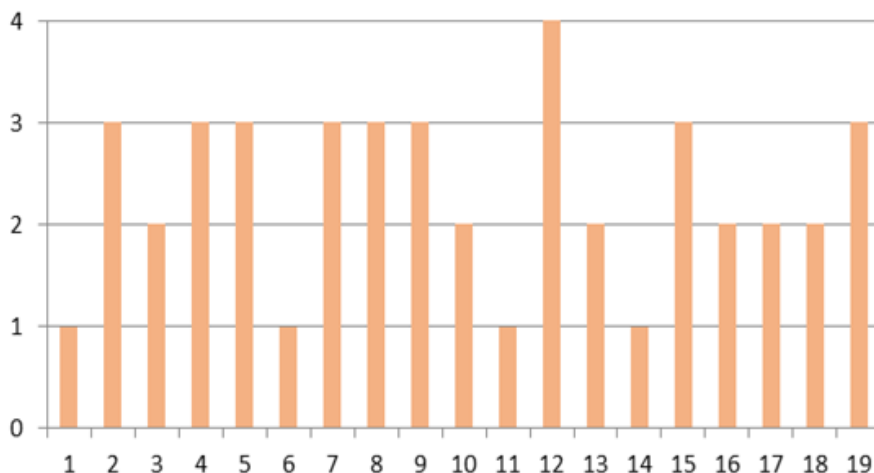
The next key factor listed by FDA in its Expedited Programs Guidance is how well-understood the relationship is between the drug’s effect on the unvalidated surrogate and the disease process. To analyze this factor, the authors searched FDA reviews for evidence of reliance upon epidemiological associations (*see, e.g.*, Sirturo and Makena) and the effect of another drug in the same or pharmacologically similar class of therapy on both the unvalidated surrogate and the disease (*see, e.g.*, Tysabri and Celebrex). In several cases, there was only relatively weak support for the relationship between the unvalidated surrogate and the disease process, such as in the cases of Fabrazyme (in

³³ EXPEDITED PROGRAMS GUIDANCE, *supra* note 3, at 20.

^{||} 1. Northera; 2. Sirturo; 3. Ferriprox; 4. Makena; 5. Soliris; 6. Promacta; 7. Exjade; 8. Tysabri; 9. Luveris; 10. Fabrazyme; 11. Remodulin; 12. Celebrex; 13. Synercid; 14. Remicade; 15. Prifitin; 16. Sulfamylon; 17. ProAmatine; 18. Biaxin; 19. Betaseron.

which little had ever been shown between clearance of substrate in particular cell types and renal function), Promacta, Remodulin, Synercid, Remicade, and Biaxin. *See* Figure 4.

Figure 4: Understanding of the Relationship Between the Drug's Effect on Unvalidated Surrogate and the Disease (0–4)^{††}



Again, as in the case of Part 2, a weaker showing in this particular factor was not a bar to Subpart H qualification.

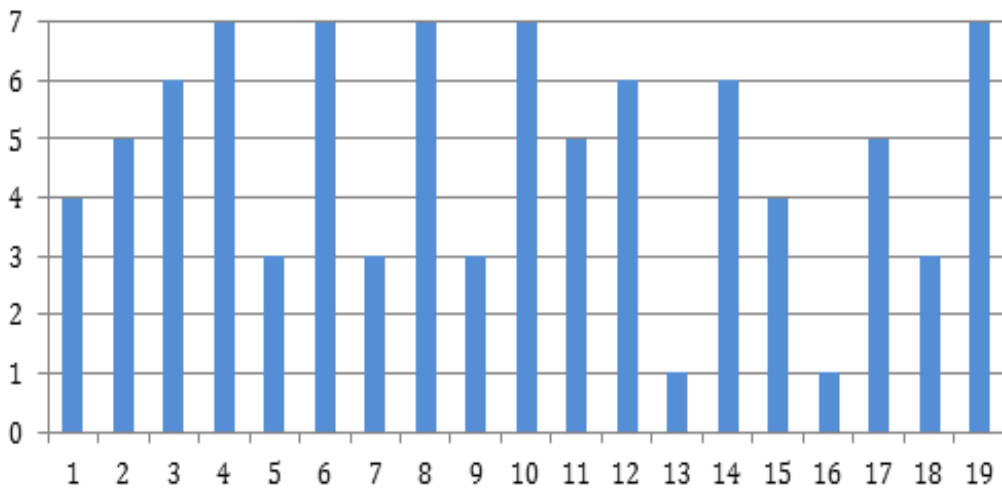
D. Part 4: Clinical Evidence on the Unvalidated Surrogate as Well as on Benefit

Finally, in its Expedited Programs Guidance, FDA noted the critical role of the clinical strength of evidence of the drug primarily on the unvalidated surrogate and/or ICE, but also to a lesser extent on the clinical benefit as well. Therefore, the authors' analysis of clinical evidence is divided into two components of unequal weight: the clinical evidence of the drug on the unvalidated surrogate and/or ICE and the clinical evidence of the drug on the clinical benefit.

1. Clinical Evidence on Surrogate or ICE

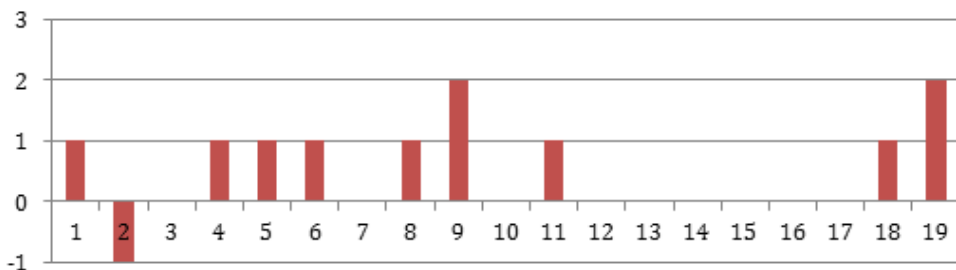
With regard to the strength of clinical evidence on their unvalidated surrogate endpoints or ICE, it is not surprising that five of the nineteen precedents garnered the highest rating of "7" on a scale of 0 to 7. (Note that this factor was given the greatest weight in the overall analysis because it is viewed, by FDA and the authors, as the single most important factor.) However, even therapies such as Sulfamylon and Synercid, which had extremely weak strength of clinical evidence on their respective unvalidated surrogates, were judged by FDA as appropriately qualified for Subpart H, carried mainly on the strength of the other factors described in FDA's Expedited Programs Guidance. *See* Figure 5.

^{††} 1. Northera; 2. Sirturo; 3. Ferriprox; 4. Makena; 5. Soliris; 6. Promacta; 7. Exjade; 8. Tysabri; 9. Luveris; 10. Fabrazyme; 11. Remodulin; 12. Celebrex; 13. Synercid; 14. Remicade; 15. Prifitin; 16. Sulfamylon; 17. ProAmatine; 18. Biaxin; 19. Betaseron.

Figure 5: Clinical Evidence on Surrogate or ICE (0–7)^{}**

2. Clinical Evidence of Clinical Benefit

The second half of the assessment of overall clinical evidence was the strength of evidence of clinical benefit. Since there is no requirement for any such evidence, it was not anticipated that these scores would be high. In line with these expectations, the Subpart H approval precedents generally contained relatively little clinical evidence of benefit in the clinical data sets that were the basis for each approval. Ten of the nineteen precedents had essentially no substantial positive evidence of clinical benefit, and one of the precedents actually had a fairly strong negative numerical “lean” in clinical outcome evidence, suggesting that the therapy may have a negative impact on long-term clinical benefit. See Figure 6.

Figure 6: Clinical Evidence of Clinical Benefit (0–3)^{†††}

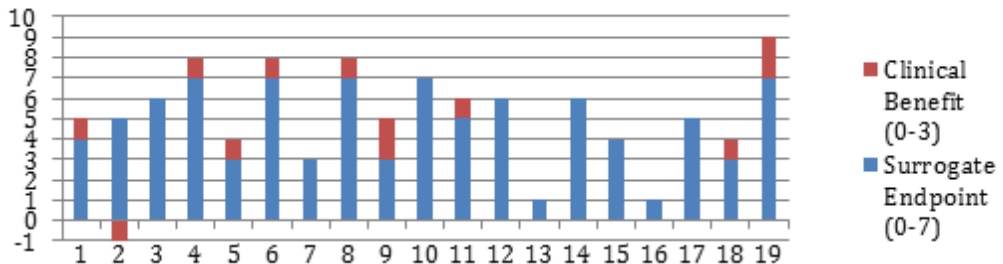
^{**} 1. Northera; 2. Sirturo; 3. Ferriprox; 4. Makena; 5. Soliris; 6. Promacta; 7. Exjade; 8. Tysabri; 9. Luveris; 10. Fabrazyme; 11. Remodulin; 12. Celebrex; 13. Synercid; 14. Remicade; 15. Priftin; 16. Sulfamylon; 17. ProAmatine; 18. Biaxin; 19. Betaseron.

^{†††} 1. Northera; 2. Sirturo; 3. Ferriprox; 4. Makena; 5. Soliris; 6. Promacta; 7. Exjade; 8. Tysabri; 9. Luveris; 10. Fabrazyme; 11. Remodulin; 12. Celebrex; 13. Synercid; 14. Remicade; 15. Priftin; 16. Sulfamylon; 17. ProAmatine; 18. Biaxin; 19. Betaseron.

3. Combined Clinical Evidence

The combined clinical evidence (including clinical evidence on surrogate or ICE and clinical evidence of clinical benefit) varied greatly. On a scale of 0 to 10, one therapy received a combined score of “1” while another therapy received a combined score of “9.” The average score was approximately “5.” See Figure 7.

Figure 7: Combined Clinical Evidence (0–10)^{*}**



V. DISCUSSION

Regulatory ingenuity, if not outright genius,³⁴ led FDA on its own to create the concept of the Subpart H approval in order to address at first, the emerging AIDS epidemic in the 1980s and since then, all other serious conditions for which there is an unmet medical need. The linchpin of FDA Subpart H system was, and is, the unvalidated surrogate endpoint that is “reasonably likely to predict clinical benefit,” and/or the ICE that is “an effect on a clinical endpoint that can be measured earlier than an effect on irreversible morbidity or mortality, that is reasonably likely to predict an effect on irreversible morbidity or mortality or other clinical benefit.”³⁵

There have been many misunderstandings of the Subpart H system. Some have thought FDA’s Subpart H regulations mean that the quantum or quality of evidence was somehow reduced for certain therapies and the statutory requirement of “substantial evidence of effectiveness” was in some way, in whole or in part, skirted or deferred by the establishment of the Subpart H pathway. On the other extreme, others have thought that unless the unvalidated surrogate is validated, it cannot be relied upon in a Subpart H approval decision. This view is just as likely as the first to not “serve the

^{***} 1. Northera; 2. Sirturo; 3. Ferriprox; 4. Makena; 5. Soliris; 6. Promacta; 7. Exjade; 8. Tysabri; 9. Luveris; 10. Fabrazyme; 11. Remodulin; 12. Celebrex; 13. Synercid; 14. Remicade; 15. Priftin; 16. Sulfamylon; 17. ProAmatine; 18. Biaxin; 19. Betaseron.

³⁴ This genius was inspired, at least in part or mainly, by the urgency of the voice of the AIDS patients’ community.

³⁵ FOOD & DRUG ADMIN., DRAFT GUIDANCE FOR INDUSTRY: LABELING FOR HUMAN PRESCRIPTION DRUGS AND BIOLOGICAL PRODUCTS APPROVED UNDER THE ACCELERATED APPROVAL REGULATORY PATHWAY 2 (Mar. 2014), <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm390058.pdf>.

public well”³⁶ and is sometimes found in FDA reviews that conclude that the Sponsor’s evidence failed to satisfy the standard for approval because the showing on clinical benefit was not robust enough to validate the drug’s effect on the unvalidated surrogate.

Between these two extremes, there has existed a fundamental question that has begged to be addressed for well over twenty years and that is this: what is the regulatory and evidentiary foundation for FDA’s determination that an unvalidated surrogate endpoint and/or ICE is capable of supporting a Subpart H approval?

This analysis attempts to apply the factors in FDA’s Expedited Programs Guidance to the nineteen Subpart H approvals (that are not for AIDS or cancer or based on the “Animal Efficacy Rule”) in order to discern the types and patterns of evidence that FDA has found adequate to be the foundation for past Subpart H approvals.

A. Overview of Lessons Learned for Subpart H Actions

1. “You Don’t Need to Knock It Out of the Park”—Part 1: Sponsors Do Not Need Strong Evidence on All Four Risk Factors

Almost all of the nineteen precedents were designated as orphan drugs and, therefore, the scores under Part 1 were uniformly high. As for the relative strength of FDA factors, which this analysis housed under headings of Parts 2, 3, and 4, there were some noteworthy inconsistencies, especially within Part 2 (understanding of the disease process) and the component of Part 4 on the clinical evidence of the drug’s effect on the unvalidated surrogate. Also, of note, a weak assessment or contribution from Part 2, Part 3, or even Part 4 did not prove to be a barrier to qualifying for Subpart H. The weak clinical evidence on the unvalidated surrogate for Synercid and Sulfamylon were the greatest surprises to the authors in that this is regarded as the single most important factor in a Subpart H approval. What this illustrates is the extraordinary exercise of regulatory flexibility by FDA officials manifested in these approval actions.³⁷

2. “You Don’t Need to Knock It Out of the Park”—Part 2: FDA Exercises Extraordinary Flexibility

As with the prior analysis of FDA’s orphan drug precedents by one of the authors, this analysis of FDA’s Subpart H precedents testifies to FDA’s flexibility in applying its standards to therapies under its review. In 2013, both Congress and the president additionally and strongly exhorted FDA to extend and expand its use of Subpart H, especially beyond AIDS and cancer. By interpreting and applying the factors FDA laid out in its Expedited Programs Guidance to these precedents, the authors hope that this analysis will help propel that endeavor.

B. Learnings Concerning Intermediate Clinical Endpoints

FDA’s Expedited Programs Guidance defines “intermediate clinical endpoint” (or ICE) as “a measurement of a therapeutic effect that can be measured earlier than an effect on IMM and is considered reasonably likely to predict the drug’s effect on IMM

³⁶ PCAST REPORT, *supra* note 6, at 59.

³⁷ The authors cannot conduct a parallel assessment of therapies that were considered by FDA under Subpart H, or could have been considered by FDA under Subpart H and have not been approved. What such an analysis would show is undeterminable.

or other clinical benefit.”³⁸ When the Subpart H precedents that relied on ICE are compared to those that relied on unvalidated surrogate endpoints, it is evident that Parts 2 and 3 are often less critical to approval when ICE are employed. *See* Table 3. Three of the four therapies that scored the lowest (i.e., 1 out of 4 points) in “Understanding of the Disease Process” were approved using ICE. These therapies were Northera (for neurogenic orthostatic hypotension), Makena (for pre-term birth), and Remicade (for Crohn’s disease). Similarly, three of the four therapies that scored the lowest (i.e., 1 out of 4 points) in “Understanding the Relationship Between the Drug’s Effect on [Unvalidated] Surrogate and the Disease” were approved using ICE. These therapies were Northera, Remodulin (for pulmonary arterial hypertension), and Remicade.

³⁸ EXPEDITED PROGRAMS GUIDANCE, *supra* note 3, at 18.

Table 3: Subpart H Analysis of Unvalidated Surrogate Endpoints and Intermediate Clinical Endpoints

Drug	Basis of Accelerated Approval		Ultimate Clinical Benefit	Part 2 Score (0-4)	Part 3 Score (0-4)
	Surrogate Endpoint	Intermediate Clinical Endpoint (ICE)			
1. Northera	-	Short term dizziness	Long term dizziness	1	1
2. Sirturo	Sputum culture conversion at 6 months	-	Long term outcomes of failure or relapse or death at least 6 months after all MDR-TB treatment is completed.	3	3
3. Ferriprox	Serum ferritin	-	Long-term safety and clinical benefit, including efficacy in prolonging survival or improving disease-related symptoms or function	4	2
4. Makera	-	Reduction of preterm birth defined as less than 37 weeks	Infant outcome, including death	1	3
5. Soliris	-	Effects on TMA & renal function	Renal failure, death, and stroke	2	3
6. Promacta	An increase from the baseline platelet count to a count greater than or equal to 50,000/mcl	-	Excessive bleeding/bruising		
7. Exjade	Reduction of liver iron concentration	-	Evidence of control or improvement in measures of iron overload (such as serum ferritin) and/or clinical manifestations of iron overload	3	1
8. Tysabri	-	Decrease in relapse rate over the course of one year	Decrease in relapse rate over two years	4	3
9. Luvris	Follicular development as defined by three parameters: appropriate estradiol levels, ultrasound follicular measurement, and mid-luteal progesterone levels (all of which had to be satisfied)	-	Ovulation and clinical pregnancy	2	3
10. Fabrazyme	Intracellular substrate accumulation in the vascular endothelium	-	Progression of renal disease and other significant clinical events; effects of Fabryzyme on creatine over time	1	2

Table 3 (continued): Subpart H Analysis of Unvalidated Surrogate Endpoints and Intermediate Clinical Endpoints

Drug	Basis of Accelerated Approval		Ultimate Clinical Benefit	Part 2 Score (0-4)	Part 3 Score (0-4)
	Surrogate Endpoint	Intermediate Clinical Endpoint (ICE)			
11. Remodulin	-	Combined exercise/Borg score analysis	Time to first occurrence of death, hospitalization for complications of pulmonary hypertension, need for esoprostenol, or other clear evidence of deterioration	4	1
12. Celebrex	-	Percent change in the number of colorectal adenomas	Impact on incidence of FAP-related events (e.g., polypectomy, surgery, cancer, desmoids, death)	4	4
13. Synercid	Clearance of VREF bacteremia	-	Cure of underlying infection	4	2
14. Remicade	-	Clinical response, defined as a reduction from baseline in the CDAI score of at least 70 points, at the 4-week evaluation	Maintaining a sustained clinical outcome in patients with moderately to severely active Crohn's disease, including patients with draining enterocutaneous fistula(s)	1	1
15. Pritin	Negative sputum culture at 6 month post-treatment (6 month relapse rate)	-	Negative sputum culture up to 2 years post-treatment. Inability to produce sputum is considered a "cure."	3	3
16. Sulfamylon	Necessity to change topical antimicrobial treatment during the first 5 days of application due to infection or colonization	-	Primary endpoint of Phase 4 trial is treatment failure. Acceptable outcomes: Sulfamylon is proven to be superior or equivalent to standard of care in the primary endpoint.	4	2
17. ProAmatine	Increase in 1-minute standing systolic blood pressure	-	Improved ability to carry out activities of daily living	4	2
18. Blaxin	Clearance of bacteremia	-	Clinical cure/mortality	3	2
19. Betaseron	MRI evaluations of brain lesions	Rate and extent of MS exacerbations	4-6 year study using disability as measured by EDSS as a primary endpoint; usefulness of continued treatment beyond two years; and correlation between MRI imaging and the two endpoints above	2	3

This analysis shows that, when FDA approves a therapy via the Subpart H pathway using ICE instead of an unvalidated surrogate, the Agency reduces the criticality or significance of two of the three factors it considers in its Subpart H approval decisions (i.e., the need to have a clear understanding of the disease as well as the need to establish the relationship between the Subpart H endpoint and the disease). Despite this, FDA has only approved seven Subpart H drugs on the basis of ICE alone. ICE, therefore, presents a largely untapped opportunity for Subpart H approval.

FDA's most recent non-AIDS and non-cancer Subpart H approval, on February 18, 2014, provides an illustrative example. Northera was approved for the treatment of "orthostatic dizziness, lightheadedness, or the 'feeling that you are about to black out' in adult patients." It was approved on trials that relied on an ICE, specifically: a short-term benefit or acute improvement on dizziness, which is the main symptom and disability of the disease. By relying upon this ICE, FDA had greater confidence that the confirmatory Phase 4 study will confirm the longer term, durable, or chronic continued benefit of the *same* symptom improvement that was shown in the acute setting in clinical trials and was the basis for Northera's Subpart H approval.

Because Northera's ICE is also the ultimate clinical benefit (though in an acute setting only), understanding of the disease process and understanding of the relationship between the drug's effect on the Subpart H endpoint and the disease were less important than in Subpart H approval decisions that rely on unvalidated surrogate endpoints. In other words, because the primary endpoint of short-term dizziness (i.e., the ICE) is the primary endpoint that will be tested in a chronic setting in the confirmatory Phase 4 study, the degree of regulatory uncertainty is reduced relative to an approval based on an unvalidated surrogate. Therefore, the amount of evidence needed for understanding the disease and understanding the relationship between the endpoint and the disease is lessened.

One additional takeaway from the analysis concerning ICE is that an unvalidated surrogate may be coupled with an intermediate clinical endpoint to provide the "substantial evidence" needed, as was done with Betaseron. In that case, approval was based on a finding on an unvalidated endpoint (MRI lesion volume) combined with an intermediate clinical endpoint (rate and extent of multiple sclerosis exacerbations).

Therefore, if Sponsors and FDA base their Subpart H programs and approvals on ICE, either alone or coupled with evidence from an unvalidated surrogate, the demands on Sponsors and FDA during the review process and, more importantly, in FDA approval decisions may be reduced. Most critically, patients in need of these therapies for their serious diseases where there is no available therapy will have therapies that have been approved even where the understanding of the disease's pathophysiology may not yet be clear.

C. Use of Multiple Surrogates and/or ICE

FDA has used multiple unvalidated surrogate endpoints as a basis for Subpart H approval for three of the nineteen non-cancer, non-AIDS drugs: Betaseron, Luveris, and Soliris. While the FD&C Act generally discusses Subpart H in terms of a single unvalidated surrogate endpoint (e.g., it discusses the "effect on a surrogate"

and “evidence to support that an end-point is reasonably likely to predict”),³⁹ FDA has on at least three occasions determined that two or more unvalidated surrogate endpoints, collectively, were “reasonably likely to predict” clinical benefit.

One example is Betaseron, which in 1993 represented a breakthrough therapy for patients with multiple sclerosis. For that approval, Bayer Healthcare Pharmaceuticals was able to show the efficacy of its therapy based on not only a reduction of exacerbations (an ICE), but also on improvement on magnetic resonance imaging (MRI) (an unvalidated surrogate). Based on the results from *two* non-irreversible morbidity or mortality endpoints (i.e., an unvalidated surrogate *plus* an intermediate clinical endpoint), FDA approved Betaseron using the Subpart H pathway. FDA stated in its review memoranda:

It was also clear that the Committee as a whole placed great weight on the MRI findings in their deliberations. Specifically, although the clinical benefit, as measured by the proportion of exacerbation-free patients and exacerbation frequency, was considered real and of value clinically, the Committee considered the size of the treatment effect relatively small.⁴⁰

In the Betaseron data there is a second kind of replication, the MRI results, which are more or less persuasive, depending on one’s beliefs. At a minimum, as Dr. Leber says, these [MRI surrogate] data are an independent measurement that supports the clinical finding, a kind of “within-study” replication. At best, they are evidence of an effect far more important than the modest effect on exacerbations. I certainly am not qualified to choose between these interpretations, but our advisors seem to believe the latter, even though all would agree that, strictly, the correlation of improved clinical outcome and improved MRI has not been made. . . . It would be possible, I believe, to grant approval under the Accelerated Approval Regulations, which allow this procedure where a surrogate or clinical, but *non-ultimate endpoint* is the basis for approval.⁴¹

As FDA found during its review of Betaseron, findings on two or more unvalidated surrogate and/or ICE endpoints can, when they are independent measures, serve as a type of replication that can work to strengthen the clinical evidence.

D. Weighing Uncertain Benefits and Risks

Given the low number of Subpart H approvals for non-cancer, non-AIDS drugs since 1993, it is apparent that FDA has taken a risk-adverse approach towards approving drugs via the Subpart H pathway. FDA’s Expedited Guidance Document states the Agency’s concerns:

³⁹ FD&C Act § 506(c)(1), 21 U.S.C. § 356 (2015).

⁴⁰ Memorandum from Russell Katz, Supervisory Clinical Review of Prod. Licensing Application (PLA) 92-0495 for Betaseron for the Treatment of Patients with Exacerbating Remitting Multiple Sclerosis, (May 28, 1993), as reprinted in Ctr. for Drug Evaluation & Research, Application No. 103471 Med. Review at 356–57.

⁴¹ Memorandum from Robert Temple, Dir., Office of Drug Evaluation I, Ctr. for Drug Evaluation & Research, to Janet Woodcock, Dir., Office of Therapeutics Research & Review, Ctr. for Biologics Evaluation & Research, (June 3, 1993), as reprinted in CTR. FOR DRUG EVALUATION & RESEARCH, Application No. 103471 Admin. & Correspondence Documents, at 329–30.

The principal risk of this approach is the possibility that patients will be exposed to a drug that ultimately will not be shown to provide an actual clinical benefit. In addition, there generally will be fewer, smaller, or shorter clinical trials than is typical for a drug receiving traditional approval, which may generally mean there is less information about the occurrence of rare or delayed adverse events. Uncertainty about whether clinical benefit will be verified and the possibility of undiscovered risks are the primary reasons that accelerated approval is reserved for drugs intended to treat a serious condition and that appear to provide a meaningful advantage over available therapy.⁴²

While FDA's task of preventing harm is of paramount importance, rather than generally ignoring or rejecting the use of the Subpart H pathway, the Agency could work to understand its potential safety concerns on a case-by-case, granular basis and find ways to address those specific concerns. By extension, if the trials underlying a Subpart H approval are not, in fact, "fewer, smaller, or shorter . . . than is typical for a drug receiving traditional approval," then FDA should conclude that the possibility of "undiscovered risk" has been addressed. In addition, Sponsors and FDA should work to develop confirmatory and other postmarketing studies that adequately address potential "undiscovered risk." Given that the Subpart H pathway is reserved for therapies with the potential to address unmet medical needs for serious and life-threatening diseases, the commensurate risk of not utilizing the Subpart H pathway (and rejecting a potentially beneficial therapy) must be weighed.

VI. CONCLUSION

This article intends to contribute to an understanding of the strength of scientific and clinical evidence in FDA's reaching its nineteen Subpart H approval determinations over the past twenty-four years. It is also intended, at least partially, to lift the veil obscuring the basis for FDA's determination when an unvalidated surrogate and/or ICE is "reasonably likely to predict clinical benefit." The authors conclude that, in practice, FDA exercises extraordinarily more regulatory flexibility than either FDA's foundational statutes or even FDA's most recent 2014 Subpart H guidance explicitly express.

⁴² EXPEDITED PROGRAMS GUIDANCE, *supra* note 3, at 16.

CASE STUDY ANALYSES

I. NORTHERA (droxidopa)

FDA's February 18, 2014 approval for treatment of orthostatic dizziness, lightheadedness, or the "feeling that you are about to black out" in patients with symptomatic neurogenic orthostatic hypotension caused by primary autonomic failure (Parkinson's disease, Multiple System Atrophy, and Pure Autonomic Failure), dopamine beta-hydroxylase deficiency, and non-diabetic autonomic neuropathy relied upon the effectiveness at one week as the surrogate endpoint. Applying the terms of the Expedited Programs Guidance, this would be an intermediate clinical endpoint that would be reasonably likely to predict the longer term benefit.

Part I. Regulatory Factors Weighing into FDA Determination

a. Severity of the Condition

Orthostatic hypotension may be a severely disabling condition which can seriously interfere with the quality of life of afflicted subjects. Some patients become confined to a wheelchair and some become bedridden. There are no currently available therapeutic options that have been demonstrated to have symptomatic benefit.⁴³

b. Rarity of the Condition

Northera was designated as an orphan drug on January 17, 2007.

c. Lack of Available Therapy

NOH is a rare condition, but many of the patients who are afflicted are profoundly symptomatic and have few treatment options.⁴⁴ Midodrine is the only other approved treatment for symptomatic neurogenic orthostatic hypotension (NOH). . . . Midodrine received accelerated approval in 1996 on the basis of an increase in standing SBP, a surrogate endpoint reasonably likely to predict clinical benefit; however, subsequent clinical trials have not shown that midodrine improves symptoms.⁴⁵ The limitations of currently available therapeutic options, and the incapacitating nature and often progressive downhill course of disease, point to the need for an improved therapeutic alternative.⁴⁶

⁴³ Clinical Review from Melanie Blank 16 (Jan. 27, 2012), *as reprinted in* Ctr. for Drug Evaluation & Research, Application No. 203202Orig1s000 Medical Review(s), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/203202Orig1s000MedR.pdf.

⁴⁴ Memorandum from Ellis F. Unger, Acting Dir., Office of Drug Evaluation I, Office of New Drugs, Ctr. for Drug Evaluation & Research 16 (Mar. 28, 2012), *as reprinted in* Ctr. for Drug Evaluation & Research, Application No. 203202Orig1s000 Medical Review(s), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/203202Orig1s000MedR.pdf.

⁴⁵ Clinical Review from Shari L. Targum 8 (Dec. 5, 2013), *as reprinted in* Ctr. for Drug Evaluation & Research, Application No. 203202Orig1s000 Medical Review(s), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/203202Orig1s000MedR.pdf.

⁴⁶ Blank, *supra* note 43, at 20–21.

d. Use of External Expertise

FDA's Cardiovascular and Renal Drugs Advisory Committee met on February 23, 2012, and supported approval of the treatment with a vote of 7–4–1 (with one additional member not voting).

Part 2. Understanding of the Disease Process

The biological means by which NOH progresses is complex and multifaceted. “Autonomic dysfunction disorders, encompassing Pure Autonomic Failure (PAF . . .), Multiple System Atrophy (MSA . . .), Parkinson's Disease (PD), Non-Diabetic Autonomic Neuropathy (NDAN), or Dopamine Beta Hydroxylase (DBH) Deficiency . . . differ in etiology and pathophysiology; however, each condition is accompanied by a deficiency of [norepinephrine (NE)].”⁴⁷ “The diseases classified under primary autonomic failure (PAF, MSA, and PD) are all neurodegenerative and of unknown etiology.”⁴⁸ “[NDAN] can be caused by a number of factors, including autoimmune, environmental, and infectious agents. These conditions are associated with either degradation of peripheral NE nerve function or failure of the central mechanism controlling the release of NE. The cause of DBH Deficiency is a rare genetic mutation that results in the loss of this key enzyme in NE production, resulting in a global NE deficiency and a surplus of NE precursor dopamine.”⁴⁹

“While the mechanism is not well characterized, NE presumably binds to alpha adrenergic receptors in the vascular smooth muscle of arterioles causing vasoconstriction and consequent elevation of systolic blood pressure. Norepinephrine may also have an effect on venous vascular resistance. By elevating the blood pressure, it promotes the maintenance of cerebral blood flow, thereby lessening the symptoms of neurogenic orthostatic hypotension; primarily dizziness, lightheadedness and syncope. Droxidopa crosses the blood brain barrier and therefore may exert its effect both peripherally and centrally by increasing NE production.”⁵⁰ FDA has, however, noted that increases in 1-minute standing systolic blood pressure in midodrine have not meaningfully predicted the clinical benefit of the drug.

Part 3. Understanding of the Relationship Between OHSA Item 1 at Week 1 and OHSA Item 1 Over the “Longer Term”

The understanding of the relationship between OHSA Item 1 at Week 1 and OHSA Item 1 over the “longer term” is not well-understood. In fact, there is no generally established definition for “longer term” for this therapy. In Northera's Package Insert, FDA defined “longer-term” as 8 weeks and 3 months. During the review, the Medical Reviewer indicated that a durable effect may be considered “more than 4 weeks.”⁵¹ “Droxidopa is a synthetic catecholamine acid analogue that is metabolized by dopa decarboxylase to norepinephrine, which is thought to increase blood pressure (BP) through binding and activation of adrenergic receptors. The applicant also asserts that

⁴⁷ *Id.* at 21.

⁴⁸ *Id.* at 17.

⁴⁹ *Id.* at 21.

⁵⁰ *Id.* at 26.

⁵¹ *Id.* at 11.

droxidopa increases neuronal levels of NE, which could lead to sustained effects.”⁵² Despite this, the link between short term and longer term effects are unclear. “Droxidopa . . . has been approved in Japan since 1989 for essentially the same indication.”⁵³

*Part 4. Clinical Evidence on OHS A Item 1 at Week 1
(Surrogate Endpoint) and Over the Longer Term (Ultimate
Clinical Endpoint)*

a. Surrogate: OHS A Item 1 at Week 1

The primary endpoint in Study 306B was OHS A Item 1 score, defined as “dizziness, lightheadedness, feeling faint, and feeling like you might black out” at Week 1. Study 306B was statistically significant on its primary endpoint ($p=0.028$). Patients in Study 306B also had a greater improvement in Week 1 lowest standing blood pressure within three minutes after standing compared with patients on placebo ($p=0.032$).⁵⁴ In Study 301, the primary endpoint was Orthostatic Hypotension Questionnaire (OHQ), a patient reported outcome that measures symptoms of NOH and their impact on patients’ daily life. OHQ includes OHS A Item 1 in its composite score. There was no statistically significant effect on OHQ ($p=0.19$); however, on OHS A Item 1, there was a decrease in dizziness at one week compared with placebo ($p=0.06$).⁵⁵

b. Ultimate Clinical Outcome: OHS A Item 1 over “Longer Term”

Study 302 ($n=101$) was a placebo-controlled [two]-week randomized withdrawal study of Northera in patients with symptomatic NOH. Study 303 ($n=75$) was an extension of Studies 301 and 302, where patients received their titrated dose of Northera for [three] months and then entered a [two]-week randomized withdrawal phase. Neither study showed a statistically significant difference between treatment arms on its primary endpoint.⁵⁶ [FDA thus concluded that] the effectiveness of Northera beyond two weeks is uncertain, and patients should be evaluated periodically to determine whether Northera is continuing to provide a benefit.⁵⁷

2. SIRTURO (bedaquiline)

This December 28, 2012 approval for treating multi-drug resistant tuberculosis (MDR-TB) was based on a surrogate of time to sputum culture conversion.

⁵² Unger, *supra* note 44, at 2.

⁵³ *Id.* at 3.

⁵⁴ NORTHERA (DROXIDOPA) LABEL 7 (Feb. 2014), http://www.accessdata.fda.gov/drugsatfda_docs/label/2014/2032021bl.pdf.

⁵⁵ *Id.* at 9.

⁵⁶ *Id.*

⁵⁷ *Id.*

Part 1. Regulatory Factors Weighing into FDA Determination

a. Severity of the Condition

“Overall mortality still exceeds 10%, with a range of 8 to 21% for patients enrolled into good treatment programs.”⁵⁸

b. Rarity of the Condition

FDA granted Sirturo orphan drug designation on January 10, 2005. Furthermore, in FDA’s determination that the time to sputum culture conversion is an acceptable surrogate on which to base accelerated approval, it appears that FDA may have taken into account specifically the rarity of MDR-TB in this country in that FDA acknowledged that, “[i]n the United States, the total number of primary MDR-TB cases has fluctuated from 88 to 132 cases [since] 1993, with 88 cases reported in 2010.”⁵⁹

c. Lack of Available Therapy

Treatment of MDR-TB is more complex [than treating drug-susceptible TB or DS-TB] and prolonged and typically has a favorable outcome rate [of only] 41-70%. Cases of MDR-TB are currently treated with at least five second-line anti-TB drugs for an extended period of time that may last up to two years. . . . The challenges of the treatment of MDR-TB include toxicities of the drugs, decreased potency, cost (50-200 times more expensive than DS-TB) and the need for possible hospitalization.⁶⁰

d. Use of External Expertise

FDA did turn to the Anti-Infective Drug Advisory Committee for external expertise, which on June 3, 2009, “voted 18 to 1, recommending that sputum culture conversion . . . could be used as a surrogate . . . [t]herefore, the committee recommended that approval of an antimycobacterial drug could be done under Subpart H regulations (Accelerated Approval) using sputum culture conversion as a surrogate endpoint. Further, traditional endpoints used to evaluate treatment response such as relapse, failure, and mortality should still be used . . . for traditional approval.”⁶¹

Part 2. Understanding of the Disease Process

The pathophysiology of MDR-TB is well-understood.

⁵⁸ Sirturo (bedaquiline), Clinical Review, NDA 204384, 22 (Dec. 29, 2012), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2012/204384Orig1s000MedR_.pdf.

⁵⁹ *Id.*

⁶⁰ *Id.*

⁶¹ *Id.* at 28.

Part 3. Understanding of the Relationship Between Sputum Culture Conversion and Relapse, Long-Term Response, and Mortality

Epidemiologic evidence exists that supports the relationship between sputum culture conversion and clinical outcome, in particular, mortality.⁶²

Part 4. Clinical Evidence of Sirturo's Effect on Sputum Culture Conversion and on Relapse and Mortality

The FDA Medical Reviewer noted the existence of the epidemiological evidence, but stressed that the clinical evidence provided by the Sponsor both on the surrogate and on traditional endpoints of clinical benefit, especially mortality, would be “most persuasive.” In this case, the Medical Reviewer listed these traditional endpoints as relapse, long-term response, and mortality.⁶³

There were two Phase 2 clinical trials that comprised the clinical evidence for this drug on the surrogate and on clinical benefit, but only one of which was considered to be the single, pivotal trial: Study C208 Stage 2. Study C208 Stage 2 was a randomized, double-blinded, placebo-controlled trial with a 24-week treatment period in which both the drug and “placebo” arms received an optimized background regimen.⁶⁴

a. Sputum Culture Conversion

The primary endpoint, which was the surrogate endpoint, of the time to sputum culture conversion was highly statistically significant (p-value of 0.0005) (N=160 randomized, with 67 and 66 subjects in the drug and placebo arms in the mITT analysis, respectively). Sputum culture conversion at week 24 was a key secondary endpoint (as well as another supportive measure of the surrogate endpoint of sputum culture conversion), and it too was statistically significant (p-value = 0.014) with 78% and 58% of drug and placebo arm subjects, respectively, achieving sputum culture conversion at week 24.⁶⁵ “Lastly culture conversions data after all patients completed 72 weeks in the study showed a statistically significant but diminishing improvement in the time to sputum culture conversion for [Sirturo]-treated patients compared to placebo-treated patients.”⁶⁶

b. Relapse and Mortality

Relapse is a “traditional” measure of clinical benefit. The Medical Reviewer notes that in “the mITT population, five subjects (7.6%) in the [drug] group and eight subjects (12.1%) in the placebo group experienced relapse. . . . [However,] the subjects in the placebo group appear to take a longer time from culture conversion to relapse

⁶² See Shama D. Ahuja et al., *Multidrug Resistant Pulmonary Tuberculosis Treatment Regimens and Patient Outcomes: An Individual Patient Data Meta-analysis of 9,153 Patients*, 9 PLOS MED. e1001300 (2012).

⁶³ Clinical Review, *supra* note 58, at 16.

⁶⁴ Sirturo (bedaquiline), Statistical Review, NDA 204384, 6 (Dec. 4, 2012), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2012/204384Orig1s000StatR.pdf.

⁶⁵ *Id.*

⁶⁶ Clinical Review, *supra* note 58, at 44.

than those in the [drug] group.”⁶⁷ Therefore, the Medical Reviewer conducted an alternative analysis, and in this analysis, “the two treatment arms become more comparable with respect to relapse with 5 relapses on [drug] and 4 on placebo.”⁶⁸

Survival is the most objective and clinically meaningful benefit in MDR-TB. In the pivotal study, 9 of 79 in the drug arm died (11.4%) compared to 2 of 81 (2.5%) in the placebo arm.⁶⁹ Both placebo subjects died of TB as did 5 of the 9 subjects in the drug group.⁷⁰ Signals of QT prolongation and serum transaminase elevation, with one death due to liver injury in the drug arm, were also observed.⁷¹

In the “summary and conclusions” section of the statistical review, FDA observed: “There was a statistically significant increase in mortality in the [drug] group. Despite the observed treatment benefit in time to culture conversion, it did not lead to a benefit in patient survival. This was a major concern both for efficacy and safety.”⁷²

The relationship between the traditional clinical endpoints of relapse and survival and the surrogate endpoint of sputum culture conversion were not robust in this case. In fact, the clinical evidence on survival was actually and strongly in the wrong numerical direction.⁷³ Notwithstanding this, FDA appears to have, as noted in its Expedited Programs Guidance, relied in part on the “external expertise” of the June 2009 Anti-Infective Drug Advisory Committee as well as took “into account” these three factors that were listed in FDASIA: (1) the “severity” of the disease; (2) the “rarity” of the disease; and (3) the “lack of alternative treatments.”^{74, 75}

3. *FERRIPROX (deferiprone)*

FDA approved Ferriprox on October 14, 2011 as an iron chelator for the treatment of patients with transfusional iron overload due to thalassemia syndrome when current chelator therapy is inadequate. Ferriprox was approved on the basis of its showing on an unvalidated surrogate, serum ferritin.

⁶⁷ *Id.* at 59-60.

⁶⁸ *Id.* at 60.

⁶⁹ *Id.* at 70.

⁷⁰ *Id.*

⁷¹ *Id.* at 70-71.

⁷² Statistical Review, *supra* note 64, at 60.

⁷³ This is the reason for the authors scoring clinical evidence on the actual clinical benefit as -1 on scale of 0 to 3. The scale was set up under the assumption that, at worst, there would be an absence of any clinical evidence of benefit, or if clinical evidence, then not even any “lean” in favor of the investigational treatment, which then would have been rated as “0.”

⁷⁴ *See, e.g.*, Clinical Review, *supra* note 58, at 59.

⁷⁵ Guidance for Industry, *supra* note 3, at 15. In addition to Dr. Porcalla’s Medical Review reaching this conclusion, every other review unanimously supported a recommendation for approval. For instance, the statistical review by Dr. Xianbin Li concluded, “The efficacy in terms of a surrogate endpoint, sputum culture conversion, was supported by the pivotal study C208 and supportive study C209. There was a significantly elevated mortality risk in the [Sirturo] group. This should be considered in an approval decision and use of this regimen.” Statistical Review, *supra* note 64, at 60. The reviews of the Cross-Discipline Team Leader, Dr. Navarro (Dec. 21, 2012), the Deputy Division Director, Dr. Laessig (Dec. 27, 2012) and the Office Director, Dr. Cox (Dec. 28, 2012) all recognized the robust finding on the surrogate endpoint of sputum culture conversion and recommended approval despite serious consideration of the clinical safety results, especially the survival results in the pivotal study. This unanimity of support for a Subpart H approval decision within the entirety of the internal FDA expert review team was not always observed in the other 18 Subpart H precedents.

Part I. Regulatory Factors Weighing into FDA Determination

a. Severity of the Condition

Persons with certain inherited anemias, especially sickle cell anemia and thalassemia, require frequent red blood cell (RBC) transfusions because they are unable to manufacture hemoglobin. Each unit of packed RBCs contains 200 mg of iron, which is an extreme excess of iron as compared with the dietary intake of 1 mg of iron necessary to maintain normal total body iron stores in healthy individuals. Without a way for the body to excrete excess iron, persons receiving these regular transfusions of RBCs build up massive iron overload which leads to morbidity and often eventually death due to cardiac damage.⁷⁶

b. Rarity of the Condition

FDA designated Ferriprox as an orphan drug on December 21, 2001.

c. Lack of Available Therapy

At the time of Ferriprox's approval, there were two other approved therapies for iron overload due to transfusions: Desferal (deferoxamine) and Exjade (deferasirox). Ferriprox was given fast track designation in January 2004, before Exjade was approved. Exjade, an orally active iron chelator, was approved in 2005. In January 2004, Desferal was the only available therapy and requires continuous infusion over many hours, every day.

The Sponsor first submitted its NDA seeking an indication for "all transfusion-dependent anemias for whom the use of other iron chelators has been considered inappropriate." A complete response letter was issued in November 2009 and a resubmission was made in April 2011 for essentially the same second-line use. However, the data submitted were almost exclusively from thalassemia patients and FDA's October 2011 approval is for "patients with transfusional iron overload due to thalassemia syndromes when direct chelator therapy is inadequate." For this specific use, there is a lack of available therapy.

d. Use of External Expertise

FDA appears to have given consideration to two types of external expertise. First, FDA seems to have given some weight to the "expertise" of clinical practice that uses serum ferritin to monitor the patient's iron status. While serum ferritin is a non-specific endpoint for which FDA noted that "the relationship between the serum ferritin and clinical outcome is not well-established,"⁷⁷ FDA nevertheless appears to give serum ferritin some weight because serum ferritin is "a commonly used parameter for following body iron burden in patients undergoing chronic red blood cell transfusions,"⁷⁸ and because "in clinical practice, measurements of serum ferritin and

⁷⁶ Ferriprox (deferiprone), Medical Review, NDA 021825, 1-2 (Sept. 27, 2011), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021825Orig1s000MedR.pdf.

⁷⁷ Ferriprox (deferiprone), Medical Review, NDA 021825, 34 (Sept. 16, 2011), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021825Orig1s000MedR.pdf.

⁷⁸ Medical Review, *supra* note 76, at 12.

[liver iron concentration] have been the generally accepted methods of evaluation of the efficacy of therapy in persons with iron overload.”⁷⁹

Second, the Oncology Drugs Advisory Committee recommended Ferriprox for approval on September 14, 2011 by a vote of 10 to 2 for treating patients in whom current chelator therapy is inadequate.

Part 2. Understanding of the Disease Process

In this case, the pathophysiology by which iron overload leads to deposition of iron in tissues and leads to iron-catalyzed peroxidation of membrane lipids, which then leads to morbidity and death due to cardiac damage, is well-known.⁸⁰

Part 3. Understanding of the Relationship Between the Effect on Serum Ferritin and Cardiotoxicity and Death

The mechanism of the drug’s action is well-known, that is, binding to iron in a 3:1 complex which is excreted in the urine, and the reduction in iron in these persons is needed to avoid iron overload morbidities.⁸¹ However, serum ferritin is non-specific and “changes in serum ferritin are difficult to interpret because serum ferritin is subject to variations induced by a number of mechanisms that are unrelated to total body iron.”⁸² Most of all, “[t]he relationship between the serum ferritin and clinical outcome is not well established.”⁸³

This part was scored a 2 on a scale of 0 to 4, mainly on the basis of the biologic plausibility that this drug, due to its mechanism, would reduce iron stores, notwithstanding the weakness of serum ferritin itself as a surrogate, due to its lack of specificity as a measure of iron stores. The non-specificity of serum ferritin and the lack of understanding of the relationship between the surrogate and outcomes led to a score of 2 instead of 3.⁸⁴

Part 4. Clinical Evidence of Ferriprox’s Effect on Serum Ferritin and on Clinical Outcome

It is of value here to note that FDA rejected the original NDA submitted in 2009 for Ferriprox because the “primary efficacy endpoint of the single major controlled trial . . . was the change in cardiac MRI T2* which was said to measure iron content within the heart. FDA stated that this endpoint was a surrogate endpoint and *there were no data to support the incremental changes in the values as predictive of clinical benefit.*”⁸⁵ Moreover, “secondary endpoints [of serum ferritin and liver iron

⁷⁹ Ferriprox (deferiprone), Medical Review, NDA 021825, 5 (Nov. 25, 2009), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021825Orig1s000MedR.pdf.

⁸⁰ Medical Review, *supra* note 76, at 1.

⁸¹ *Id.* at 2.

⁸² Ferriprox (deferiprone), Clinical Review, NDA 021825, 15 (Oct. 19, 2009), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021825Orig1s000MedR.pdf.

⁸³ Ferriprox (deferiprone), Medical Review, NDA 021825, 34 (Sept. 27, 2011), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021825Orig1s000MedR.pdf.

⁸⁴ Others may score this differently, perhaps even only as a “1” given the non-specificity of serum ferritin and lack of well-established relationship between surrogate and outcomes.

⁸⁵ Ferriprox (deferiprone), Medical Review, NDA 021825, 10 (Sept. 27, 2011), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021825Orig1s000MedR.pdf (emphasis added).

concentration] also were not consistently corroborative of the primary endpoint [MRI T2*] results.”⁸⁶ Overall, “the study did not find a significant correlation between change in cardiac MRI T2* and measures of cardiac function and there were no differences between treatments in change in liver iron concentration (LIC).”⁸⁷ The statistical review observed that “the patients in this study were not followed for clinical outcomes and therefore, this study was not designed to obtain internal validation of MRI T2* change as a surrogate for any clinical outcome indicative of reduced cardiac iron.”^{88, 89}

Although the data from this study provided statistically significant evidence . . . in MRI T2* . . . this study was not designed to and therefore does not provide evidence that change in MRI T2* is reasonably likely to predict clinical benefit due to lack of long-term follow-up of these patients.⁹⁰

In response to FDA’s rejection of the original NDA, the Sponsor “conducted an analysis of a subpopulation of patients drawn from its previously conducted studies and defined as being inadequately treated with current chelator therapy.”⁹¹ In this analysis, approximately 50% met the primary efficacy endpoint of having a 20% or greater decline in serum ferritin. Of additional importance, the Sponsor-defined “success rate” in this same analysis was 42% for liver iron concentrate (LIC).⁹² FDA noted that “change in LIC using liver biopsy has generally been considered to be the standard measure of efficacy in response to iron chelation therapy.”⁹³

Overall, FDA first rejected the original NDA on grounds that the primary endpoint of the key pivotal study, MRI T2* changes, was not sufficiently correlated with any clinical outcome to warrant being the basis for even an accelerated approval, notwithstanding the disease being severe, rare, and without adequate therapy. However, FDA approved a second resubmission that was based on an analysis of a commonly used measure in clinical practice of patients with transfusion-related iron overload, serum ferritin, which itself was supported internally by a positive finding in the same population on liver iron concentration which is the “standard measure of efficacy in response iron chelator therapy.”⁹⁴

FDA’s actions on Ferriprox illustrate the fatal flaws in a clinical program attempting to rely upon a surrogate (MRI T2*), the factors to be considered and the clinical

⁸⁶ *Id.* at 5.

⁸⁷ *Id.* at 2.

⁸⁸ Ferriprox (deferiprone), Statistical Review, NDA 021825, 7 (Nov. 10, 2009), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021825Orig1s000StatR.pdf.

⁸⁹ Note that FDA states that this study could provide both evidence of the effect of the drug on an unvalidated surrogate and at the same time, in the same study, evidence of the effect of the drug on clinical outcome, thereby “validating” that surrogate.

⁹⁰ Ferriprox (deferiprone), Statistical Review, NDA 021825, 3 (Nov. 20, 2009), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021825Orig1s000StatR.pdf.

⁹¹ Ferriprox (deferiprone), Medical Review, NDA 021825, 10 (Sept. 27, 2011), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021825Orig1s000MedR.pdf.

⁹² *Id.*

⁹³ Ferriprox (deferiprone), Clinical Review, NDA 021825, 15 (Oct. 19, 2009), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021825Orig1s000MedR.pdf.

⁹⁴ *Id.*

evidence that were found by FDA to be of sufficient merit to allow FDA, as a matter of its judgment, to conclude that serum ferritin is reasonably likely to predict clinical benefit, even without any clinical trial results on any cardiac outcomes, such as heart failure or mortality, and notwithstanding an FDA acknowledgement that serum ferritin is a non-specific measure. However, FDA's Subpart H approval here was based clinically on the corroboration of the serum ferritin results by the liver iron concentrate results and bolstered by the known mechanistic action of the drug (i.e., that by its mechanism of action the drug leads to iron excretion in the urine).

Overall, the clinical evidence of the surrogate was scored a 6 out of a possible 7 due to the strength of evidence on serum ferritin which itself was buttressed by the clinical findings on LIC. However, since there was no clinical evidence on any ultimate clinical outcome, the score for clinical evidence of outcome benefit is 0.

4. MAKENA (hydroxyprogesterone caproate)

FDA's February 3, 2011 approval of Makena to reduce the risk of preterm birth (PTB) was based on a surrogate of reducing preterm birth as defined as those births occurring at less than 37 weeks of gestation. "Preterm birth <37 weeks gestation . . . was a surrogate⁹⁵ for pregnancy outcome (neonatal/infant morbidity and mortality)."⁹⁶

Part 1. Regulatory Factors Weighing into FDA Determination

a. Severity of the Condition

The risks of miscarriage, stillbirths, and neonatal mortality are associated with delivery prior to full-term gestation, as well as neonatal morbidities and adverse maternal outcomes.

b. Rarity of the Condition

Makena was designated as an orphan drug on January 25, 2007.

c. Lack of Available Therapy

"Currently there is no drug product approved in the United States to reduce the risk of preterm birth; however, [the active ingredient in Makena] is compounded by pharmacists and is used widely for this indication in women at high risk."⁹⁷ In 1956, FDA approved an NDA for Delalutin, which had the same active ingredient as Makena, for treating pregnant women for "habitual and recurrent abortion, threatened abortion."⁹⁸ In 2000, FDA withdrew the approval of Delalutin at the request of the NDA Sponsor because it no longer marketed Delalutin. In a June 25, 2010 Federal Register notice, FDA announced its determination that Delalutin was not withdrawn from marketing for safety or efficacy reasons.

⁹⁵ While FDA Medical and Statistical Reviews refer to PTB <37 weeks as a "surrogate," preterm birth is a clinical event and, therefore, in the terminology of the Expedited Programs Guidance, PTB <37 weeks is an "intermediate clinical endpoint."

⁹⁶ Makena (hydroxyprogesterone caproate), Medical Review, NDA 021945, 14 (Feb. 3, 2011), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945Orig1s000MedR.pdf.

⁹⁷ *Id.* at 11.

⁹⁸ *Id.* at 12.

d. Use of External Expertise

With Makena, FDA relied upon two forms of external expertise and FDA reached its “informed judgment” that the surrogate endpoint of preterm birth less than 37 weeks was reasonably likely to predict clinical benefit, that is, pregnancy outcome or neonatal infant and maternal morbidity and mortality. These two forms of external advice are summarized in the Medical Review: (1) 2006 Advisory Committee; and (2) subsequent scientific papers published in the literature.

i. The surrogate endpoints of reductions of [preterm birth] at <35 and <32 weeks were thought by the Advisory Committee to predict a reduction in neonatal mortality and morbidity. At the time of the Advisory Committee meeting in 2006, the endpoint PTB at <37 weeks was not believed to be an adequate surrogate for neonatal outcome.^{99, 100}

ii. The Applicant submitted a single phase 3 clinical trial which demonstrated a statistically strong ($p < .001$) reduction in the incidence of preterm births prior to 37 weeks gestation, the protocol pre-specified primary endpoint. There is recent evidence that “late preterm births” (births between 34^{0/7} and 36^{6/7}), which comprise 71.3% of all preterm births, are increasing, and suffer greater neonatal and childhood morbidity and mortality than previously thought [5 papers are cited that were published between the time of the 2006 Advisory Committee and the Medical Review]. These data indicate that “preterm birth prior to 37 weeks” is a surrogate endpoint that is reasonably likely to predict clinical benefit.¹⁰¹

Part 2. Understanding of the Disease Process

Here the disease process is complex and has multiple pathophysiologic pathways, and therefore, this militates against reliance upon any surrogate. The biological means by which the gestational process progresses to premature delivery is complex and multifaceted. Therefore, the surrogate endpoint of PTB <37 weeks is likely more analogous to the prostate-specific antigen (“PSA”) example than the enzyme replacement example in the Expedited Programs Guidance¹⁰² in that PTB <37 weeks is not on the pathophysiological causal pathway and is not the biologic mechanism that causes the neonatal mortality and morbidity, even though, like PSA, it is correlated with increased risk.

⁹⁹ *Id.* at 6.

¹⁰⁰ “The Committee stated that a reduction of preterm birth <37 weeks was not an adequate surrogate (Yes: 5; No: 16) but that reductions in preterm birth <35 weeks (Yes: 13; No: 8) and <32 weeks (Yes: 20; No: 1) were adequate surrogates.” Makena (hydroxyprogesterone caproate), Medical Review, NDA 021945, 7 (Jan. 23, 2009), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945Orig1s000MedR.pdf.

¹⁰¹ Makena (hydroxyprogesterone caproate), Medical Review, NDA 021945, 5 (Feb. 3, 2011), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945Orig1s000MedR.pdf.

¹⁰² See Expedited Programs Guidance, *supra* note 3, at 19-20.

Part 3. Understanding of the Relationship Between PTB and Pregnancy Outcomes

a. Epidemiological Evidence

The epidemiological evidence is strong with Makena. The 2006 Advisory Committee assessed the epidemiological evidence supporting the relationship between PTB and pregnancy outcomes and found that this evidence was strong enough to support the endpoints of PTB <32 weeks and PTB <35 weeks as surrogate endpoints but not PTB <37 weeks. However, additional evidence published subsequent to the 2006 Advisory Committee permitted the Medical Officer, Dr. Barbara Wesley, to conclude that PTB <37 weeks was also a reliable, consistent and acceptable surrogate endpoint.^{103, 104}

b. Effect of Drugs in the Same or Closely Related Pharmacologic Class to Affect Pregnancy Outcomes

Since there are no drugs in any pharmacologic class approved for reducing the risk of PTB, there are no analogous therapies here on which to draw support directly for reducing the risk of PTB. However, other progesterones including the active ingredient in Makena have been approved for aiding in assisted reproductive technologies and other conditions supporting the maintenance of pregnancy.

Part 4. Clinical Evidence of Makena's Effect on PTB <37 Weeks and on Pregnancy Outcomes

a. PTB <37 Weeks

The surrogate of PTB <37 weeks was highly statistically significant ($p < 0.001$).

b. Pregnancy Outcomes

The proportion of babies with at least one event on the [secondary] composite index of neonatal morbidity and mortality was lower in the [Makena] group (11.9%, 35/295 infants) than in the vehicle group (17.2%, 26/151 infants) but the between-group differences was not statistically significant (nominal p -value of 0.1194).¹⁰⁵ Approximately 6.5% of the

¹⁰³Makena (hydroxyprogesterone caproate), Medical Review, NDA 021945, 5 (Feb. 3, 2011), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945Orig1s000MedR.pdf.

¹⁰⁴It is also likely that the Advisory Committee was opining on PTB <32 weeks, PTB <35 weeks and PTB <37 weeks as *validated* surrogates which would have qualified Makena for traditional approval, not Subpart H approval. Outside of AIDS and cancer, FDA has not often asked Advisory Committees to opine on whether clinical evidence on a particular endpoint would qualify a therapy for Subpart H approval. For example, note that the August 5, 2013 Cardiorenal Advisory Committee, addressing the approvability of tolvaptan, a vasopressin V2 receptor antagonist, was *not* asked whether total kidney volume would qualify as an unvalidated surrogate that may support a Subpart H approval if the Advisory Committee found that total kidney volume is “reasonably likely to predict clinical benefit,” which, in this case, clinical benefit would likely be end-stage renal disease and/or clinically meaningful outcomes such as significant worsening of renal function or kidney pain. However, there are exceptions outside of AIDS and cancer. For instance, the Oncology Drugs Advisory Committee (ODAC) was asked whether familial adenomatous polyposis (“FAP”) was an adequate “unvalidated” surrogate, that is, to qualify Celebrex (Precedent #12) for Subpart H approval. But even this case was before ODAC, and while FAP is not cancer, the ultimate clinical benefit was prevention of colon cancer, so even this “exception” is not fully outside of AIDS and cancer.

¹⁰⁵Makena (hydroxyprogesterone caproate), Medical Review, NDA 021945, 6 (Feb. 3, 2011),

women in each treatment group experienced a fetal or neonatal deaths The results . . . show that despite the treatment groups having about the same rate of fetal and neonatal deaths, the losses occur earlier among [Makena] women.¹⁰⁶

This impact on fetal or neonatal deaths was stated another way by the Medical Reviewer: “There was a trend toward an increased risk of miscarriage and stillbirths in the [Makena] treatment arm and a trend toward a decrease in neonatal death, with no overall net survival benefit.”¹⁰⁷

Overall, the secondary endpoint of a composite measure of neonatal morbidity/mortality leaned in favor of the Makena group while the separate analysis of neonatal mortality showed essentially no numerical difference and had a nominal p-value of 0.6887.¹⁰⁸ The clinical evidence for the ultimate clinical benefits in the single pivotal trial was not strong.

5. SOLIRIS (*eculizumab*)

FDA’s September 23, 2011 approval of Soliris for the treatment of patients with atypical hemolytic uremic syndrome (aHUS) to inhibit complement-mediated thrombotic microangiopathy was based on thrombotic microangiopathy (TMA) endpoints.

Endpoints related to TMA included the following:

- Platelet count change from baseline;
- Hematologic normalization (maintenance of normal platelet counts and LDH levels for at least four weeks);
- Complete TMA response (hematologic normalization plus at least a 25% reduction in serum creatinine for a minimum of four weeks);
- TMA-event free status (absence for at least 12 weeks of a decrease in platelet count of >25% from baseline, plasma exchange or plasma infusion, and new dialysis requirement); and
- Daily TMA intervention rate (defined as the number of plasma exchange or plasma infusion interventions and the number of new dialyses required per patient per day).¹⁰⁹

Part 1. Regulatory Factors Weighing into FDA Determination

a. Severity of the Condition

The risks of aHUS are associated with TMA, as well as the formation of blood clots in small blood vessels throughout the body, which can lead to stroke, heart attack, kidney failure, and death.

http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945Orig1s000MedR.pdf.

¹⁰⁶Makena (hydroxyprogesterone caproate), Statistical Review, NDA 021945, 20 (Oct. 19, 2006), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945Orig1s000StatR.pdf.

¹⁰⁷Makena (hydroxyprogesterone caproate), Medical Review, NDA 021945, 7 (Feb. 3, 2011), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945Orig1s000MedR.pdf.

¹⁰⁸*Id.*

¹⁰⁹Soliris (eculizumab) Label, BLA 125166, 18 (Sept. 2011), http://www.accessdata.fda.gov/drugsatfda_docs/label/2011/125166s172lbl.pdf.

b. Rarity of the Condition

Soliris was designated as an orphan drug on April 29, 2009.

c. Lack of Available Therapy

No available therapies.

d. Use of External Expertise

No advisory committee.

Part 2. Understanding of the Disease Process

Here the disease process is complex and not well understood, and therefore, this militates against reliance upon any surrogate.

Part 3. Understanding of the Relationship Between TMA Endpoints and Renal Failure, Death, and Stroke

The understanding of the relationship between TMA endpoints and renal failure, death, and stroke is fairly good.

Part 4. Clinical Evidence of Soliris's Effect on TMA Endpoints and on Renal Failure, Death, and Stroke

While data appeared positive in the three registration trials on TMA endpoints, little evidence is available on the clinical evidence of renal failure, death, and stroke from FDA approval documents.

6. PROMACTA (eltrombopag)

FDA approved Promacta on November 20, 2008 on “short term platelet count response as a surrogate marker for longer platelet count responses (platelet counts are recognized as acceptable measures of clinical benefit for patients with ITP [idiopathic thrombocytopenic purpura]).”¹¹⁰ The two clinical trials of Promacta administered drugs over 6 weeks or less (this is the meaning of “short term” in the Reviewer’s statement above). Had the Promacta trials studied and established the drug’s effect on platelet counts out to 6 months, this approval would have been a traditional approval and not one under Subpart H.

Part 1. Regulatory Factors Weighing into FDA Determination

a. Severity of the Condition

Chronic ITP is a serious medical condition.¹¹¹ The frequency of death from hemorrhage in patients with platelet counts below 30,000/mcl is estimated to be between 1.6 and 3.9% per patient year.¹¹²

¹¹⁰Promacta (eltrombopag), Medical Review, NDA 022291, 3 (Oct. 29, 2008), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2008/022291s000_MedR_P1.pdf.

¹¹¹*Id.*

¹¹²Promacta (eltrombopag), Medical Review, NDA 022291, 17 (Sept. 12, 2008), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2008/022291s000_MedR_P1.pdf.

b. Rarity of the Condition

FDA designated Promacta as an orphan drug on May 5, 2008.

c. Lack of Available Therapy

[Promacta] approval would provide a meaningful therapeutic benefit to patients over existing treatments because of its minimal risk for immunogenicity (based upon [its] small molecule characteristics). The labeling for romiplostim, the only currently marketed TPO receptor agonist, includes information regarding the risks for immunogenicity. These risks are not applicable to [Promacta].¹¹³

d. Use of External Expertise

In the medical and statistical reviews, the authors found no evidence of any reliance on special government employees (SGEs), an Advisory Committee for Promacta, or specific published literature.

Part 2. Understanding of the Disease Process

The clinical hallmark of the disease is an increased tendency to bleed.¹¹⁴ [Furthermore, the relationship of platelet count to bleeding is well-established:] Patients with platelet counts between 30,000/mcl and 10,000/mcl are generally considered treatment candidates due to slightly increased risk of spontaneous bleeding or increased risk of bleeding due to potential trauma.¹¹⁵

Part 3. Understanding of the Relationship Between the Drug's Effect on Short-Term Platelet Counts and Increased Risk of Bleeding

There was no epidemiological evidence cited in FDA review documents to support the surrogate – which is “short term” (that is, six weeks) increase in platelet count – as reasonably likely to predict long-term, chronic increase in platelet count – which is generally established in six month trials or generally on increased risk of bleeding. While there was no evidence to support the use of this surrogate, there was a therapy approved from the same pharmacologic class but based on an endpoint of six-month duration. Earlier in 2008 (the year FDA approved Promacta), FDA had approved romiplostim, a biological product that is a member of the same pharmacologic class – thrompoietin (TPO) receptor agonists – and this approval for the same indication (that is, to treat ITP) was a traditional approval based on two clinical trials, each of six-months duration.

¹¹³Promacta (eltrombopag), Medical Review, NDA 022291, 3 (Oct. 29, 2008), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2008/022291s000_MedR_P1.pdf.

¹¹⁴Promacta (eltrombopag), Medical Review, NDA 022291, 17 (Oct. 29, 2008), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2008/022291s000_MedR_P1.pdf.

¹¹⁵*Id.*

*Part 4. Clinical Evidence of Promacta's Shorter-Term
(Surrogate) Effect and Long-Term Effect on Platelets
and/or Bleeding*

Both of Promacta's pivotal studies showed a robust short-term (surrogate) effect on platelets ($p < 0.001$).¹¹⁶

As for clinical evidence that FDA had at the time of the approval that Promacta's short-term (six weeks) impact on platelet counts would predict either clinical benefit of long-term impact on platelet counts or on bleeding, there was mixed evidence.

As supportive evidence that the platelets produced by Promacta behaved in a physiologically "normal" way, the Sponsor had conducted "an exploratory clinical study that demonstrated [that Promacta] prompted platelet count increases in healthy subjects. These drug-stimulated platelets had in vitro platelet function characteristics generally typical of platelets. Hence, this study supported the generally accepted use of platelet counts as an 'accepted' measure of clinical benefit for clinical studies of TPO receptor agonists among patients with chronic ITP."¹¹⁷

As Promacta was only administered for six weeks (or less) in the two pivotal trials, there is no clinical evidence as to the impact long-term on platelet counts if Promacta was administered chronically (for which a trial of six-months duration would have been relied upon). Furthermore, of some concern, "discontinuation of [Promacta] at the end of the study resulted in an unacceptable amount of serious hemorrhage."¹¹⁸ Also, the statistical reviewer observed that within two weeks after the subjects on drug were off treatment, there was a return to placebo levels of platelet counts.¹¹⁹

As for bleeding events, there was a numerical lean in favor of Promacta, but in neither trial was this statistically significant with p-values of 0.121 and 0.088 for the between-group difference on bleeding events in the two pivotal trials, respectively.¹²⁰

7. EXJADE (deferasirox)

The FDA approval of Exjade for treating chronic iron overload due to blood transfusions on November 2, 2005 was based on a surrogate endpoint of improvement in liver iron concentration (LIC).

Part 1. Regulatory Factors Weighing into FDA Determination

a. Severity of the Condition

"Chronic iron overload due to requisite blood transfusion is a serious and life-threatening condition."¹²¹

¹¹⁶CTR. FOR DRUG EVALUATION & RESEARCH, NDA 022291, STATISTICAL REVIEW (Sept. 18, 2008), at 19, 27, http://www.accessdata.fda.gov/drugsatfda_docs/nda/2008/022291s000_StatR.pdf.

¹¹⁷Promacta (eltrombopag), Medical Review, NDA 022291, 2-3 (Oct. 29, 2008), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2008/022291s000_MedR_P1.pdf.

¹¹⁸*Id.* at 3.

¹¹⁹*Id.* at 10.

¹²⁰*Id.* at 8-9.

¹²¹Exjade (deferasirox), Medical Review, NDA 021882, 2 (Nov. 2, 2005), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2005/021882_s000_Exjade_medr.pdf.

b. Rarity of the Condition

Exjade was granted orphan drug designation on November 21, 2002.

c. Lack of Available Therapy

At the time FDA was reviewing the Exjade NDA, the Medical Team Leader, Dr. Dwaine Rieves, stated: “Deferoxamine, the only available therapy for this condition, presents unique compliance and infectious risks due to the need for prolonged parenteral administration of the drug. [Exjade] is an orally administered drug that provides a meaningful therapeutic benefit over the existing therapy.”¹²²

d. Use of External Expertise

FDA sought the advice of the Blood Products Advisory Committee (BPAC) and at its September 29, 2005 meeting, the BPAC found that “the applicant [had] provided substantial evidence of the effectiveness of [Exjade] in the reduction of liver iron concentration, an outcome indicative of a clinical benefit The Sponsor’s major clinical evidence of [Exjade] effectiveness . . . is based upon alterations in liver iron content, an endpoint the BPAC discussants regarded as a measure of clinical benefit. In this context, the endpoint is not regarded as a surrogate endpoint rather as an endpoint other than survival or irreversible morbidity¹²³, as cited in the Subpart H regulations.”¹²⁴

Part 2. Understanding of the Disease Process

See Part 2 under Ferriprox.

Part 3. Understanding of the Relationship Between LIC and Cardiac Outcomes, Including Mortality

Although accepted by the Division as a clinically meaningful endpoint, the primary endpoint [of LIC] is technically a surrogate endpoint since it does not necessarily address clinically significant morbidity or mortality. The main mortality on β -thalassemia is due to cardiac dysfunction whose etiology in β -thalassemia is probably multifactorial. Nonetheless, most of the literature in β -thalassemia has used LIC as a marker for morbidity for other organ involvement and as a surrogate for mortality. There is some information, however, that LIC does not completely correlate to the extent of cardiac hemosiderosis, the primary cause of mortality. Obviously,

¹²²*Id.*

¹²³An “intermediate clinical endpoint” (rather than a surrogate) is the term used in the Draft Guidelines for this kind of endpoint; however, during the later FDA approval of Ferriprox, FDA Reviewers refer to both serum ferritin and LIC as “surrogates,” and in an earlier medical review of Exjade, FDA refers to LIC in this pivotal trial as a “surrogate” (see CTR. FOR DRUG EVALUATION & RESEARCH, NDA 021882, MEDICAL REVIEW(S) (Oct. 26, 2005), at 38, http://www.accessdata.fda.gov/drugsatfda_docs/nda/2005/021882_s000_Exjade_medr.pdf). Therefore, this analysis will refer to LIC as a surrogate and not as an intermediate clinical benefit.

¹²⁴Exjade (deferasirox), Medical Review, NDA 021882, 2 (Nov. 2, 2005), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2005/021882_s000_Exjade_medr.pdf.

repetitive biopsy of the myocardium to measure iron concentrations in the heart is not acceptable.¹²⁵

As for understanding the relationship between drugs in the same pharmacologic class as LIC, the single pivotal trial for Exjade was a noninferiority study design which used as its active comparator, deferoxamine, and therefore, FDA had evidence from a within-study comparison of the only other member of the same or closely related class on the surrogate endpoint of LIC.

Part 4. Clinical Evidence of Exjade's Effect on LIC and/or Cardiac Outcomes Including Mortality

FDA, in its review of this NDA, noted that LIC as “[t]he primary endpoint is acceptable and was agreed to by the Division in the Special Protocol Assessment. It should be remembered, however, that the LIC is a surrogate marker and that the effects of Exjade on morbidity/mortality, which are the truly important clinical endpoints, are not likely to be demonstrated in this short trial.”¹²⁶

Rather than bolster LIC results by seeing trends on irreversible morbidity and mortality in this “short” trial, FDA looked to find support from other critical surrogate markers such as serum ferritin.¹²⁷

As for LIC, the protocol had specified that “non-inferiority of [Exjade] to [deferoxamine] was to be established if the two sided 95% confidence interval of the difference in success rate between the two study groups was above -15%. The basis for the choice of this non-inferiority margin was unclear in the submission. Notably, FDA had questioned the meaningfulness of this margin during the study’s protocol review.”^{128, 129}

The primary efficacy result was a point estimate difference of -13.5%, with a lower 95% confidence interval of -21.6% (or, in other words, the margin defining success of the trial was *not* met). About this, FDA concluded: “Given that the original basis for the non-inferiority margin was poorly substantiated, little clinical meaningfulness could be assigned to failure to achieve the primary endpoint. The primary endpoint data did establish that both [Exjade and deferoxamine] lowered LIC over a 12 month period of time, a time period during which subjects would have been expected to have increases in LIC due to continuing blood transfusions. This observation provides evidence of a treatment effect for [Exjade].”¹³⁰

With respect to serum ferritin, FDA concluded that “[s]erum ferritin values declined in a dose-related manner for subjects receiving [Exjade], a pattern similar to that for subjects receiving [deferoxamine].”¹³¹

¹²⁵CTR. FOR DRUG EVALUATION & RESEARCH, NDA 021882, MEDICAL REVIEW(S) (Oct. 26, 2005).

¹²⁶*Id.* at 31.

¹²⁷The authors must inform the reader that this trial was a year-long trial, and, therefore, by many would not be considered “short;” however, even a year long study is too “short” to see effects on mortality and irreversible morbidity.

¹²⁸Exjade (deferasirox), Medical Review, NDA 021882, 4 (Nov. 2, 2005), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2005/021882_s000_Exjade_medr.pdf.

¹²⁹Query, though, how FDA nevertheless had accepted the design of this pivotal study under an SPA.

¹³⁰Exjade (deferasirox), Medical Review, NDA 021882, 5 (Nov. 2, 2005), http://www.accessdata.fda.gov/drugsatfda_docs/nda/2005/021882_s000_Exjade_medr.pdf.

¹³¹*Id.*

8. *TYSABRI (natalizumab)*

FDA approved Tysabri on November 23, 2004 for treating relapsing-remitting multiple sclerosis (RRMS), relying upon the reduction in MS relapse rates at one year as the surrogate endpoint. Applying the terms of the Expedited Programs Guidance, this would be an intermediate clinical endpoint that would be reasonably likely to predict the benefit at two years. All previous MS therapies were approved on the basis of two-year relapse rate reduction and “the clinical meaningfulness of a decrease in the relapse rate through only one year is uncertain.”¹³²

Part 1. Regulatory Factors Weighing into FDA Determination

a. Severity of the Condition

Relapsing-remitting multiple sclerosis is a serious, life-threatening condition.

b. Rarity of the Condition

While Tysabri was not designated an orphan drug for RRMS, the statutory threshold for qualifying as an orphan drug was, in part, set in the 1984 amendment to the Orphan Drug Act specifically to include *all* of multiple sclerosis as an orphan disease, not just the subset of RRMS. This was because, in considering how to amend the original 1983 Orphan Drug Act to make it less difficult to garner orphan drug designation, key Senators caucused with the National Organization for Rare Diseases (NORD) and mutually determined that the maximum number of Americans with a condition which would still qualify as an “orphan” would be 200,000. This number was chosen, specifically, to make sure that MS would be an “orphan” disease, and in 1984 there were just under 200,000 Americans diagnosed with MS. However, soon after FDA approved the first therapy for multiple sclerosis (Betaseron in August 1993, which was also the first non-AIDS Subpart H approval), the number of Americans diagnosed with multiple sclerosis dramatically increased. So, while Tysabri was never designated as an orphan drug for RRMS, the authors, fully cognizant of the intent of the 1984 orphan drug amendment, view Tysabri as, nevertheless, falling within the “penumbra” of orphan drug status and score Tysabri a “2” on rarity.

c. Lack of Available Therapy

Accelerated approval requires that the new drug provide evidence of the potential to address an unmet medical need. Many MS patients continue to have exacerbations while taking one of the available first-line MS therapies. None of the currently available therapies have proven efficacy when used as add-on therapy. [One of the two pivotal Tysabri studies] provides evidence that [Tysabri] is effective as add-on therapy for subjects who continue to have relapses while on a first-line therapy (Avonex). Therefore, [Tysabri] has the potential to address an unmet medical need.¹³³

¹³²CTR. FOR DRUG EVALUATION & RESEARCH, NDA 125104, MEDICAL REVIEW (Nov. 23, 2004), at 6, http://www.accessdata.fda.gov/drugsatfda_docs/nda/2004/125104s000_Natalizumab_Medr_P1.pdf.

¹³³*Id.*

d. Use of External Expertise

FDA did not rely on an advisory committee during its initial review of Tysabri. However, Tysabri was withdrawn from the market by the manufacturer in February 2005 after three patients developed progressive multifocal leukoencephalopathy (PML). Subsequently, FDA convened an Advisory Committee to consider the reintroduction of Tysabri in March 2006. Furthermore, FDA had convened and considered the input from several earlier advisory committees on other multiple sclerosis therapies.

Part 2. Understanding of the Disease Process

“Multiple sclerosis is a chronic, inflammatory, possibly autoimmune, demyelinating disease of the central nervous system.”¹³⁴ Note that FDA’s review states that multiple sclerosis may be “possibly autoimmune.” Given that Tysabri’s mechanism of action is as an immunomodulator, having a more definitive view of the causative role of autoimmunity in the pathophysiology of this disease would have been more compelling.

Part 3. Understanding of the Relationship Between the One-Year Relapse Rate and Two-Year Relapse Rate

The effect of [Tysabri] on relapse rate in [the pivotal study on Tysabri’s use as first-line therapy] was approximately twice the effect observed with current first-line drugs for this indication. Such comparisons of different agents across studies are problematic However, the magnitude of [Tysabri’s] effect is sufficient that the effect at one year is reasonably likely to predict a clinical benefit at two years.¹³⁵

As for understanding the relationship of drugs in the same or closely-related pharmacologic class on rate and extent of exacerbations at one year of treatment as predictive of their two year effectiveness, at the time of Tysabri’s approval, there were four other approved immunomodulators approved for treatment of MS: Betaseron, Avonex, Rebif, and Copaxone. While each of these was approved on the basis of two-year studies of impact on reducing rate and extent of MS exacerbations, their impacts after one year of therapy, while generally more modest than at the end of two years, were predictive of their two year results.

Part 4. Clinical Evidence on One-Year and Two-Year Relapse Rates

For other MS products, FDA has required two-year data. . . . A salutary effect on relapse rate at one year is not a validated surrogate for benefit at two years. However, the apparent treatment effect of [Tysabri] with respect to relapse rate at one year is *unprecedented* in the MS field, and its magnitude is reasonably likely to predict clinically meaningful effectiveness at two years. If, in fact, the benefit on clinical relapses is shown to be durable through two years, the product may be substantially more efficacious than currently approved MS therapies. . . . It is possible,

¹³⁴*Id.* at 11.

¹³⁵*Id.* at 102.

however, that the magnitude of [Tysabri's] effect on relapse rate, when assessed through one year, may substantially overestimate [Tysabri's] benefit on relapse rate through two years. . . . In particular, the treatment effect appears to wane with the development of [anti-Tysabri] antibodies, which may increase with time.¹³⁶

9. LUVERIS (*lutropin alfa*)

On October 8, 2004, FDA approved Luveris for stimulating follicular development in infertile hypogonadotropic hypogonadal women with profound LH deficiency (LH < 1.2). "The Division Director further concluded that in this orphan population of women with severe LH deficiency (LH \leq 1.2), the surrogate endpoint of follicular development (as defined by the Sponsor) was reasonably likely to predict clinical benefit [with respect to pregnancy]."¹³⁷

Part 1. Regulatory Factors Weighing into FDA Determination

a. Severity of the Condition

The inability to ovulate due to profound luteinizing hormone (LH) deficiency includes, among other serious consequences, the inability to become pregnant. "The Director believes that infertility in the context of hypogonadotropic hypogonadism and profound LH deficiency is a serious condition with very limited options for pregnancy."¹³⁸

b. Rarity of the Condition

Luveris was granted orphan drug designation by FDA on October 7, 1994.

c. Lack of Available Therapy

Luveris would be the only LH-alone product . . . on the U.S. market. There are no approved drug products that have the indication of treatment of infertility in women with hypogonadotropic hypogonadism.¹³⁹

d. Use of External Expertise

The Reproductive Health Advisory Committee considered Luveris on September 30, 2003. "After hearing presentations from experts in Reproductive Endocrinology on the subject of female hypogonadotropic hypogonadism . . . the Committee voted 15 to 0 that the Sponsor's data did *not* demonstrate efficacy for Luveris in ovulation induction when the primary endpoint was ovulation rate. The Committee voted 8 to 7 that the Sponsor's data demonstrated efficacy for Luveris in ovulation induction when the primary endpoint was follicular development. Finally, the Committee voted 11 to

¹³⁶*Id.* at 53 (emphasis added).

¹³⁷CTR. FOR DRUG EVALUATION & RESEARCH, NDA 021322, MEDICAL REVIEW (Oct. 7, 2004), at 2, http://www.accessdata.fda.gov/drugsatfda_docs/nda/2004/021322s000_Luveris_medr.pdf.

¹³⁸CTR. FOR DRUG EVALUATION & RESEARCH, NDA 021322, DIVISION DIRECTOR'S MEMORANDUM, (Oct. 8, 2004), at 7, http://www.accessdata.fda.gov/drugsatfda_docs/nda/2004/021322s000_Luveris_admincorres.pdf.

¹³⁹CTR. FOR DRUG EVALUATION & RESEARCH, NDA 021322, MEDICAL REVIEW (Sept. 28, 2004), at 17, http://www.accessdata.fda.gov/drugsatfda_docs/nda/2004/021322s000_Luveris_medr.pdf.

3 . . . that the Sponsor's data demonstrated efficacy for Luveris for follicular development when the primary endpoint was follicular development."¹⁴⁰

Part 2. Understanding of the Disease Process

FDA's medical review suggests that the disease process is complex and multifactorial: "the role of LH in hypogonadal female infertility patients is clouded by the spectrum of clinical disorders that cause hypogonadotropic hypogonadism with the differing patterns of gonadotropin secretion may further confound clinical outcome results."¹⁴¹

Part 3. Understanding of the Relationship Between Follicular Development and Fertility

"The Division believed that although both follicular development and ovulation are surrogates for pregnancy (the clinically meaningful outcome), ovulation is more temporally proximate to pregnancy and therefore more appropriate as a surrogate."¹⁴² Nevertheless, follicular development is on the causal pathway, as is ovulation. However, there was no epidemiological evidence cited in FDA review documents linking follicular development to pregnancy.

As for understanding the relationship of drugs in the same or closely-related pharmacologic class on follicular development: "Recognition of the therapeutic potential of gonadotropins began in the 1950's with the extraction and purification of human menopausal gonadotropins (both follicle stimulating hormone and luteinizing hormone) from both human pituitaries and urine sources. Successful clinical pregnancies resulting from the use of these human-derived gonadotropins were first reported in the 1960s. In the 1990s cells that are capable of producing biologically-active LH in culture produced LH. This recombinant derived LH is from in vitro cultured cells."¹⁴³

Part 4. Clinical Evidence of Luveris on Follicular Development and Fertility

The primary efficacy parameter for both Studies 6905 and 6253 was "follicular development" as defined by three co-primary endpoints (follicle size as measured on ultrasound, pre-ovulatory serum estradiol levels and mid-luteal progesterone levels). The Sponsor's analysis demonstrated that in Study 6253, 75 IU of Luveris was numerically better than 25 IU of Luveris or placebo for follicular development in women with LH <1.2 IU/L.¹⁴⁴ The Division's analysis of Study 6905 demonstrated . . . [the] placebo was as efficacious as 75 IU of Luveris.

¹⁴⁰CTR. FOR DRUG EVALUATION & RESEARCH, NDA 125104, MEDICAL REVIEW (Nov. 23, 2004), at 6, http://www.accessdata.fda.gov/drugsatfda_docs/nda/2004/125104s000_Natalizumab_Medr_P1.pdf.

¹⁴¹CTR. FOR DRUG EVALUATION & RESEARCH, NDA 021322, MEDICAL REVIEW (Sept. 28, 2004), at 19, http://www.accessdata.fda.gov/drugsatfda_docs/nda/2004/021322s000_Luveris_medr.pdf.

¹⁴²Luveris Division Director's Memo, *supra* note 138, at 5.

¹⁴³CTR. FOR DRUG EVALUATION & RESEARCH, NDA 021322, MEDICAL REVIEW (Sept. 28, 2004), at 17, http://www.accessdata.fda.gov/drugsatfda_docs/nda/2004/021322s000_Luveris_medr.pdf.

¹⁴⁴Luveris Division Director's Memo, *supra* note 138, at 3.

Therefore, in the opinion of the Division, Luveris was not demonstrated to be effective.¹⁴⁵

Therefore, the Sponsor planned and conducted a third study, Study 21008, with follicular development as the Sponsor's prespecified primary endpoint, despite the Division's recommendations that ovulation rate be the primary endpoint. The Sponsor's "evaluable patient analysis of Study 21008 demonstrated that 67% of patients receiving 75 IU of Luveris achieved follicular development compared to 20% of patients receiving placebo."¹⁴⁶ "The Director [Dr. Shames] further concluded that the results from Studies 21008 and 6253 provide substantial evidence that Luveris 75 IU, when administered concomitantly with FSH, induces follicular development in this population of infertile women. These studies, however, do not demonstrate a positive effect on clinical pregnancy, etc. Study 21415 evaluated titrable FSH dosing with the dose of Luveris fixed at 75 IU and demonstrated a 36% clinical pregnancy rate after one cycle. While reassuring, this finding is not definitive because there was no placebo comparator group in Study 21415, and the finding has not been replicated in a second trial."¹⁴⁷ Study 21415 also reported follicular development rates of 63% "in all cycles combined."¹⁴⁸ Therefore, in Study 21415, there was within-study clinical evidence both on follicular development, the surrogate, as well as on pregnancy, the ultimate clinical outcome.

10. FABRAZYME (agalsidase beta)

FDA approved Fabrazyme on April 24, 2003 to treat Fabry's disease. This approval was based on a surrogate endpoint of near-elimination of all accumulation of enzyme in renal capillary endothelium, one type of vascular endothelium.

Part 1. Regulatory Factors Weighing into FDA Determination

a. Severity of Condition

[W]ith age, the principal manifestations of concern in Fabry's disease are in the kidney, heart, and brain. Renal disease is manifested by proteinuria, hypertension, and progressive azotemia; the principal cause of death in Fabry's disease in the past was renal failure. . . . The median age of death for hemizygous males is 50 years.¹⁴⁹

b. Rarity of the Condition

FDA designated Fabrazyme as an orphan drug on January 19, 1988.

c. Lack of Available Therapy

"There is no specific treatment for Fabry's disease."¹⁵⁰

¹⁴⁵*Id.*

¹⁴⁶*Id.* at 4.

¹⁴⁷*Id.* at 7.

¹⁴⁸CTR. FOR DRUG EVALUATION & RESEARCH, NDA 021322, MEDICAL REVIEW (Sept. 28, 2004), at 29–30, http://www.accessdata.fda.gov/drugsatfda_docs/nda/2004/021322s000_Luveris_medr.pdf.

¹⁴⁹CTR. FOR DRUG EVALUATION & RESEARCH, NDA 103979, MEDICAL REVIEW (Apr. 21, 2003), at 4, <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/TherapeuticBiologicApplications/ucm091354.pdf>.

¹⁵⁰*Id.*

d. Use of External Expertise

Vessels (capillaries in this case) that are essentially near-normal in appearance that may well lead to an altered development of vascular occlusion, and thus to an alteration in expression of the clinical impairments of the disease. The [January 2003] Advisory Committee has also supported this assessment of the potential impact of near-absence of capillary accumulation, as well as concurring that the evidence submitted by [the Sponsor has] demonstrated this effect on capillary endothelium.¹⁵¹

Part 2. Understanding of the Disease Process

The underlying basis of Fabry disease is well understood; it is an X-linked enzyme deficiency leading to a lipid storage disorder. Lipid storage occurs in a wide variety of cell types, and consequently there are a wide variety of signs and symptoms from different organ systems. . . . However, [there] is widespread belief that a number of the organ injury manifestations are related to vascular injury. It is believed that while this may not be the sole pathologic process, progressive substrate accumulation within vascular walls will ultimately lead to local vessel occlusion, with organ impairment as a consequence.¹⁵²

Part 3. Understanding of the Relationship Between Near-Elimination of Substrate in the Renal Capillary Endothelium and the Outcomes of Fabry's Disease Including Renal Failure and Mortality

Vascular injury does appear to be an important mechanism of promoting the progressive organ impairment, and substrate accumulation within vascular walls is the basis for this. The exact (quantitative) relationship between the amount of substrate accumulation and the degree or rate of vascular ischemia is unknown and not addressed in any information submitted by [the Sponsor]. It is unknown if reducing substrate accumulation by half might slow vascular injury by half, or if there is a threshold effect, wherein some specific amount of accumulation will invariably lead to vascular occlusion and thus no change in the clinical expression of the disease. However, by focusing upon a near-elimination of all accumulation within a specific cell type [the Sponsor's] data appear to overcome these concerns.¹⁵³

Following FDA requests to [the Sponsor], additional data were submitted which demonstrated that while not all cell types show a marked decrease in substrate accumulation (e.g., renal podocytes, with a limited degree of reduction in substrate accumulation) there are a variety of cell

¹⁵¹*Id.* at 52.

¹⁵²CTR. FOR DRUG EVALUATION & RESEARCH, NDA 103979, MEDICAL REVIEW (Apr. 23, 2003), at 3, <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/TherapeuticBiologicApplications/ucm091347.pdf>.

¹⁵³*Id.*

types with moderate and several that show marked reduction in substrate accumulation.¹⁵⁴

As for understanding the relationship of drugs in the same or closely related pharmacologic class on near-elimination of substrate in specific cell types and Fabry's disease, there were no other drugs approved at that time, and there was only one other drug with controlled clinical studies in Fabry's disease: Replagal.

Part 4. Clinical Evidence on Substrate Reduction in Certain Cell Types and Fabry's Disease Outcomes

a. Substrate Reduction

The primary endpoint in the 58-patient, placebo-controlled randomized trial was clearance (i.e., elimination) of kidney intestinal capillary endothelium GL-3 inclusions (or substrate). While none of the 29 placebo subjects achieved a score of "zero" GL-3 inclusions over the 5 month duration of the trial, 20 of the 29 Fabrazyme subjects "cleared" all substrate ($p < 0.001$).¹⁵⁵

b. Clinical Outcomes

The clinical trials failed to show clinical benefit on a wide range of tests of neurologic, renal, and cardiac function. This finding weakens confidence in the clinical importance of the reduction of kidney interstitial capillary endothelial cell GL-3 [enzyme substrate] levels that constituted the primary endpoint of the pivotal trial.¹⁵⁶

In the pivotal study, there was only one secondary endpoint that assessed a clinical outcome, and that was pain. In the five ways in which pain was assessed, the placebo group outperformed the treated group in 4 of the 5 measures of pain.¹⁵⁷ There were tertiary endpoints that assessed clinical outcomes and in eight of these, there were no numerical between-group differences, and in one measure of neuropathy, the placebo group fared somewhat better and in two measures (symptom-free days and episode-free days), the Fabrazyme group fared somewhat better. Of interest, renal function was assessed by Inulin-GFR and by serum cystatin-C, and on both of these measures of renal function, there were essentially no numerical differences between placebo and Fabrazyme groups. Among "other" endpoints, there were ophthalmic assessments, and "the ophthalmological findings, like the tertiary endpoints, did not show a clinical change effected by the product."¹⁵⁸

11. REMODULIN (treprostini)

The May 21, 2002 approval of Remodulin for treating pulmonary hypertension (now referred to as pulmonary arterial hypertension or PAH) was based on an intermediate clinical endpoint of 6-minute walk ("6MW") test, a measure of exercise

¹⁵⁴*Id.* at 1.

¹⁵⁵*Id.* at 30.

¹⁵⁶*Id.* at 74.

¹⁵⁷*Id.* at 35–36.

¹⁵⁸*Id.* at 39–42.

capacity that is a clinical endpoint, but not the ultimate clinical outcome of this serious disease.

Part 1. Regulatory Factors Weighing into FDA Determination

a. Severity of the Condition

PAH is a serious, life-threatening condition.

b. Rarity of the Condition

FDA designated Remodulin for PAH an orphan drug on June 4, 1997.

c. Lack of Available Therapy

The only other therapy approved before Remodulin was Flolan, whose labeling states that “8 of 40 patients receiving standard therapy alone died, whereas none of the 41 patients receiving Flolan died (p=0.003).”¹⁵⁹ This same Medical Review states also that Flolan’s “use is difficult and inconvenient. The infusion of Flolan requires the insertion of an indwelling central catheter with the . . . subsequent risk of catheter infection Any inadvertent interruption of the infusion is potentially life-threatening.”¹⁶⁰

d. Use of External Expertise

The Cardiovascular and Renal Drugs Advisory Committee, on August 9, 2001, voted 6 to 3 in favor of approving Remodulin.

Part 2. Understanding of the Disease Process

The pathophysiology of PAH is well-understood.

Part 3. Understanding of Relationship Between 6MW Results and Clinical Worsening of PAH

Exercise capacity as measured by the 6MW test was judged by FDA as reasonably likely to predict clinical benefit, which was determined to be clinical worsening of PAH symptoms. Confirmation of FDA’s decision to rely upon the 6MW test results as predictive of clinical benefit was later seen in that this same measure, 6MW, was the basis for the approval of several subsequent PAH therapies, especially after this Sponsor’s successful completion of its Phase 4 confirmatory trial established Remodulin’s effect on preventing clinical worsening (p<0.001). The Sponsor’s Phase 4 trial results on clinical worsening demonstrated the positive predictive value of the 6MW test results with Remodulin.

Part 4. Clinical Evidence on 6MW and on Clinical Worsening or Mortality

The primary endpoint of the pivotal trials was “change in [6 minute] walking distance from baseline at the end of week 12 The database was to be considered demonstrating a benefit for [Remodulin] if either both studies were by themselves significant at the p<0.049 or if one study was significant (P<0.049) and the pooled

¹⁵⁹CTR. FOR DRUG EVALUATION & RESEARCH, NDA 021272, MEDICAL REVIEW (Mar. 9, 2001), at 14, http://www.accessdata.fda.gov/drugsatfda_docs/nda/2002/21-272_Remodulin_medr_P2.pdf.

¹⁶⁰*Id.*

studies had a p-value of less than 0.01 Neither of the studies demonstrated a p-value of <0.049 (p=0.06 for both studies), although the pooled studies demonstrated an overall p-value of <0.01 (p=0.006 for the pooled studies).¹⁶¹ In the pivotal [Remodulin] studies, the drug demonstrated no mortality benefit.¹⁶²

12. CELEBREX (*celecoxib*)

FDA's December 23, 1999 approval of a supplemental NDA for Celebrex to reduce the risk of colorectal cancer in patients with familial adenomatous polyposis (FAP) was based on a surrogate endpoint, reduction in colorectal polyps.

Part 1. Regulatory Factors Weighing into FDA Determination

a. Severity of the Condition

"The average life expectancy for patients with untreated FAP has been estimated to be 42 years."¹⁶³

b. Rarity of the Condition

"The frequency of the FAP gene has been estimated on the basis of disease prevalence to be 1 in 5,000 to 1 in 7,500."¹⁶⁴ Although the prevalence of FAP is sufficiently low, the Sponsor did not seek orphan drug designation.

c. Lack of Available Therapy

"Surgical therapy is the only acceptable option for patients with FAP after colonic polyps have been detected."¹⁶⁵

d. Use of External Expertise

Here are the recommendations of the Oncologic Drugs Advisory Committee that met on December 14, 1999:

i. Do you believe that a reduction in colorectal polyp count in FAP patients in focal areas of some magnitude is "reasonably likely" to predict benefit?

Yes: 13 No: 0 Abstain: 2

ii. Do you believe that the observed reduction (about 25% at 6 months) is likely to predict benefit in FAP patients?

Yes: 12 No: 0 Abstain: 3

iii. Do you recommend approval of Celebrex under the accelerated approval rule for some treatment of FAP?

Yes: 14 No: 0 Abstain: 1¹⁶⁶

¹⁶¹*Id.* at 10.

¹⁶²*Id.* at 14.

¹⁶³CTR. FOR DRUG EVALUATION & RESEARCH, NDA 020998/S007, MEDICAL REVIEW (Dec. 22, 1999), at 25, http://www.accessdata.fda.gov/drugsatfda_docs/nda/99/21156-S007_Celebrex_medr.pdf.

¹⁶⁴*Id.* at 22.

¹⁶⁵*Id.* at 26.

¹⁶⁶*Id.* at 76-77.

Part 2. Understanding of the Disease Process

FAP is characterized by the presence of hundreds to thousands of colorectal adenomatous polyps and the inevitable development of colon cancer. . . . The disease results from germ line mutations of the APC gene. . . . The APC gene is thus believed to be a tumor suppressor gene.¹⁶⁷ A significant body of evidence suggests that the cellular expression of COX-2 is prominent in several types of tumors, including colon . . . as well as pre-cancerous changes such as Barrett's esophagus, the adenomatous polyp and actinic keratosis.¹⁶⁸

Part 3. Understanding of the Relationship Between Reducing Polyp Counts and Colon Cancer

"Celebrex was evaluated in two models of colon cancer. The Min mouse model represents a genetic model of human FAP Adenomas and adenocarcinomas of the colon can be chemically induced in rats by administration of azoxymethane." Celebrex was shown to prevent or inhibit colorectal tumor development in both of these animal models.¹⁶⁹

As for understanding the relationship of drugs in the same or a closely-related class on FAP polyp counts, "[s]tudies have shown that Sulindac, one of the non-selective NSAIDs, induces apoptosis Recent study of COX-2 inhibitors showed that inhibition of COX-2 produced sequential increases in arachidonic acid and ceramide, the latter a potent stimulant of apoptosis. Furthermore, *in vitro* evidence exists that angiogenesis is regulated by COX-2 expression in colon cancer cells. Therefore, another mechanism by which tumor growth may be inhibited by COX-2 inhibitor is through blockade of angiogenesis and tumor vascularization."¹⁷⁰

Part 4. Clinical Evidence on Polyp Counts and on Colon Cancer

"A single, randomized, double-blind, placebo-controlled study has been submitted. A total of 83 patients received treatment with either placebo, Celebrex 100mg BID, or Celebrex 400mg BID for 6 months (with a 1:2:2 randomization) The mean reduction in colorectal polyps count was 28% on the Celebrex 400mg BID arm, 15% on the Celebrex 100mg BID arm and 5% on placebo. Only treatment with Celebrex 400mg BID was associated with a statistically superior mean reduction in polyp counts, with p=0.003."¹⁷¹ In a six-month study there were, as expected, no cases of colon cancer in any arm of the trial.

¹⁶⁷*Id.* at 22–23.

¹⁶⁸Medical Review, *supra* note 163 at 15.

¹⁶⁹*Id.* at 16–17.

¹⁷⁰*Id.* at 15–16.

¹⁷¹*Id.* at 1–2.

13. SYNERCID (*dalfopristin/quinupristin*)

The FDA approval of Synercid on September 21, 1999 was for treating patients with vancomycin-resistant *Enterococcus faecium* (VREF) and was based on a surrogate showing of clearance of the VREF bacteremia.

Part 1. Regulatory Factors Weighing into FDA Determination

a. Severity of the Condition

“The mortality rates in both [pivotal] studies [were] approximately 50%.”¹⁷²

b. Rarity of the Condition

The Sponsor has no intention of developing Synercid for this use, but a “rise in the United States in both the number of nosocomial infections due to *E. faecium* and in the proportion of strains of this pathogen found to be vancomycin-resistant, led to increasing requests for the emergency use of Synercid.”¹⁷³ Synercid appears not to have been granted orphan drug designation. Given the Sponsor’s reluctance to submit an NDA for this use, the Sponsor likely never had applied for designation, even though the condition was rare.

c. Lack of Available Therapy

Those patients who enrolled in the two pivotal trials were only those “infected with VREF who did not have any other therapeutic option.”¹⁷⁴

d. Use of External Expertise

On February 19, 1998, the Anti-Infective Drugs Advisory Committee voted 9 to 1 in favor of approval of Synercid for VREF.

Part 2. Understanding of the Disease Process

The understanding of the pathophysiology of infections with vancomycin-resistant strains of *Enterococcus faecium* is well-known.

Part 3. Understanding of the Relationship Between Clearance of the VREF Bacteremia and Mortality (and Other IDSA/FDA Guideline Clinically Meaningful Endpoints)

The VREF literature is clear that VREF bacteremia . . . should be treated and that clearance of VREF from the bloodstream can be seen as beneficial to the patient . . . [T]here is consensus that bacteremia should be treated. Thus, while clearance of bacteremia is not a clinical benefit by itself, it can be seen as likely to predict clinical benefit. Thus, it is

¹⁷²CTR. FOR DRUG EVALUATION & RESEARCH, NDA 050747, STATISTICAL REVIEW (Mar. 5, 1998), at 18, http://www.accessdata.fda.gov/drugsatfda_docs/nda/99/50747_Synercid_statr.pdf.

¹⁷³CTR. FOR DRUG EVALUATION & RESEARCH, NDA 050747, MEDICAL REVIEW (July 10, 1998), at 2, http://www.accessdata.fda.gov/drugsatfda_docs/nda/99/50747_Synercid_medr_P2.pdf.

¹⁷⁴CTR. FOR DRUG EVALUATION & RESEARCH, NDA 050747, STATISTICAL REVIEW (Mar. 5, 1998), at 3, http://www.accessdata.fda.gov/drugsatfda_docs/nda/99/50747_Synercid_statr.pdf.

proposed that the clearance of VREF bacteremia be viewed as a surrogate endpoint likely to predict clinical resolution of infection.¹⁷⁵

Part 4. Clinical Evidence on VREF Bacteremia Clearance and Mortality

FDA concluded that the four emergency use VREF studies did not provide evidence of an improvement in mortality or resolution of infection due to a host of issues. None of these four studies had a concurrent control and, while FDA had advised that the lack of concurrent control would be acceptable because it would be unethical to include a placebo arm, FDA had stipulated that the studies either: (1) had to show a “dramatic improvement in overall mortality as compared to a historical perspective”¹⁷⁶ and these studies did not (these four studies had mortality rates of 48.8%, 49.5%, 53.8% and 54.0% compared to the VREF literature reporting “all-cause” mortality rates in the range of 30% to 70%);¹⁷⁷ or (2) had to have a historical control and this was not established.¹⁷⁸

While two of the four studies, according to FDA Medical Reviewer, established clearance of VREF bacteremia, only 18% of the patients in these emergency use studies were “evaluable” due primarily to missing data, and there was a low response rate as well.¹⁷⁹ In addition, “in the unevaluable patients who died on therapy but with negative blood cultures, there is the ‘apparent’ clearance of the organism.”¹⁸⁰

14. REMICADE (infliximab)

The August 24, 1998 FDA approval of Remicade to treat patients with Crohn’s disease was based on an intermediate clinical endpoint of a clinical response defined as a reduction in the Crohn’s Disease Activity Index (CDAI) of at least 70 points at the 4-week evaluation.

Part 1. Regulatory Factors Weighing into FDA Determination

a. Severity of the Condition

The prognosis for Crohn’s disease is generally unfavorable The mortality rate increases with the duration of disease and most likely ranges from 5% to 10%. Most deaths occur from peritonitis and sepsis.¹⁸¹

¹⁷⁵CTR. FOR DRUG EVALUATION & RESEARCH, NDA 050747, MEDICAL REVIEW (July 10, 1998), at 33, http://www.accessdata.fda.gov/drugsatfda_docs/nda/99/50747_Synercid_medr_P2.pdf.

¹⁷⁶*Id.* at 31.

¹⁷⁷*Id.* at 19.

¹⁷⁸*Id.* at 31–32.

¹⁷⁹*Id.* at 30, 32.

¹⁸⁰*Id.* at 33.

¹⁸¹CTR. FOR DRUG EVALUATION AND RESEARCH, BLA 98-0012, MEDICAL REVIEW (July 10, 1998), at 5, <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/TherapeuticBiologicApplications/ucm107702.pdf>.

b. Rarity of the Condition

“In . . . the United States, . . . the prevalence is estimated at 20 to 40 per 100,000.”¹⁸² Remicade was designated as an orphan drug on November 14, 1985.

c. Lack of Available Therapy

The FDA Medical Review surveys all the therapies being used and at the time, no robustly effective therapies were available. “Because its cause is unknown, medical management of the disease is largely empirical and is designed to reduce inflammation.”¹⁸³

d. Use of External Expertise

On May 28, 1998, the Anti-Infective and Gastrointestinal Drug Advisory Committees voted unanimously in favor of approval for both: treatment of patients with moderate-severe inflammatory disease refractory to conventional therapy, and treatment of patients with fistulizing Crohn’s disease for the reduction in the number of draining enterocutaneous fistula(s).

Part 2. Understanding of the Disease Process

Crohn’s disease most likely represents a heterogeneous group of disorders. After much effort that has focused on the identification of a specific pathogenic cause, it is being recognized that disease manifestations could result from a combination of any, or all of, a number of factors.¹⁸⁴

Part 3. Understanding of the Predictive Potential of a 70-Point Change in CDAI at Week 4 on Crohn’s Disease

Pathologic review of biopsy . . . often can aid in . . . measurement of extent and severity of disease. Pathologically, Crohn’s disease is described as a transmural disease with focal or microscopic skip areas of inflammation in the lamina propria. The degree of inflammation in the most heavily involved area often is an accurate assessment of the severity of disease Disease activity indices are used to objectively measure the activity of disease for judgment of response in clinical trials. The [CDAI] was developed . . . [in] 1979 . . . to objectively assess response to therapy Although imperfect and cumbersome, e.g., requirement of recording of symptoms for 7 days and for hematocrits, the CDAI remains the most commonly [used] index.¹⁸⁵

As for understanding the relationship between drugs in the same pharmacologic class, Remicade is a chimeric monoclonal antibody to Tumor Necrosis Factor (TNF). As such, Remicade was the first of this kind in a new class of immunomodulatory drugs. Other immunomodulatory drugs, including azathioprine, mercaptopurene, cyclosporine, and methotrexate were accepted for use for long-term treatment of some

¹⁸²*Id.* at 4.

¹⁸³*Id.* at 6.

¹⁸⁴*Id.* at 3.

¹⁸⁵*Id.* at 5.

Crohn's patients. "The mechanism of action of these drugs may involve inhibition of lymphocyte function, primarily that of T cells."¹⁸⁶ As such, they have a different mechanism of action than Remicade.

Part 4. Clinical Evidence on CDAI and on Long-Term Clinical Benefit

Study T16, a placebo-controlled, dose-ranging (n=108) study:

[W]as designed as a Phase 2 trial to determine an effective dose in the acute treatment of patients with active Crohn's disease no[t] responding to immunosuppressant therapy and to explore maintenance therapy with a single dose in patients who respond[ed] initially. This clinical trial became the pivotal trial for licensure of [Remicade] for this indication. . . .¹⁸⁷ 65.1% of the [Remicade] treated patients achieved a clinical response (\geq 70-point reduction from baseline in the CDAI) at the week 4 evaluation compared to 16.7% of the placebo patients ($p < 0.001$). . . . There was no apparent relationship between [Remicade] dose [5mg/kg, 10mg/kg, 20mg/kg] and the proportion of patients responding; the highest clinical response was observed in the 5mg/kg dose group (81.5%; $p < 0.001$ vs placebo).¹⁸⁸

In the Medical Review's Summary Conclusions on the Review of the Safety and Efficacy Data, the Medical Reviewer stated that:

The Sponsor has presented phase 2 clinical data results to support licensing of a potent, novel immunomodulating agent for the management of patients with Crohn's disease, a chronic debilitating disease The number of patients with moderate to severe disease who have received the proposed dose of 5mg/kg . . . is very low (n=28) and no patients have received chronic retreatment with 5mg/kg every 8 weeks as proposed in the original submission. The effects of a single dose [last] approximately 12-16 weeks, compatible with the half-life of the compound. For patients with fistula, although the majority of patients experienced stoppage of drainage in two weeks, there are no data on internal healing of the fistula canal. Once [Remicade] was stopped the effect of therapy was lost. In summary, there are inadequate data to support the long-term benefit of [Remicade] in patients with either fistulizing or moderate/severe disease.¹⁸⁹

From the conclusions of the Medical and Statistical Reviews, there appear to have been some concerns among FDA Reviewers as to the appropriateness of the short-term (CDAI improvement after 4 weeks) surrogate endpoint as being adequate to predict long-term benefit in a chronic disease. The conclusion of the Statistician on Study T16 in moderate to severe Crohn's disease patients was redacted from the

¹⁸⁶*Id.* at 6.

¹⁸⁷CTR. FOR DRUG EVALUATION & RESEARCH, BLA 98-0012, MEDICAL REVIEW (July 10, 1998), at 11, <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/TherapeuticBiologicApplications/ucm107702.pdf>.

¹⁸⁸*Id.* at 20.

¹⁸⁹*Id.* at 82.

publicly available version of the Statistician's Review. However, there was a second Phase 2 study in patients with Crohn's disease with fistula, Study T20, which is referred to in the conclusions of the Medical Review. From the information in the Statistician's Review of Study T16 that was made publicly available, it would seem that the Statistician's conclusions with respect to Study T20 may have closely paralleled those for Study T16. With respect to Study T20, here are the Statistician's conclusions:

Although the differences in response rates between the placebo group and the [Remicade] treated groups were statistically significant, questions remain about the durability of response. Patients received doses at weeks 2, 4, and 6, but this dosing strategy should be thought of as one-time dosing. After 6 months of follow-up, the drug effect had disappeared and the proportion of responding patients in the placebo arm was similar to the proportions in the treatment arms. The data suggest, therefore, that although this agent has an initial beneficial effect on Crohn's disease, a single set of doses is unlikely to provide durable benefit in this chronic disease. There are no data to assess chronic use of [Remicade] for this indication. There is no information regarding the formation of neutralizing antibodies (HACA) with repeated dosing and how this may affect the efficacy of this product. There is also no safety data to allay concerns of a possible increase in malignancies or serious infections. The Agency should carefully weigh the observed early benefits seen with this product against the paucity of information regarding the safety and efficacy of repeated use for this chronic indication.¹⁹⁰

15. PRIFTIN (*rifapentine*)

On June 22, 1998, FDA approved Priftin for treating pulmonary tuberculosis (TB), and this approval was based on a surrogate of a 6-month relapse rate as contrasted with the standard 2-year relapse rate information for a traditional approval.

Part 1. Regulatory Factors Weighing into FDA Determination

a. Severity of the Condition

"[TB] is the leading infectious cause of morbidity and mortality worldwide."¹⁹¹

b. Rarity of the Condition

"In 1990, there were 25,701 new cases of TB reported in the [U.S.]"¹⁹² Priftin was designated as an orphan drug on June 9, 1995.

c. Lack of Available Therapy

During development of rifapentine for TB, the applicant was encouraged to submit 6-month follow-up data from one study, under the accelerated approval regulations (21

¹⁹⁰CTR. FOR DRUG EVALUATION & RESEARCH, BLA 98-0012, STATISTICAL REVIEW (Aug. 5, 1998), at 13, <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/TherapeuticBiologicApplications/ucm107710.pdf>.

¹⁹¹CTR. FOR DRUG EVALUATION & RESEARCH, NDA 021024, MEDICAL REVIEW, PT. 1 (June 19, 2008), at 15, http://www.accessdata.fda.gov/drugsatfda_docs/nda/98/021024a-1a.pdf (pt. 1).

¹⁹²*Id.*

CFR 314 Subpart H). There is a need for new anti-tuberculosis medications, and for medications which will potentially increase the adherence to dosing thereby decreasing the potential for the development of resistant organisms. It was anticipated that rifapentine would be such an agent. Six-month relapse data would serve as a surrogate for two-year relapse data predictive of long term clinical benefit.¹⁹³

FDA had previously approved rifampin for use in treating TB.

d. Use of External Expertise

At the Anti-Viral Advisory Committee Hearing on May 5, 1998, “[t]he committee voted to recommend approval of [Priftin] for the treatment of pulmonary tuberculosis, with only one dissenting vote.”¹⁹⁴

Part 2. Understanding of the Disease Process

The pathophysiology of TB is well-understood.

Part 3. Understanding of the Relationship Between Six-Month Relapse Rate and Two-Year Relapse Rate and Mortality

The Medical Review stated that “[i]t is expected that the majority of relapses will occur by 6 months of follow-up, however, the ‘gold standard’ is 2 year relapse rate.”¹⁹⁵ However, the pattern of relapses for Priftin does not appear to reflect the same showing of relapses in the latter half of six-month follow-up that was seen for rifampin in the pivotal study. See discussion of results under Part 4.

Part 4. Clinical Evidence on Six-Month and Two-Year Relapse Rates

The single pivotal trial was an open-label, randomized, two-arm parallel, rifampin-controlled trial with 570 patients in the modified ITT analysis.

The primary efficacy endpoint for this accelerated approval review was treatment outcome at the end of 12 months (6 months of active treatment + 6 months of follow-up). This was a binary variable with success defined as achieving a negative sputum culture during active treatment and sustaining it to the end of [6] months of follow-up.¹⁹⁶

“There is essential equivalence for [negative sputum culture] rates at the end of [the 6-month active treatment] between the rifampin [(83% negative sputum cultures)] and [Priftin (88% negative sputum cultures)] arms.”¹⁹⁷ However, “[t]here is a statistically

¹⁹³*Id.* at 18.

¹⁹⁴CTR. FOR DRUG EVALUATION & RESEARCH, NDA 021024, MEDICAL REVIEW, Pt. 2 (June 19, 2008), at 36, http://www.accessdata.fda.gov/drugsatfda_docs/nda/98/021024a-1b.pdf (pt. 2).

¹⁹⁵CTR. FOR DRUG EVALUATION & RESEARCH, BLA 98-0012, MEDICAL REVIEW (July 10, 1998), at 29, <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/TherapeuticBiologicApplications/ucm107702.pdf>.

¹⁹⁶CTR. FOR DRUG EVALUATION & RESEARCH, NDA 021024, STATISTICAL REVIEW (July 27, 1998), at 6, http://www.accessdata.fda.gov/drugsatfda_docs/nda/98/021024a-3-StatReview.pdf.

¹⁹⁷CTR. FOR DRUG EVALUATION & RESEARCH, NDA 021024, MEDICAL REVIEW, Pt. 2 (June 19, 2008), at 14, http://www.accessdata.fda.gov/drugsatfda_docs/nda/98/021024a-1b.pdf (pt. 2).

significant difference between the treatment arms for relapse The risk is 5% for rifampin . . . and 11% for [Priftin].”¹⁹⁸ The Statistical and Medical Reviews agree that while 10 of the 11 relapses on rifampin occurred within the first 6 months of follow-up, 7 relapses occurred in the Priftin arm at time points between 6 and 12 months of follow-up. (Note: While the endpoint was at 6 months of follow-up, almost all subjects had had 12 months of follow-up, so FDA analyzed the 12 months of follow-up data as well and noted that the Priftin arm continued to experience sizable numbers of relapses beyond the first 6 months of follow-up, which was much different than the pattern of relapses observed for rifampin).

Despite the above discrepancy between the rifampin and Priftin arms in relapse rate beyond 6 months, FDA reviewers (and the Advisory Committee members) seemed to believe that this may reflect a lack of optimized dosing of Priftin, rather than a lack of confidence in the prognostic surrogate of 6-month relapse rate predicting 2-year relapse rate, and eventually, survival. However, at the time of approval there appears to be no clinical evidence of Priftin on 2-year relapse rate or on mortality.

16. SULFAMYLON (mafenide acetate)

FDA approved Sulfamylon on June 5, 1998, “to control bacterial infection when used under moist dressings over meshed autografts on excised burn wounds.” The approval was based on an intermediate clinical endpoint of evidence derived from patients who were burned over up to 20% of their total body surface area (TBSA) with a Phase 4 commitment to conduct a confirmatory trial in patients with 20% to 60% TBSA thermal injuries.

Part 1. Regulatory Factors Weighing into FDA Determination

a. Severity of the Condition

The Medical Review commenting on the results of the single pivotal trial (done exclusively in children) observed the following: “It is remarkable that so many of these severely burned children survived to leave the hospital It is not unexpected that survival rates fall as TBSA burned increases.”¹⁹⁹ Large [TBSA] burns are serious and life-threatening.”²⁰⁰

b. Rarity of the Condition

The number of persons in the country in need of such care is small, thankfully, very small. FDA designated Sulfamylon as an orphan drug for this use for two different Sponsors at separate times: on August 29, 1995 and on July 18, 1990.²⁰¹

c. Lack of Available Therapy

“There is no existing approved treatment for these burn patients who require excision and meshed autografts.”²⁰²

¹⁹⁸*Id.* at 15.

¹⁹⁹CTR. FOR DRUG EVALUATION & RESEARCH, NDA 19-832, MEDICAL REVIEW (Nov. 28, 1997), at 37, http://www.accessdata.fda.gov/drugsatfda_docs/nda/98/19832-1.pdf.

²⁰⁰CTR. FOR DRUG EVALUATION & RESEARCH, NDA 19-832, APPROVAL PACKAGE (June 5, 1998), at 66, http://www.accessdata.fda.gov/drugsatfda_docs/nda/98/19832-1.pdf.

²⁰¹*Id.* at 23.

²⁰²*Id.* at 67.

d. Use of External Expertise

Sulfamylon [was] discussed by FDA Anti-Infective Drug Products Advisory Committee [on July 24, 1996]. The Committee concluded that since topical antimicrobial solutions had evolved to a standard of care [(SOC)] over the last 20 years, a placebo-controlled study would be unethical.²⁰³

Part 2. Understanding of the Disease Process

There is adequate evidence available in the literature to establish that wounds (including burn wounds) may be expected to progress satisfactorily if the microbial load present is reduced to less than 10^5 organisms per gram of tissue . . . [I]t may be said that if a topical antimicrobial is successful in maintaining low bacterial levels on a newly placed skin graft until the graft is adequately vascularized, the antimicrobial has contributed to take of the graft.²⁰⁴

Part 3. Understanding of the Relationship Between the Treatment Failures in Those with <20% TBSA Burned and Treatment Failures in Those with >20% TBSA Burned

“The applicants have been reluctant to use a vehicle control on the grounds that failure to treat a burn patient with a [TBSA] burn of larger than 10-20% would be unethical.”²⁰⁵ This was supported by the deliberations of the Advisory Committee. Therefore, while the single pivotal trial enrolled all patients with burns, regardless of how extensively the body was burned, there was:

[N]o protocol-specified assignment of patients to treatment with [either Sulfamylon or standard of care (SOC)]. This was a medical decision, made by the attending physician . . . The reviewers separated the results into patient groups by TBSA burned. All patients who had burns covering more than 40% TBSA were treated with [Sulfamylon] . . . It is impossible to assess the effect of [Sulfamylon] in this group. In the 20-40% TBSA burn group, there were a few patients who received [SOC] but . . . the contribution of [Sulfamylon is] difficult to quantify. However, there [were] sufficient [SOC] patients in the 0-20% TBSA burn group to permit comparison of the two treatment regimens.²⁰⁶

As for understanding the relationship between drugs in the same pharmacologic class as Sulfamylon, “Sulfamylon for 5% Topical Solution” is the drug product that was the subject of this NDA. However, “[s]ulfamylon cream is currently approved for use in the treatment of second and third degree burns and the proposed indication for the Sulfamylon 5% Solution is related.”²⁰⁷ “Because of the pain caused by the

²⁰³*Id.* at 23.

²⁰⁴*Id.* at 62.

²⁰⁵CTR. FOR DRUG EVALUATION & RESEARCH, NDA 19-832, MEDICAL REVIEW (Nov. 28, 1997), at 24, http://www.accessdata.fda.gov/drugsatfda_docs/nda/98/19832-1.pdf.

²⁰⁶*Id.* at 65.

²⁰⁷*Id.* at 66.

cream, burn physicians began to make a 5% solution using mafenide acetate powder in the mid-1970s . . . [A]nd the 5% solution has become the standard of use in some burn units for maintaining skin grafts in the period between graft placement and graft take.”²⁰⁸

Part 4. Clinical Evidence on Those with <20% and Those with >20% TBSA Burned

The single pivotal efficacy study was an unblinded, retrospective, non-randomized, parallel group study with an active control of standard of care (SOC) and was conducted at a single site and with a single investigator: Dr. Glenn Warden at Shriners’ Burn Institute in Cincinnati, Ohio.

In this study, among the 229 procedures in persons with less than 20% TBSA burned, there were 19 (19%) who were “treatment failures” in those treated with Sulfamylon compared to 33 (26%) who failed on SOC. However, those treated with Sulfamylon had more serious burns, that is, third-degree burns (6.5% vs. 3.3% SOC), a higher percentage of the body surface area burned (10.6% vs. 7.0% SOC), and fewer participants with only less serious burns, that is, those with second-degree burns only (4.4% vs. 17.3% SOC).

In her review, the reviewing Statistician, Dr. Yulan Li, reached the following conclusion: “Based on the Cincinnati study, the applicant has demonstrated that the use of [Sulfamylon] is associated with the decreasing of treatment failure in the subgroup of patients with 0-20% TBSA burns after separately adjusting for etiology and degree of burn. However, it is unknown whether . . . treatment failure reflects the benefit of [Sulfamylon] due to non-random treatment assignment and investigator knowledge of treatment at the time treatment failure was assessed.”²⁰⁹

While there appears to be no disagreement in any FDA review as to the intermediate clinical endpoint of effect in those with less than 20% TBSA burned as “reasonably likely to predict benefit” in those with burns over more than 20% TBSA; there were concerns expressed, especially by the Statistician, as to the strength of the efficacy evidence for the findings in those with less than 20% TBSA burned.²¹⁰

17. PROAMATINE (midodrine hydrochloride)

FDA approved Proamatine for treating “symptomatic orthostatic hypotension” on September 6, 1996 on the basis of “increases in 1-minute standing systolic blood pressure, a surrogate marker likely to correspond to a clinical benefit” (as stated in FDA-approved labeling).²¹¹

²⁰⁸*Id.* at 24.

²⁰⁹*Id.* at 26.

²¹⁰While scored as a “1,” the strength of clinical evidence on the surrogate here with Sulfamylon could reasonably be scored as either “1” or “zero,” and the same may be said of the strength of clinical evidence for the surrogate in Synercid, Precedent #13.

²¹¹All of the formation in this analysis is drawn from FDA-approved labeling, as no Medical or Statistical Reviews from FDA were publicly available.

Part 1. Regulatory Factors Weighing into FDA Determination

a. Severity of the Condition and Lack of Alternative Therapy

Although the review documents for Proamatine are not publicly available on FDA's website, the Agency's Subpart H approval of Proamatine has to mean that FDA assessed the condition as rather serious and lacking available therapy.

b. Rarity of the Condition

Proamatine was designated as an orphan drug on June 21, 1985.

c. Use of External Expertise

There is no evidence from documents currently available, including approved labeling and trade press, whether FDA sought the advice of an Advisory Committee.

Part 2. Understanding of the Disease Process

For FDA to have approved Proamatine on the basis of a change in 1-minute systolic blood pressure suggests that FDA must have considered that there was a sound understanding of the pathophysiology of the disease.

Part 3. Understanding of the Relationship Between Change in 1-Minute Systolic Blood Pressure and the Ability to Perform Life Activities

Since there were no other drugs in any class approved for this condition, FDA could not have relied upon their effects on this disease. However, many drugs are approved on changes in blood pressure as a validated surrogate based upon both robust epidemiology and multiple interventions affecting serious cardiovascular outcomes such as major adverse cardiac events (MACE), and FDA may have relied upon this strong association for support of the power of a change in 1-minute systolic blood pressure in this disease to predict clinical benefit in this disease.

Part 4. Clinical Evidence on 1-Minute Systolic Blood Pressure and Clinical Outcome

Midodrine has been studied in 3 principal controlled trials, one of 3-weeks duration and two of 1-to-2 days duration. All studies were randomized, double-blind and parallel-design trials in patients with orthostatic hypotension of any etiology and supine-to-standing fall of systolic blood pressure of at least 15 mmHg accompanied by at least moderate dizziness/lightheadedness In a 3-week study in 170 patients . . . the midodrine-treated patients . . . had significantly higher (by about 20 mmHg) 1-minute standing systolic pressure 1 hour after dosing . . . for all 3 weeks. After week 1, midodrine-treated patients had small improvements in dizziness/lightheadedness/unsteadiness scores and global evaluations, but these effects were made difficult to interpret by a high early drop-out rate (about 25% vs. 5% on placebo). Supine and sitting blood pressure rose 16/8 and 20/10 mmHg, respectively, on average. In a 2-day study, after open-label midodrine, known midodrine responders received midodrine 10 mg or placebo at 0, 3, and 6 hours. One-minute standing systolic blood pressures were increased 1 hour after each dose by about 15 mmHg and 3 hours after each dose by about 12 mmHg;

3-minute standing pressures were increased also at 1, but not 3, hours after dosing. There were increases in standing time seen intermittently 1 hour after dosing, but not at 3 hours. In a 1-day, dose-response trial, single doses of 0, 2.5, 10, and 20 mg of midodrine were given to 25 patients. The 10- and 20-mg doses produced increases in standing 1-minute systolic pressure of about 30 mmHg at 1 hour; the increase was sustained in part for 2 hours after 10 mg and 4 hours after 20 mg. Supine systolic pressure was =200 mmHg in 22% of patients on 10 mg and 45% of patients on 20 mg; elevated pressures often lasted 6 hours or more.²¹²

18. BIAXIN (clarithromycin)

FDA approved Biaxin on December 23, 1993 for treating disseminated mycobacterial infections due to mycobacterium avium complex (MAC) on the basis of a showing of Biaxin's effect on the surrogate of decreases in MAC bacteria.

Part 1. Regulatory Factors Weighing into FDA Determination

a. Severity of the Disease, Rarity, and Lack of Alternative Therapy

The pivotal studies were conducted in persons with CDC-defined AIDS and CD₄ counts <100 cells/ μ L, and median survival time in the one trial that was randomized and blinded was 249 days and 215 days for the two dose groups reported in the approved labeling.²¹³

While Biaxin was not designated as an orphan drug for this use, this condition was not prevalent and the absence of orphan drug status is likely due to FDA approval of Biaxin for many other prevalent diseases (such that orphan drug exclusivity would have had substantially diminished, if any, value).

b. Use of External Expertise

On May 11, 1993, the Antiviral Drugs Advisory Committee provided insight on the approvability of Biaxin for treatment of MAC.²¹⁴

Part 2. Understanding of the Disease Process

The pathophysiology of MAC in immune-compromised AIDS patients was likely understood relatively well for the extent of time that the condition had been known.

Part 3. Understanding of the Relationship Between Reducing MAC Bacteremia and Clinical Outcomes

The general axiomatic principles of infectious disease likely guided and illuminated FDA's interpretation of the prognostic value of reducing MAC bacteremia on achieving negative cultures and clinical benefit. Other antibiotic regimens had shown some value as well.

²¹²Proamitine (midodrine hydrochloride) Label, NDA 019815, 41 (Sept. 1996).

²¹³There were no FDA medical or statistical reviews publicly available and nearly all information is from FDA approved labeling.

²¹⁴Based on public documents currently available, it is unclear what the outcome of this Advisory Committee was.

Part 4. Clinical Evidence on Reducing MAC Bacteremia and Clinical Outcomes Including Mortality

Of the 3 studies conducted from May 1991 to March 1992, Study 500 was the only one to be blinded and randomized (dose comparison trial of 3 different doses of Biaxin). Study 500 showed a reduction in MAC bacteremia with the lowest dose having the smallest decrease in colony-forming units (CFUs). There was seemingly no survival benefit, as FDA-approved labeling reported that: “The median survival times for these [Biaxin] dosages were similar to recent historical controls with MAC when treated with combination therapies.” However, there was some evidence of improvement in other signs and symptoms of MAC infection including night sweats, fever, and weight loss.

19. BETASERON (interferon beta-1b)

FDA approved Betaseron as the first therapy to treat multiple sclerosis (MS) on July 23, 1993 on the basis of a showing on both rate and extent of exacerbations and on improvement in MRI-measured lesion area.

Part 1. Regulatory Factors Weighing into FDA Determination

a. Severity, Rarity, and Lack of Available Therapy

MS is a serious disease for which, prior to Betaseron, there was no FDA approved treatment. Betaseron was designated as an orphan drug on November 17, 1988.

b. Use of External Expertise

The FDA Peripheral and Central Nervous System Advisory Committee on March 19, 1993 voted 7-2 to recommend approval of Betaseron.

Part 2. Understanding of the Disease Process

The pathophysiology of multiple sclerosis was known to a fair degree at the time of the conduct of the pivotal trial which permitted the Sponsor, in collaboration with the lead FDA CBER official, Dr. Janet Woodcock, and her office, to have general agreement on co-primary endpoints of clinical utility related to exacerbations, as well as somatic measures of the putatively key causal biologic marker, MRI lesion volume.

Part 3. Understanding of the Relationship Between MRI Lesion Volume and Multiple Sclerosis

[I]t was also clear that the Committee as a whole placed great weight on the MRI findings in their deliberations. Specifically, although the clinical benefit, as measured by the proportion of exacerbation-free patients and exacerbation frequency, was considered real and of value clinically, the Committee considered the size of the treatment effect to be relatively small.

However, it was obvious that great emphasis was placed on the MRI findings. Specifically, the Committee appeared convinced by the firm’s presentation that the drug had an important effect on the underlying pathology as measured by total lesion area as seen on MRI. The statistically significant decrease in the total lesion area in the high dose group as compared to placebo patients over the course of the study that the sponsor claimed was demonstrated was interpreted by the Committee,

in my view, as powerful support for the conclusion that the drug was having an important effect on the underlying disease process. While the Committee stopped short of declaring that the data proved that the drug had an effect on the progression of the disease, I believe it is fair to characterize their view with a quote, made at the meeting, by Dr. McFarland, who said at one point, that, while the sponsor had not proved that the drug had an effect on the course of the disease, "I would be amazed if it didn't change the course of disease." A number of Committee members explicitly referred to Dr. McFarland's comments in this regard when explaining their votes.²¹⁵

That is, it appears clear that the Committee felt that the MRI results not only were consistent with the clinical benefit observed (that is, the changes seen corresponded to the exacerbation rate data at a given point in time), but that they could be relied upon to accurately "predict" patients' future courses. In other words, the MRI data were considered, for all intents and purposes, as a surrogate marker for disease progression.²¹⁶

If the lesions detected on MRI are taken to be a better index of the "activity" of the pathologic process than are clinical manifestations of MS, (a not unreasonable possibility given the knowledge that lesions detected on MRI may be unaccompanied by clinical signs/symptoms when they occur in so-called "silent" regions of the CNS) *and if* the rate of clinical progression of MS (in the sense of increasing physical disability) is a positive function of the activity of that pathologic process, *it follows logically* that any drug suppressing this "activity" "must"²¹⁷ have some beneficial effect on the progression of MS (as manifest by increasing physical disability). Although the clinical evidence collected²¹⁸ in Study TB01-35(6/8)86 does not provide convincing affirmative support for this hypothesis, that does not necessarily undercut its appeal or its psychological impact on those asked to render an opinion about the therapeutic potential of Betaseron.

During the PCNS meeting, the sponsor's representatives, several members of the Committee and, in particular, Dr. Henry McFarland, who was attending the meeting as the Agency's expert consultant on neuro-imaging and MS, espoused the hypothesis just described. Although

²¹⁵CTR. FOR DRUG EVALUATION & RESEARCH, PLA 92-04952, MEDICAL REVIEW (May 28, 1993), at 10-11, http://www.accessdata.fda.gov/drugsatfda_docs/nda/pre96/103471s0000_MedR.pdf.

²¹⁶*Id.* at 13-14.

²¹⁷"Must" appears in quotations as a reminder of prior occasions in the history of therapeutics where perfectly logical extrapolations based on beliefs about the pathophysiology of a disease and the postulated mechanism of a drug's action have led experts to reach totally incorrect conclusions about the promise of a particular drug (e.g., CAST: the suppression of ventricular ectopy "must" save lives.) [Footnote is part of quotation].

²¹⁸In their report of the study, the Sponsor asserts that the correlation between EDSS disability scores and MRI lesion areas detected at both baseline ($r=0.169$) and at the end of year two ($r=0.2$) establishes that MRI 'burden' predicts disability (EDSS score). Although these statements are correct in a statistical sense, the correlation does not tell us what we really seek to learn: whether a treatment reducing the extent of MRI area increase over time will reduce the extent of clinical worsening, as judged by EDSS, over the same interval or in a future one. [Footnote is part of quotation].

virtually all proponents of this hypothesis acknowledged that the link between MRI lesion frequency/intensity/area and subsequent outcome (progression in level of physical disability) in MS was not proven, almost all affirmed that they would be very surprised if the link was not eventually demonstrated. Thus, for many experts, the number and area of lesions detected on MRI are tantamount to a “surrogate” endpoint that predicts disease progression in MS.²¹⁹

In the Betaseron data there is a second kind of replication, the MRI results, which are more or less persuasive, depending on one’s beliefs. At a minimum, as Dr. Leber says, these data are an independent measurement that supports the clinical finding, a kind of “within-study” replication. At best, they are evidence of an effect far more important than the modest effect on exacerbations. I certainly am not qualified to choose between these interpretations, but our advisors seem to believe the latter, even though all would agree that, strictly, the correlation of improved clinical outcome and improved MRI has not been made.

It would be possible, we believe, to grant approval under the accelerated approval regulations, which allow this procedure where a surrogate or clinical, but *non-ultimate endpoint* is the basis for approval.²²⁰

Part 4. Clinical Evidence on MRI Lesion Volume and on Reduction in Exacerbations of MS

The trial was designed as a randomized, double-blind, and placebo-controlled study to evaluate the safety and efficacy of Betaseron in the treatment of patients with relapsing-remitting MS. The protocols proposed that the primary efficacy evaluations will be based on reduction in frequency of exacerbations per subject and proportion of exacerbation-free subjects.²²¹

The proportions of exacerbation-free subjects in the three arms of the study are given in Table 1. If we consider all reported exacerbations, 18 of the 112 placebo patients (16.1%) and 36 of the 115 45 mIU Betaseron patients (31.3%) were exacerbation-free. This difference was significant at $p=0.008$.²²²

The second primary endpoint, prospectively specified in the protocol, was the frequency of exacerbation per subject If we consider the outcomes in all six categories of exacerbations (i.e., 0, 1, 2, 3, 4, and 5+)

²¹⁹CTR. FOR DRUG EVALUATION & RESEARCH, PLA 92-04952, MEDICAL REVIEW (May 28, 1993), at 22–23, http://www.accessdata.fda.gov/drugsatfda_docs/nda/pre96/103471s0000_MedR.pdf.

²²⁰CTR. FOR DRUG EVALUATION & RESEARCH, PLA 92-04952, ADMINISTRATIVE AND CORRESPONDENCE DOCUMENTS (1993), at 13–14, http://www.accessdata.fda.gov/drugsatfda_docs/nda/pre96/103471s0000_AdminCorres.pdf.

²²¹CTR. FOR DRUG EVALUATION & RESEARCH, PLA 92-04952, STATISTICAL REVIEW (May 20, 1993), at 11, http://www.accessdata.fda.gov/drugsatfda_docs/nda/pre96/103471s0000_StatR.pdf.

²²²*Id.* at 13.

then the probability of better response on Betaseron therapy is 63%. It is significantly different ($p=0.0004$) from 50%.²²³

As for the MRI lesion volume results, depending upon the analysis used by FDA reviewer, Dr. Jay Siegel, the p-value for the comparison between Betaseron and placebo arms ranges from a p-value of 0.03 to a p-value of 0.001.²²⁴

²²³*Id.* at 15–16.

²²⁴CTR. FOR DRUG EVALUATION & RESEARCH, PLA 92-04952, ADMINISTRATIVE AND CORRESPONDENCE DOCUMENTS (1993), at 17.