

# Quantum of Effectiveness Evidence in FDA's Approval of Orphan Drugs

## Cataloguing FDA's Flexibility in Regulating Therapies for Persons with Rare Disorders

by Frank J. Sasinowski, M.S., M.P.H., J.D.<sup>1</sup>

Chairman of the Board

National Organization for Rare Disorders

One of the key underlying issues facing the development of all drugs, and particularly orphan drugs, is what kind of evidence the Food and Drug Administration (FDA) requires for approval. 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." 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 its benefit to patients.

The sole law that applies specifically to orphan drugs, the Orphan Drug Act of 1983, provided financial incentives for drug companies to develop orphan drugs, which is legally defined as products that treat diseases that affect 200,000 or fewer patients in the U.S. But the Orphan Drug Act, whose enactment was championed by the National Organization for Rare Disorders (NORD), did not amend or revise the statutory standards in the law for establishing that a new medicine is safe and effective for its proposed use. From a strict regulatory standpoint, the standard for orphan drugs is identical to the standard required for all other drugs, namely that "substantial evidence" demonstrates the effectiveness of the drug for its intended uses.

In the past decades FDA has moved in two broad formal ways to establish policies that provide greater flexibility for medical reviewers in assessing applications for new drugs. Neither of these efforts was designed specifically for orphan therapies. First, in response to the AIDS crisis and need for new cancer therapies, FDA established regulatory systems that formally recognized the need for flexibility in FDA's review of therapies for serious diseases for which there is an unmet medical need. Such systems found expression in FDA's promulgation in 1988 of the IND Subpart E regulation (21 C.F.R. Part 312) and in 1992 of the NDA Subpart H regulation (21 C.F.R. Part 314) (sometimes referred to as the "accelerated approval" regulation). Second, in its pursuit of good regulatory science, FDA announced a seminal guidance in May 1998 on "Providing Clinical Evidence of Effectiveness" in which FDA described nine different ways for a new therapy to get approved on the basis of a single adequate and well-controlled trial. With this guidance, FDA created new regulatory tools for addressing the needs of patients while meeting the legal obligations to ensure that all new therapies are both safe and

effective for their intended uses.

FDA has for many decades acknowledged that there is a need for flexibility in applying its standard for approval. For example, one of FDA's regulations states that: "FDA will approve an application after it determines that the drug meets the statutory standards for safety and effectiveness... While the statutory standards apply to all drugs, the many kinds of drugs that are subject to the statutory standards and the wide range of uses for those drugs demand flexibility in applying the standards. Thus FDA is required to exercise its scientific judgment to determine the kind and quantity of data and information an applicant is required to provide for a particular drug to meet the statutory standards." 21 C.F.R. § 314.105(c).

FDA publicly has expressed sensitivity to applying this flexibility to new therapies for rare disorders. For example, in his testimony to the United States Senate on June 23, 2010, Dr. Jesse Goodman, FDA Chief Scientist and Deputy Commissioner for Science and Public Health, testifying on "FDA's Efforts on Rare and Neglected Diseases," said: "FDA is fully committed to applying the requisite flexibility in the development and review of products for rare diseases, while fulfilling its important responsibility to assure that the products are safe and effective for these highly vulnerable populations. There are numerous examples of drugs approved for treating rare diseases where FDA's flexibility and sensitivity to the obstacles of drug development for rare diseases has brought forth a successful treatment. Many of the 357 approved orphan drugs have been successfully tested on extremely limited numbers of patients, serving as a testament to FDA's commitment to these patients. This is possible when the best science is flexibly applied and when therapies are truly effective."

Dr. Goodman cited as successful examples the following:

\*\*\* Carbaglu (carglumic acid) for the treatment of NAGS deficiency, the rarest of the Urea Cycle Disorders (UCDs): This disease affects fewer than 10 patients in the U.S. at any given time and fewer than 50 patients worldwide. This drug was approved in March 2010 based on a case series derived from fewer than 20 patients and comparison to a historical control group.

\* VPRIV (velaglucerase) for the treatment of Gaucher disease, a rare genetic disorder: This disease affects approximately 2,000 people in the U.S. and approximately 5,000 worldwide. This drug was approved in February 2010 based on a development program that included about 100 patients and a pivotal study of 25 patients.

<sup>1</sup> Frank J. Sasinowski is a director with the law firm of Hyman, Phelps & McNamara, P.C. and Chairman of the Board of Directors of the National Organization for Rare Disorders (NORD).

\* Myozyme (alglucosidase alfa) for the treatment of infantile variant, a rapidly fatal form of Gaucher disease: The variant of this disease affects about 1,000 patients in the U.S. and about 3,000 patients worldwide. This drug was approved in April 2006 based on a clinical development program of fewer than 80 patients and a pivotal study that included 18 patients.

\* Ceprotin (human plasma derived protein C concentrate) for the treatment of severe congenital Protein C deficiency: There are fewer than 20 known patients with this disorder in the United States. This biological drug product was approved in March 2007 based on a study of 18 patients using comparison to historical control data.”<sup>2</sup>

### PURPOSE OF THIS STUDY

NORD designed this study to examine closely how much flexibility FDA provides in reviewing orphan drugs – that is, to determine whether FDA requires that orphan drug applications provide the conventional or traditional level of proof of effectiveness that is ordinarily expected for most drugs for more prevalent diseases. This issue is especially critical because the patient population available for testing of orphan drugs is by definition more limited than for drugs for more prevalent diseases. The National Institutes of Health estimates that there are as many as 7,000 rare diseases, with some affecting only a handful of patients. The numbers of persons with such disorders can vary, as for example, cystic fibrosis which affects 35,000 Americans, or infant botulism, which affects, at most, only a few hundred infants per year. This study examines whether FDA exercises flexibility when reviewing applications for these diseases and, if so, illustrates the nature and scope of that flexibility.

This paper specifically examines the quantum of effectiveness evidence that provided the basis for FDA's approval of the 135 non-cancer orphan drug new chemical entities that were approved between the orphan drug law's creation in 1983 and June 30, 2010. The intent was to catalogue each of the 135 orphan drugs according to whether its approval had demonstrated any exercise of scientific judgment or flexibility by FDA in reaching its conclusion that the statutory requirement for demonstrating that drug's effectiveness had been met. The study aims to determine, based on an examination of the publicly-available information used to support approval, whether the amount of data presented would have satisfied the conventional requirements for proving the effectiveness of the drug.

The examination of 135 orphan drugs found that 90 approvals were based on some exercise of flexibility by FDA. That is, the study supports the FDA assertion that it exercises flexibility when reviewing applications for orphan drugs. This study also catalogues the types of situations in which the FDA has elected to exercise that flexibility.

### METHODS

<sup>2</sup> FDA Deputy Commissioner Dr. Jesse Goodman, Testimony before U.S. Senate Appropriations Committee Agriculture Subcommittee, at p.2 (June 23, 2010).

To identify the non-cancer orphan drugs approved as new chemical entities, NORD relied upon FDA's publicly-available documents for drugs approved by FDA from January 1983 to June 30, 2010.

For each approved drug, NORD sought to access 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 the FDA officials who approved the products, and the reviews of the medical and statistical officers. While such documents were retrievable in most cases, only subsets of these documents were recoverable for some drugs, especially for several of the earliest approved orphan therapies.

The evidence explaining the basis for each drug's approval was analyzed and classified, in the judgment of NORD, as whether or not it would have met the usual and customary conventional showings of effectiveness that would ordinarily be expected for any disorder, including a common or prevalent disorder. In addition to this classification, the category of 90 non-cancer orphan drugs whose approval was based on some exercise of FDA flexibility was further analyzed and subdivided into either those which were based on a formal, expressed FDA system for flexibility (“administrative flexibility”) or were not based on any such formal FDA expression of flexibility (“case-by-case flexibility”).

In summary, this paper classifies the 135 orphan drug approvals into one of three categories based on the analysis of the quantum of effectiveness evidence:

1. “conventional” or traditional quantum of evidence;
2. evidence consistent with some formal FDA system for exercising discretion or “administrative flexibility”; or
3. evidence that is consistent with a “case-by-case flexibility”.

The first two of these classifications are described below and the third category is one by exclusion. All available source documents were gathered and analyzed for each FDA approval in order to classify each approved orphan therapy approval as “Conventional,” “Administrative Flexibility” or “Case-by-Case Flexibility” (see Figure 1).

#### 1. Conventional or Traditional Showing of Effectiveness

This category is for those drugs whose quantum of effectiveness evidence would satisfy the usual, conventional, traditional showing of effectiveness, which most often is colloquially and commonly referred to as “the two adequate and well-controlled studies” standard.

The 1962 Amendments to the FD&C Act added the requirement that for FDA to approve for commercial marketing any drug, it had to conclude that there exists “substantial evidence...consisting of adequate and well-controlled investigations, including clinical investigations” such that “experts

qualified by scientific training and experience to evaluate the effectiveness of the drug involved" could "fairly and responsibly" conclude that the drug will have the effects that the drug purports or claims to have in the sponsor's proposed labeling for that therapy. FD&C Act § 505(d). FDA has interpreted "adequate and well-controlled studies" to mean generally a minimum of two such studies. FDA has promulgated regulations defining the types of trial designs that are "adequate and well-controlled studies." 21 C.F.R. § 314.126. Traditionally, FDA has accepted two adequate and well controlled trials when each meets its primary endpoint by its prespecified primary analysis with a p value of less than 0.05.

## 2. Administrative Flexibility in Formal Expressed Systems

There are three major expressions of formal ways in which FDA exercises scientific discretion in assessing the effectiveness evidence of all drugs, not just for orphan therapies: 1. FDA Guidance for Industry: "Providing Clinical Evidence of Effectiveness" (May 1998) ("Evidence Guidance"); 2. FDAMA 115 "one adequate and well-controlled clinical investigation and confirmatory evidence"; and 3. Subpart H, 21 C.F.R. Part 314, or "accelerated approval" regulations ("Subpart H").

### A. Evidence Guidance and FDAMA 115

In its May 1998 Evidence Guidance, FDA describes nine<sup>3</sup> circumstances in which a single trial may meet the statutorily-required effectiveness evidence. Generally FDA had set a standard of requiring at least two adequate and well-controlled studies, following the language of the 1962 Amendments which used the plural "investigations" to describe the basic requirement for effectiveness. There had been times prior to 1998 when FDA had approved drugs based on a single study, especially when the AIDS crisis was just starting, but for most diseases the agency held drug approval to the "at least two" studies standard. The 1998 Evidence Guidance described circumstances in which a single study might be sufficient, such as where it may be unethical to conduct a second study and where the single study has a "statistically very persuasive finding" with other indicia of reliability, such as a multi-center trial with no single center dominating the results.

At the same time that FDA was developing its May 1998 guidance, Congress was enacting an amendment to the 1962 effectiveness standard that created a new alternative statutory standard for establishing a drug's effectiveness. This new alternative statutory standard is: "one adequate and well-controlled study and confirmatory evidence." This provision of the law is referred to as FDAMA 115 (after the section in the law called the FDA Modernization Act or FDAMA that established this alternate statutory standard for substantial evidence of effectiveness). The May 1998 Evidence Guidance and FDAMA 115 can be seen as qualitatively similar, in that both spoke to new ways of establishing substantial evidence of effectiveness, and both were issued almost simultaneously.

### B. Subpart H and Fast Track

The same 1997 law that created FDAMA 115 also created the statutory authority for "Fast Track" drugs, which is a modest elaboration by Congress of an FDA regulation known by its section of the drug regulations, Subpart H of 21 C.F.R. Part 314, or the so-called "accelerated approval" regulations (for biologics, the parallel regulation is at 21 C.F.R. Part 601, Subpart E). Both Fast Track and Subpart H are programs whereby a therapy for a serious or life-threatening disease for which there is no FDA-approved "available therapy" may be approved based either on an unvalidated surrogate that is reasonably likely to predict ultimate clinical outcome, or on an outcome other than irreversible morbidity or mortality. However, in such cases, there is also an additional post-approval requirement to conduct a study to establish the ultimate clinical outcome benefit, and if that study fails to do so, FDA may withdraw its approval as an expedited basis.

Subpart H represents a formal FDA system established to introduce an element of flexibility in executing FDA's responsibilities for ensuring that investigational therapies have adequately demonstrated their treatment benefit prior to marketing authorization. FDA created this system in response to the need of patients contracting HIV infections in the 1980's and the attendant public health crisis. This paper notes which orphan drugs were approved under Subpart H as well as which ones were designated as Fast Track therapies by way of a footnote in Figure 1.

## RESULTS

Figure 1 records the classification for each of the 135 non-cancer orphan therapies approved as new chemical entities from the enactment of the Orphan Drug Act in 1983 through June 30, 2010, with 45 classified as "Conventional", 32 as "Administrative Flexibility", and 58 as "Case-by-Case Flexibility". In Appendix 1, there is a narrative text that briefly describes the basis for each "case-by-case flexibility" classification, except for two therapies in that category: #75 Lanreotide Acetate and #116 Sodium Phenylbutyrate. In addition, there are textual comments about particular aspects of interest regarding eight other therapies: #6 Ambrisentan, #7 Amifostin, #8 Anagrelide, #35 Coagulation Factor IX, #58 Fosphenytoin, #59 Gallium, #60 Ganciclovir, and #79 Levomethadyl Acetate. Textual comments are included for these eight even though they are not classified as "case-by-case flexibility" in order to provide breadth of perspective and depth of understanding to the analytical processes employed. All of the therapies are listed alphabetically by chemical names.

<sup>3</sup> See Appendix 2 for more detailed discussion of how these nine types of single study approval examples apply to orphan drugs and of how FDAMA 115 relates to these.

## DISCUSSION

When asked how much evidence of safety and effectiveness an orphan drug must provide, FDA officials have generally explained to the Agency's public Advisory Committees, patient organizations, pharmaceutical companies and Wall Street that the Orphan Drug Act did not change the statutory requirements for establishing the safety and effectiveness of a proposed new medicine. For example, in a March 2010 FDA briefing document for the FDA Advisory Committee on an orphan drug, pirfenidone, being considered for patients with a rare, fatal pulmonary condition called idiopathic pulmonary fibrosis (IPF), FDA said:

In accord with our regulations, the Agency requires substantial evidence of effectiveness. Substantial evidence consists of adequate and well-controlled investigations on the basis of which it could be concluded that the drug will have the effect it is purported or labeled to have. The Agency usually requires more than one trial to provide independent substantiation of efficacy. Although IPF is an orphan disease, the requirements to establish effectiveness are not different, with the exception that the overall database may be smaller. We ask that you consider whether the results of PIPF-004 and PIPF-006 provide substantial evidence of efficacy to support the proposed indication to reduce decline in lung function in patients with IPF.<sup>4</sup>

This statement represents the most common FDA response when FDA is asked about the quantity and quality of effectiveness evidence required of an orphan drug. However, on some occasions, FDA has noted that it has the ability to be flexible within those statutory limits. For instance, the FDA briefing document for an Advisory Committee meeting on January 13, 2010 concerning the orphan drug Carbaglu (N-carbamylglutamate) for hyperammonemia had the following statement:

FDA has been flexible within the limits imposed by the congressional scheme, broadly interpreting the statutory requirements to the extent possible where the data on a particular drug were convincing. The Code of Federal Regulations (21 C.F.R. § 314.126) allows for studies without concurrent controls to be used to provide substantial evidence of effectiveness in diseases with high and predictable mortality, or in studies in which the effect of the drug is self-evident. Thus, the evidence obtained from retrospectively reviewed case studies could be considered as substantial evidence of effectiveness under those particular circumstances. The fact that the case series presented in this application is retrospective, un-blinded, and uncontrolled, precludes any meaningful formal statistical analyses of the data. Under these conditions, any statistical inference from confidence intervals and/or p-values is uninterpretable and, consequently, should not be utilized to inform clinical decision-making. To help frame the Committee's deliberations on whether the evidence standard in this application has been met, an FDA guidance document, '[Evidence Guidance]' is provided as background on the regulatory requirements for evidence of effectiveness.<sup>5</sup>

Thus, while the norm has been for FDA to respond simply

4 FDA Pulmonary-Allergy Drugs Advisory Committee Division Memorandum, Feb. 12, 2010, at pp. 15-16.

5 FDA Briefing Document, at pp. 9-10, attached to FDA Dr. Donna Griebel's December 16, 2009 memo to the Advisory Committee. See also June 23, 2010 statement of Deputy Commissioner Goodman to the U.S. Senate hearing cited in opening paragraphs of this paper.

that the statutory standard for effectiveness was not amended by the 1983 Orphan Drug Act, there have been ample occasions on which FDA has observed that it also has the legal authority and scientific right to be flexible in applying those statutory standards to orphan drug therapies.

There have been documents in which FDA has made abundantly clear its commitment to flexibility in applying the standard of safety and effectiveness, most notably during the AIDS crisis. In the mid-1980's, FDA promulgated Subpart E of the IND regulations for "drugs intended to treat life-threatening and severely-debilitating illnesses." FDA stated:

[The purpose of Subpart E is] to establish procedures designed to expedite the development, evaluation, and marketing of new therapies intended to treat persons with life-threatening and severely-debilitating illnesses, especially where no satisfactory alternative therapy exists. As stated [in section] 314.105(c) of this chapter, while the statutory standards of safety and effectiveness apply to all drugs, the many kinds of drugs that are subject to them, and the wide range of uses for those drugs, demand flexibility in applying the standards. The FDA has determined that it is appropriate to exercise *the broadest flexibility* in applying the statutory standards, while preserving appropriate guarantees for safety and effectiveness. *These procedures reflect the recognition that physicians and patients are generally willing to accept greater risks or side effects from products that treat life-threatening and severely-debilitating illnesses, than they would accept from products that treat less serious illnesses. These procedures also reflect the recognition that the benefits of the drug need to be evaluated in light of the severity of the disease being treated.*<sup>6</sup>

The regulation that FDA references in its Subpart E regulation is section 21 C.F.R. § 314.105(c) which predates the Subpart E regulation and illustrates again FDA's historic position on applying the same statutory standards in a flexible way depending upon the circumstances. According to 21 C.F.R. § 314.105(c):

FDA will approve an application after it determines that the drug meets the statutory standards for safety and effectiveness, manufacturing and controls, and labeling. While the statutory standards apply to all drugs, the many kinds of drugs that are subject to them and the wide range of uses for those drugs demand flexibility in applying the standards. Thus FDA is required to exercise its scientific judgment to determine the kind and quantity of data and information an applicant is required to provide for a particular drug to meet them. FDA makes its views on drugs products and classes of drugs available through guidelines, recommendations and statements of policy.

An example of a formal regulatory policy or guidance that expresses this concept of "flexibility" in FDA's application of the statutory standards of safety and efficacy is seen in the International Conference on Harmonisation (ICH) of the Technical Requirements for Registration of Pharmaceuticals for Human Use E1A guidance. This FDA-adopted international guidance stipulates the minimum quantum of safety exposures necessary for FDA to even accept a marketing application for review when the medicine is intended for a chronic

6 21 C.F.R. Part 312, Subpart E, emphasis added.

condition.<sup>7</sup> Most rare disorders are chronic in nature and not acute, and so this guidance applies to most rare disorder therapies. The guidance states that the minimum number of safety exposures to meet the statutory standard for safety are 1,500 persons exposed to the investigational therapy with 300 to 600 of those exposed for at least 6 months and with at least 100 exposed for one year. However, the guidance states that these minimum safety thresholds do not apply to therapies for rare disorders. Importantly, the guidance then does NOT state what is required in the alternative, whereas it could have stated an algorithm such as at least 1% of the U.S. population with the rare disease must be exposed with half of them for at least one year. Rather, the guidance relies upon the exercise of FDA's scientific judgment to determine what is appropriate to meet the statutory standard for safety in each particular rare disorder therapy.

In other areas FDA can exercise similar flexibility. For instance, where the potential number of subjects is limited, the degree to which FDA demands dose selection to be optimized in pre-approval studies may be reduced, as can FDA's requirements for validation of a patient-reported outcome instrument in a rare disorder population or proof of the sensitivity, specificity and clinical meaningfulness of a primary endpoint.

NORD has requested that FDA issue a formal policy statement on FDA's regulation of therapies for persons with rare diseases (see footnote 9). Given that each investigational therapy for a rare disorder will present unique features, NORD understands that the granularity of the requested statement of policy may necessarily be limited. However, even cataloging the nature and scope of the orphan drug precedents that illustrate FDA's flexibility may enable key stakeholders to better understand FDA's position. That is, even while FDA states correctly that the statutory standards are the same for prevalent and rare conditions, FDA should develop and issue a formal companion statement of the equally important and consistent FDA historic position that FDA will exercise its scientific judgment to interpret and apply those statutory standards in a flexible manner, tailored to the circumstances of each investigational therapy for each rare disease and disorder.

It is this cataloguing of orphan drug precedents that is the chief purpose of this analysis and paper. This review of FDA actions on rare disorder therapy marketing applications concludes that two of every three orphan drugs approved manifests FDA's historic flexibility in applying to therapies for rare disorders the statutory standard for establishing effectiveness. By this classification, 32 of the 135 orphan drugs analyzed reflect administrative flexibility, that is, FDA application of statutes and FDA regulations and guidance documents to those particular orphan therapies, and another 58 orphan therapies were approved on a case-by-case application of flexibility.

There is an element of subjectivity and judgment in making

<sup>7</sup> Note that this paper consists of an analysis only of the quantum of evidence of effectiveness information determined to be adequate by FDA to support an approval, and the FDA-adopted ICH guidance that is the subject of this paragraph refers to a formal expression by FDA of its flexibility with respect to the quantum of safety information required for orphan drug therapies.

these classifications. NORD does not have access to non-public information, which both FDA and the sponsors have. It is therefore possible that FDA and drug manufacturers will disagree about into which one of these three categories any therapy may be classified.<sup>8</sup>

However, NORD believes that the overall thrust of the findings of this analysis is immovable – that FDA's approval actions on a considerable portion of therapies for those patients afflicted with rare disorders demonstrate a consistently applied flexibility in assessing the effectiveness of such therapies.

Ironically and unfortunately, there has not been any statement from FDA as to how that flexibility finds expression. At the first FDA public hearing on orphan drugs which was held on June 29 & 30, 2010, NORD called on FDA to issue a "clear, granular expression of FDA's historic commitment to exercise flexibility in its review of therapies for rare disorders."<sup>9</sup>

## CONCLUSION

Research resources in the universe of rare disorders are precious, with the most precious being the persons with the rare disorders who heroically volunteer to participate in a trial, usually under conditions where there is less known than in trials of therapies for prevalent diseases about the safety and potential effectiveness of the investigational therapy from animal models, animal toxicology and early human trials. So, when these trials are conducted, sometimes with designs with which all parties may not be in full concurrence, including FDA, great deference should be afforded to the design of these trials and flexibility applied in the interpretation of their results. If such a principle were to be addressed and accepted by FDA, much good would come of it.

In the more than 28 years since its enactment, the Orphan Drug Act has proven a resounding success. This is best seen in the 357 new medicines for more than 200 different rare disorders approved by FDA over the first quarter of a century of the law's existence.

NORD believes that this study's confirmation of FDA's flexibility in reviewing applications for orphan drugs reinforces the need for its public acknowledgement that orphan drugs are indeed meritorious of special consideration. Such a statement by the FDA would provide the impetus for greater attention to orphan drug therapies within the academic community as well as within the drug development and investment communities.

With health care reform measures inevitably changing how medicine is practiced and how patients are treated and reimbursed, the need for such attention to the rare disease community is especially critical. Patients with rare diseases can easily be left behind during this transitional period. FDA has demon-

<sup>8</sup> A further cautionary note is that every drug approval, whether for a rare condition or a common one, stands on a unique set of empiric evidence judged against a backdrop of specific scientific and clinical considerations in light of the relative degree of the medical needs of that particular set of patients. Therefore, caution must be exercised in any attempt to extrapolate from any one or more of these case studies to current or future therapies in development or under FDA review.

<sup>9</sup> Statement of NORD, presented by Chairman Frank Sasinowski (June 29, 2010).

strated in its review of orphan products that it recognizes the importance of therapies for persons with rare disorders. It is time for that policy to be clearly enunciated as a formal FDA policy, and for FDA medical reviewers to incorporate and recognize this flexibility in a systematic way into their evaluations of each new therapy in development and under FDA review for Americans with any rare disease. Much that is very good for all persons with rare disorders could come of this.

NORD exhorts FDA to continue to embrace even more fully the historic flexibility it has long noted and exercised in FDA's regulation of medicines for those Americans with rare disorders.<sup>10</sup>

---

<sup>10</sup> Author's note: The author commends FDA on its stellar, worldwide leadership on critical matters affecting persons with rare diseases for the past 28 years, and exhorts FDA to continue to embrace even more fully the historic flexibility FDA has long noted and exercised in FDA's regulation of medicines for those Americans with rare diseases. In the over 28 years since its enactment, the Orphan Drug Act has proven a resounding success. This is best seen in the over 350 new medicines for more than 200 different rare disorders approved by FDA. However, there are still about 6,800 disorders for which there is not one FDA-approved therapy. Perhaps most discouraging is that many affected with rare disorders do not even see any research being conducted on their conditions. It seems as though the proverbial low-hanging fruit has been harvested in the first quarter of a century of the law's existence, while the vast majority of persons with rare diseases see only that there is no medicine within their reach, and sometimes even within the reach of reasonable hope. In sum, much has been accomplished by FDA, by NIH, by medical and scientific researchers, by the pharmaceutical industry, by the financial community and by patient advocates in these first 28 years, but much more beckons each of us to respond to the needs of those with rare diseases. The author's heartfelt hope is that this analysis helps to advance the development of those medicines to aid all in need of them.

FIGURE 1

Chemical and Brand Names		Approval (mm/yy)	Type of Efficacy Evidence:		
			Conventional	Administrative Flexibility	Case-by-case Flexibility
1	Agalsidase beta - Fabrazyme <sup>1</sup>	04/2003		X	
2	Albendazole - Albenza	06/1996			X
3	Alglucerase - Ceredase	04/1991			X
4	Alitretinoin - Panretin	02/1999			X
5	Alpha1-Proteinase Inhibitor (Human) - Prolastin	12/1987			X
6	Ambrisentan - Letairis <sup>2</sup>	06/2007	X		
7	Amifostine - Ethyol <sup>1</sup>	12/1995		X	
8	Anagrelide HCl - Agrylin	03/1997	X		
9	Antihemophilic Factor Recombinant - Kogenate	02/1993		X	
10	Antithrombin III (Human) - Thrombate III	12/1991		X	
11	Antithrombin III (Human) - Atnativ	12/1989		X	
12	Antivenin, Crotalidae Polyvalent Immune Fab (Ovine) - CroFab	10/2000	X		
13	Apomorphine HCl - Apokyn <sup>3</sup>	04/2004	X		
14	Aprotinin - Trasylol	12/1993	X		
15	Artemether/Lumefantrine - Coartem <sup>3</sup>	04/2009			X
16	Atovaquone - Mepron	11/1992	X		
17	Azacitidine - Vidaza <sup>3</sup>	05/2004		X	
18	Basiliximab - Simulect	05/1998	X		
19	Sodium Benzoate and Sodium Phenylacetate - Ucephan	12/1987			X
20	Beractant - Survanta	07/1991	X		
21	Betaine HCl - Cystadane	10/1996			X
22	Bosentan - Tracleer <sup>2</sup>	11/2001	X		
23	Botulinum Toxin Type A - Botox, Botox Cosmetic	12/1991	X		
24	Botulinum Toxin Type B - Myobloc	12/2000	X		
25	Botulism Immune Globulin Intravenous (Human) - BabyBIG	10/2003		X	
26	C1 Esterase Inhibitor (Human) - Cinryze	10/2008		X	
27	C1 Esterase Inhibitor (Human) - Berinert P	10/2009		X	
28	Calcitonin-Human for Injection - Cibacalcin	10/1986	X		
29	Canakinumab - Ilaris	06/2009		X	
30	Capsaicin - Qutenza	11/2009	X		
31	Carglumic acid - Carbaglu	03/2010		X	
32	Chenodiol - Chenix	07/1983			X
33	Cinacalcet HCl - Sensipar	03/2004			X
34	Clofazimine - Lamprene	12/1986	X		
35	Coagulation Factor IX (Recombinant) - Benefix	02/1997		X	
36	Coagulation Factor IX - Mononine	08/1992		X	
37	Coagulation Factor IX (Human) - Alphanine	12/1990		X	
38	Coagulation Factor VIIa (Recombinant) - NovoSeven	03/1999			X
39	Colfosceril Palmitate, Cetyl Alcohol, Tyloxapol - Exosurf Neonatal for Intrathecal Suspension	08/1990	X		
40	Collagenase clostridium histolyticum) - Xiaflex	02/2010	X		
41	Corticotropin Ovine Triflutate - Acthrel	05/1996			X
42	Cysteamine Bitartrate - Cystagon	08/1994			X
43	Cytomegalovirus Immune Globulin (Human) - Cytogam	12/1998	X		
44	Daclizumab - Zenapax	12/1997	X		
45	Dalfampridine - Ampyra	01/2010	X		
46	Deferasirox - Exjade <sup>1,3</sup>	11/2005			X
47	Dexrazoxane HCl - Zinecard <sup>1</sup>	05/1995			X
48	Diethylenetriamine pentaacetic acid (Dtpa) - none	08/2004			X
49	Digoxin Immune Fab (Ovine) - Digibind	04/1986	X		
50	Dornase Alfa - Pulmozyme	12/1993		X	
51	Ecallantide - Kalbitor	11/2009	X		
52	Eculizumab - Soliris	03/2010	X		
53	Eflornithine HCl - Ornidyl	11/1990	X		

1-Subpart H for Efficacy; 2-Subpart H for Safety; 3-Fast Track

Quantum of Effectiveness Evidence in FDA's Approval of Orphan Drugs

		<b>Type of Efficacy Evidence:</b>		
		Conventional	Administrative Flexibility	Case-by-case Flexibility
Chemical and Brand Names	Approval (mm/yy)			
54	Eltrombopag Olamine - Promacta	X		
55	Ethanolamine Oleate - Ethamolin			X
56	Felbamate - Felbatol	X		
57	Fomepizole - Antizol			X
58	Fosphenytoin Sodium - Cerebyx		X	
59	Gallium Nitrate - Ganite		X	
60	Ganciclovir Sodium - Cytovene		X	
61	Glatiramer Acetate - Copaxone		X	
62	Glutamine - Nutrestore		X	
63	Halofantrine HCl - Halfan	X		
64	Hemin - Panhematin			X
65	Histrelin Acetate - Supprelin	X		
66	Human Fibrinogen Concentrate, Pasteurized - RiaSTAP		X	
67	Icodextrin Peritoneal Dialysis Solution - Extraneal		X	
68	Idursulfase - Elaprase		X	
69	Iloprost - Ventavis		X	
70	Imiglucerase - Cerezyme			X
71	Interferon Beta-1a - Avonex			X
72	Interferon Beta-1b - Betaseron <sup>1</sup>		X	
73	Interferon Gamma 1B - Actimmune		X	
74	Ferric Hexacyanoferrate(II) - Radiogardase			X
75	Lanreotide Acetate - Somatuline Depot			X
76	Laronidase - Aldurazyme			X
77	Lenalidomide - Revlimid <sup>2,3</sup>			X
78	Lepirudin - Refludan			X
79	Levomethadyl Acetate HCl - Orlaam		X	
80	Lodoxamide Tromethamine - Alomide	X		
81	Mecasermin Rinfabate Recombinant - Iplex			X
82	Mecasermin Recombinant - Increlex			X
83	Mesna - Mesnex		X	
84	Midodrine HCl - Proamatine (Amatine) <sup>1</sup>			X
85	Miglustat - Zavesca <sup>3</sup>			X
86	Modafinil - Provigil	X		
87	Monoctanoin - Moctanin			X
88	N-acetylgalactosamine 4-sulphatase, Recombinant Human [Galsulfase] - Naglazyme			X
89	Nitazoxanide - Alinia	X		
90	Nitisinone - Orfadin <sup>3</sup>		X	
91	Nitric Oxide - Inomax	X		
92	Oprelvekin - Neumega			X
93	Pegademase Bovine - Adagen			X
94	Pegvisomant - Somavert <sup>3</sup>		X	
95	Pentamidine Isethionate - Pentam	X		
96	Pentastarch - Pentaspan	X		
97	Pentosan Polysulfate Sodium - Elmiron	X		
98	Plerixafor - Mozobil	X		
99	Protein C Concentrate - Ceprotin			X
100	Rasburicase - Elitek			X
101	Recombinant Human Acid Alpha-Glucosidase [Alglucosidase ALFA] - Myozyme			X
102	Recombinant Human Antithrombin - ATryn			X
103	Respiratory Syncytial Virus Immune Globulin (Human) - Respigam			X
104	Rho (D) Immune Globulin Intravenous (Human) - Winrho SD			X

1-Subpart H for Efficacy; 2-Subpart H for Safety; 3-Fast Track



Quantum of Effectiveness Evidence in FDA's Approval of Orphan Drugs

Chemical and Brand Names	Approval (mm/yy)	Type of Efficacy Evidence:		
		Conventional	Administrative Flexibility	Case-by-case Flexibility
105 Rifabutin - Mycobutin	12/1992	X		
106 Rifapentine - Priftin <sup>1</sup>	06/1998			X
107 Riloncept - Arcalyst	02/2008			X
108 Riluzole - Rilutek	12/1995			X
109 Romiplostim - Nplate	08/2008	X		
110 Rifinamide - Banzel	11/2008			X
111 Sacrosidase - Sucraid	04/1998			X
112 Sapropterin Dihydrochloride - Kuvan	12/2007			X
113 Sargramostim - Leukine	03/1991	X		
114 Selegiline HCl - Eldepryl	06/1989	X		
115 Sodium oxybate - Xyrem <sup>2</sup>	07/2002	X		
116 Sodium Phenylbutyrate - Buphenyl	04/1996			X
117 Somatrem - Protopin	10/1985	X		
118 Sotalol HCl - Betapace	10/1992	X		
119 Sterile Talc Powder - Sclerosol	12/1997			X
120 Succimer - Chemet	01/1991	X		
121 Teriparatide Acetate - Parathar	12/1987		X	
122 Tetrabenazine - Xenazine	08/2008			X
123 Thalidomide - Thalomid <sup>2</sup>	07/1998			X
124 Tiopronin - Thiola	08/1988			X
125 Tranexamic Acid - Cyklokapron	12/1986			X
126 Treprostinil sodium - Remodulin <sup>1</sup>	05/2002			X
127 Trientine HCl - Syprine	11/1985			X
128 Trimetrexate Glucuronate - Neutrexin	12/1993			X
129 Vaccinia Immune Globulin (Human) Intravenous - N/A	02/2005			X
130 Velaglucerase alfa - Vpriv	02/2010	X		
131 Vigabatrin - Sabril	08/2009			X
132 von Willebrand Factor/Coagulation Factor VIII Complex (Human) - Wilate	12/2009			X
133 Zalcitabine - Hivid <sup>1</sup>	06/1992			X
134 Zidovudine - Retrovir	03/1987		X	
135 Zoledronic Acid - Zometa	08/2001	X		
<b>Sub Totals:</b>		<b>45</b>	<b>32</b>	<b>58</b>
<b>Flexibility:</b>		<b>Not Needed: 45</b>	<b>Yes: 90</b>	

1-Subpart H for Efficacy; 2-Subpart H for Safety; 3-Fast Track

## APPENDIX 1

This Appendix provides commentary on the basis for approval only for products categorized as “case-by-case flexibility.” The Appendix is keyed to the product numbering system in Figure 1 (thus, the Appendix starts with the second drug listed because there is no commentary on the first drug listed).

## 2. Albendazole - Albenza

The 1996 approval for this antihelminthic drug for treating infectious diseases caused by pork tapeworms and by dog tapeworms was based by FDA on:

1. a single study that was either a well-controlled study (Medical Review) or not well-controlled (Statistical Review, 28);
2. supporting literature (which was not all positive, one October 1995 paper in the *Annals of Internal Medicine* concluded from this study of 138 subjects over 2 years that “previous reports of favorable response to treatment of necrilocysticercosis (pork tapeworm) with...albendazole are by no means definitive and may be a reflection of the natural history of the condition”);
3. compassionate use information; and
4. existing approvals in Australia, The Netherlands, Germany, United Kingdom, South Africa, India, Japan, and Spain.

The one study characterized by the medical reviewer as “well-controlled” was the Gelman study in Peru which compared 55 subjects in those with pork tapeworm disease with approximately half of the subjects randomized either to 7 days or 14 days of Albendazole. After 90 days there was no difference between the two groups ( $p = 1.00$ , Medical Review, 42<sup>11</sup>), but at one year the group on 7 days of therapy had a statistically significant greater reduction in cysts (primary endpoint) than did the group on 14 days of therapy ( $p = 0.037$ , Medical Review, 42 .) (The statistical review had concluded that “there were no formal statistical analysis of the clinical data in the NDA, [and] only descriptive statistics were used.” Statistical Review, 28). The medical review cited multiple deficiencies in the Gelman study when it was audited (Medical Review, 98). Overall, the medical review catalogued the litany of clinical factors which hindered this regulatory review: (1) there is little or no such disease in the U.S.; (2) The natural history of the disease is not completely understood; (3) There is a lack of gold standard for diagnosis; (4) There is a lack of reliable clinical endpoints; and (5) need for long term follow-up.<sup>12</sup>

The statistical review concluded, with respect to the therapy for dog tapeworm, that “due to the very limited data available...the statistical conclusion toward the efficacy...of [albendazole] can not be reached” (Statistical Review, 30) and

11 This between group difference, even if statistically significant, hardly lends considerable weight to the biological plausibility that the drug “works” in that the group dosed for 7 days fared better than the group that was dosed for 14 days.

12 This list of factors is noteworthy because, although articulated by this FDA reviewer over 15 years ago, these same factors apply to many, many orphan diseases, yesterday, today and likely tomorrow's well.

with respect to the therapy for pork tapeworm, “the results do not sufficiently provide comprehensive evidence to confirm [albendazole] as an effective...medicine...due to the weakness of the nature of these studies. Upon considering the particularity of [albendazole] for orphan drug status, the reviewer does not preclude to endorse this application and regulatory actions will be adopted after soliciting for standpoints of clinicians” (Statistical Review, 30-31). In the medical officer's concluding statement of factors that were considered in arriving at the approval recommendation it was noted that albendazole “qualifies for orphan drug designation” (Statistical Review, 31).

## 3. Alglucerase - Ceredase

In the April 1991 approval of this lysosomal enzyme, which is deficient in those with Gaucher's disease, the medical group leader noted, as the first of three issues to be considered in approving this drug, that “no well-controlled studies were conducted” (Medical Group Leader Memo, Dec. 26, 1990, 1). She went on to explain that, to her, there were 2 studies that demonstrated efficacy. One was a study in which liver biopsies were conducted before and 44 hours after a single infusion in 22 subjects. The other (seemingly more convincing) study was a 6 month study that compared 2 groups of 12 subjects on drug and 12 not on drug (or placebo). The subjects were not randomized and there were major differences in key baseline prognostic variables. Therefore, the most compelling data from the study were the change from baseline to end of study in the 12 subjects on drug in the key clinical parameters of anemia and spleen and liver organomegaly. The Medical Group Leader concluded that “this was convincing evidence of efficacy and because of the rarity of patients and the difficulty of following placebo-treated or untreated patients with severe disease for long periods of time, randomized studies were not required.” (Medical Group Leader Memo, 1). This approval illustrated FDA's ability to exercise scientific judgment as well as to extend itself in aiding a sponsor with compiling the NDA as the FDA medical reviewer noted in his reviews.<sup>13 14</sup>

13 NORDD considered whether to classify this application as meeting the May 1998 clinical evidence of effectiveness standard for a single study because in that Guidance, FDA explains that “a single clearly positive trial can be sufficient to support approval of a replacement therapy...when it is combined with clear evidence that the condition being treated is caused by a deficiency of that factor. Demonstration of physical replacement of the deficient factor...provides strong substantiation of the clinical effect.” However, in this case, there is no “single clearly positive trial.” Evidence Guidance, 11.

14 This study could also be seen as a historically controlled study. FDA regulations recognize a historically controlled study as one of 5 enumerated types of “adequate and well-controlled studies”, 21 C.F.R. § 314.126(b)(2)(v). Use of a patient as his/her own control is a variant of the historically controlled study model. However, FDA in its regulations, notes that historically controlled studies should be reserved for “special circumstances” because pertinent variables can not be controlled and such special circumstances include where the effect of the drug is “self-evident”; however, here the effect of the drug was measured on organ volume and anemia which are less “self-evident” as caused by the investigational drug than examples given in FDA's regulation such as general anesthetics.

#### 4. Alitretinoin - Panretin

In this February 1998 approval for the treatment of AIDS-related Kaposi's Sarcoma (KS), FDA found that one of two Phase 3 trials was clearly positive but the second Phase 3 trial was stopped early and as the November 17, 1998 memo of the statistical reviewer concluded, "doubts remain as to the appropriateness of the interim analysis and robustness of the response rates in the trial that was stopped early." As noted earlier in that same review, "if the [FDA] medical reviewer's assessment would have been used as evidence for stopping the trial early, the trial would not have been stopped and one would conclude that there was no statistically significant difference between the arms" (Statistical Review, 14). In addition, the secondary endpoints measuring various "time to event" outcomes did not show numerically different results in the two treatment arms, except that in one of these secondary endpoints (median time to progression) the placebo arm results were much better (that is, took much longer for subjects to progress on placebo than on drug). The FDA reviewer noted that "this is somewhat unexpected considering the superiority of response rate in the sponsor's assessment" (Statistical Review, 13).<sup>15</sup>

#### 5. Alpha1 – Proteinase Inhibitor (Human) - Prolastin

In this December 1987 approval of a replacement protein for those who are genetically deficient in alpha-1-antitrypsin, FDA in the approved labeling cites one uncontrolled study of 19 subjects, all with the same phenotypic variant of this deficiency, the most severely affected variant of which there are many variants. In that study, there was within a few weeks a change from baseline reported in two measures, alpha-1-antitrypsin levels and antineutrophil elastase capacity, as ascertained by bronchoalveolar lavage. However, FDA also notes that the disease manifests itself as emphysema in the third or fourth decade of life but that the "pathogenesis of development of emphysema in alpha-1-antitrypsin deficiency is not well understood at this time." (Label, 1).

This approval clearly demonstrates the exercise of scientific judgment by FDA. The May 1998 FDA Guidance speaks to a single study sometimes being sufficient to support approval of a replacement therapy "when the pathophysiology of a disease and the mechanism of action of a therapy are very well understood." (Evidence Guidance, 11). Therefore, while NORD classifies this approval as "case-by-case" flexibility, if one were to conclude that the conditions of the May 1998 FDA Guidance had been met, then the classification instead would be "administrative" flexibility, which is evidence of FDA flexibility nonetheless.

#### 6. Ambrisentan - Letairis

This June 2007 approval of a drug for pulmonary hypertension was approved under 21 C.F.R. Part 314, Subpart H. However,

<sup>15</sup> To NORD, there are not two adequate and well-controlled studies clearly positive in this case, so this approval shows an exercise of scientific judgment. Also, while this indication is for cancer, this approval was closely followed and seen by the AIDS patient community as an FDA action related to AIDS, more than for cancer, and so this approval has been included in this analysis.

it was not approved under Subpart H for reasons related to its evidence of efficacy, such as its registration studies having been conducted using an unvalidated surrogate as their primary endpoint. Instead, this was approved under Subpart H restrictions on distribution for safety concerns. NORD surmises that this will be the last drug ever so approved because several months later, the Food and Drug Administration Amendments Act granted FDA authority to impose a REMS as part of marketing approval and as part of the REMS to include in some cases, Elements to Assume Safe Use (ETASU). Therefore, NORD believes that drugs formerly approved under Subpart H with safety restrictions such as Actiq, Thalomid, and Bosentan will in the future be approved with a REMS that includes ETASU. NORD includes this discussion here to illustrate a point made earlier in this paper that can otherwise be confusing and that is, Figure 1 to this paper includes footnotes that denote each drug approved under Subpart H for efficacy reasons, under Subpart H for safety concerns and under Fast Track. NORD thought it critical to include this discussion of Ambrisentan so that the diligent reader who checks on all the Subpart H orphan drug approvals and discovers some NORD would not have otherwise included because they are Subpart H orphan drugs but only because of safety concerns can now understand the reason for the apparent discrepancy.

#### 7. Amifostine - Ethyol

This drug was approved in December 1995 and illustrates all the principles that would later be articulated by FDA in its "single study" with a "statistically very persuasive finding" and where another study is likely unethical. (Evidence Guidance, 12-16). Therefore, NORD classifies this as a case of "administrative" flexibility even though this approval predated the issuance of FDA's May 1998 Guidance.

#### 8. Anagrelide HCl - Agrylin

In this March 1997 approval for treating essential thrombocytopenia, the approved indication was "to reduce the elevated platelet count and the risk of thrombosis and to ameliorate associated symptoms." The FDA-approved labeling refers to two "historically controlled, unblinded" studies in a total of about 300 subjects. The statistical review states that these two trials were both Phase 2 open-label trials that were "patient controlled," and baseline-controlled or "patient as own control" (Statistical Review, 2), which, as discussed earlier, are a form of historical control. The statistical analysis review supports this by only describing changes in each subject from that subject's baseline platelet count (without any reference to any natural history control group). While the statistical review mentions that associated symptoms were a secondary endpoint in the larger of these two Phase 2 trials, the review never mentions any results of that analysis of symptoms in its memo and moreover, there is no mention at all in either study that risk of thrombosis was assessed as measured by thrombosis events or any endpoint or instrument. (The medical reviewer's memo for efficacy is not publicly available.) While this drug's approval may illustrate some exercise of scientific judgment, NORD classifies this approval as having met the 2

adequate and well-controlled study efficacy standard or “conventional” approval.

This approval and text of its analysis is included to alert the readers that there are many cases that were classified as “conventional” which also have elements of flexibility. NORD anticipates that readers reviewing any one of these approvals classified as “case-by-case” flexibility may come to a different conclusion if asked to adjudicate that case. Therefore, for comprehensiveness, any reader would also have to re-adjudicate each of the 34 cases classified as “administrative flexibility” and 46 cases classified as “conventional” approvals in order to score all the cases, and the text of these analyses is not presented but for one or two exceptions such as this one for illustrative purposes. NORD appreciates that there is subjectivity in making these judgment calls, but the overall “gestalt” is clear.

### 15. Artemether/Lumefantrine - Coartem

In this April 2009 approval of this fixed-dose combination product for treating malaria, FDA medical and statistical reviewers both noted that there were only two trials that tested the combination against the single components and both studies were conducted at the same single center in China with a single racial group. Therefore, the statistical review of August 21, 2008 questions whether study results can be extrapolated beyond this region and this ethnic group. (Statistical Review, 11). The FDA medical review of November 25, 2008 states that the 2007 FDA draft Malaria Guidance recommends that the primary endpoint be 28-day cure as defined by FDA. (Medical Review, 34). However, the statistical review explains that “evaluation of FDA-defined cure rate is not possible [in these 2 studies] due to lack of information on clinical signs and symptoms as well as malaria-related laboratory abnormalities from the sponsor.” (Statistical Review, 8).

The key finding here is that on the primary endpoint of 28 day cure rate, even without being able to employ the FDA defined cure rate, the combination failed to beat lumefantrine: in one study the p value for this comparison was 0.49 and in the other study, there were two comparisons of the combination to the lumefantrine component because the study had both a lumefantrine capsule arm (p = 0.675) and a lumefantrine tablet arm (p = 0.16). However, there were other non-primary endpoints that showed the value of the combination over both monotherapy components. Therefore, this approval required an exercise of scientific judgment.

### 19. Sodium Benzoate/Sodium Phenylacetate - Ucephan

In this December 1987 approval to treat urea cycle disorder, FDA demonstrated flexibility in that the March 20, 1986 medical review states that about 80% of subjects on this therapy in a study of 56 subjects in 45 sites survived compared to about a 15% survival rate historically for persons on dietary modification alone. The medical review concludes by noting that, “The usual requirements for a statistical evidence of efficacy though not fulfilled, the volume of data accrued over

almost 6 years in a multicenter study appear reasonably adequate.” (Medical Review, 43-44). Furthermore, the review’s final paragraph before its approval recommendation notes that this drug has been designated as an orphan drug.<sup>16</sup>

### 21. Betaine HCl - Cystadane

In this November 1996 approval to treat an inborn error of metabolism, homocystinuria, FDA exhibited enlightened exercise of scientific judgment in that all data were drawn from published literature and there was only one randomized, double-blind placebo-controlled trial and it failed. This study looked at the effect of vertebral bone density. This six patient trial was one year in length. According to the medical review of June 19, 1996, “results showed that bone density measurements determined after 6 and 12 months Bentaine prescription did not differ from those after 6 and 12 months of placebo.” (Medical Review, 9). However, in the other 17 published trials including 78 patients, the sponsor concluded that homocysteine levels were reduced by bentaine, and 48 of these 78 patients also reported some clinical response in addition to the biochemical response. The medical reviewer’s first observation under “Discussion and Conclusions” was the following:

This NDA is generally in poor condition, and the sponsor has made relatively poor use even of the published articles.... There are several reasons for such limited exertion. First, the company is a relatively newer entity and has had little previous experience with drug development; this is in fact the first NDA it has ever submitted to FDA. Other and larger companies show little or no interest in submission of an NDA for this drug in this disorder after [the FDA Office of Orphan Products Development] inquired after a sponsor for the product. Additionally, the disorder for which this new treatment is to be indicated was described only within the past four decades, and it is rare. Homocystinuria...has been estimated that only 800-1,000 cases in total have been found and reported in the United States. It is obvious that this company was not willing and/or able to spend much on original work in homocystinuria; it has depended entirely upon knowledge already in the medical literature.

(Medical Review, 14).<sup>17</sup>

### 32. Chenodiol – Chenix

In this July 1983 approval for treating certain gallstones in patients at increased surgical risk, the preponderance of the clinical experience came from a placebo-controlled Natural Co-operative Gallstone Study (NCGS) of 916 subjects who were not at high surgical risk, and that studied two lower doses of this drug than the doses approved. The dose range approved

16 This case could alternatively be considered for classification as “administrative” rather than “case-by-case” flexibility but flexibility nevertheless.

17 To NORD, this seems to have been a candidate for Subpart H approval (because reduction of homocysteine levels would seem to be an unvalidated surrogate that would be reasonably likely to predict ultimate clinical outcome) but this drug was not classified by FDA as a Subpart H drug. However, the difficulty of conducting a confirmatory Phase 4 trial may have been a factor (although following subjects on chronic administration of drug and comparing their outcomes to natural history/historical controls could have been explored and maybe it was.)

came from several uncontrolled studies. Since there are no publicly available medical or statistical reviews (and the drug has been discontinued from marketing), the label at the time of approval is the only available source of information on the efficacy evidence in the approved orphan population of subjects at high surgical risk and the label has only the following single sentence about one study considered in that population: "In a prospective trial using 15 mg/kg/day, 31% enrolled surgical risk patients treated more than 6 months (n=86) achieved complete confirmed dissolutions." (Label, 1).

Given that most of the discussion in the labeling is of the 916 subject NCGS that was in a different type of subject and at different doses than those approved, and given that the other clinical data are several uncontrolled studies, none of which appear to be restricted to high surgical risk patients, FDA appears to have extrapolated from these clinical data set to the dose approved in the high surgical risk orphan patient population. Therefore, this seems to exhibit an exercise of some modicum of scientific judgment, although this classification necessarily has to be tentative given the lack of medical and statistical reviews in this case.<sup>18</sup>

### 33. Cinacalcet HCl - Sensipar

In this March 2004 approval for treating hypercalcemia in patients with parathyroid carcinoma, the data on patients with the orphan condition came from a Phase 2, open-label study of ten subjects. (However, there was ample clinical data from three randomized, double-blind placebo controlled trials in chronic kidney disease (CKD) patients with secondary hyperparathyroidism with about 1200 total subjects enrolled, and this clinical evidence from a prevalent disease [likely showing the drug's ability to reduce serum calcium] may have played a significant role in FDA's consideration of the orphan condition.) The primary endpoint of the Phase 2 study was a reduction in the two-to-sixteen week titration phase in serum calcium of 1.0mg/dl or more and seven of the ten subjects met this, but the medical review of February 14, 2004 went on to note that: "None of the patients, however, normalized their serum calcium levels." (Medical Review, 18). This review of the efficacy evidence for the orphan condition concludes: "To state the obvious, the data upon which Amgen is requesting approval for the treatment of parathyroid carcinomas are very limited. Yet, parathyroid carcinoma is a rare disease and patients have few treatment options for the hypercalcemia associated with the condition. Cinacalcet offers the potential to satisfy an unmet medical need in this population of seriously ill patients." (Medical Review, 18-19, emphasis added).<sup>19</sup>

### 35. Coagulation Factor IX (Recombinant) - Benefix

In this February 1997 approval for treating hemophilia B,

there were two studies that evaluated clinical results in the Summary Basis of Approval. In one study of 37 subjects who had been previously treated with moderate to severe hemophilia, 82% of all bleeding episodes in the peri-operative period required only one infusion for resolution. In the second study, the drug was administered for 13 procedures in 12 subjects. Ninety-seven percent of clinical responses were rated subjectively by the physician or patients as excellent or good, and transfusion of blood products was needed in only three of the 13 surgical procedures with hemostasis maintained throughout the surgical period without any clinical evidence of thrombotic complications.

Since there was no discussion of a historical control group nor of the prior experiences of any of the subjects in either trial (therefore, no analysis could be made using each patient as his/her own control), there was, in NORD's opinion, some exercise of scientific judgment in this approval. NORD would have classified this as "case-by-case flexibility" had NORD not consulted with a senior FDA hematologist on this and most of the other orphan blood disorder biologic approvals (see also # 9, 10, 11, 36, 37, 38, 99, 102, 104 & 132). This set of blood disorder biologic approvals is, to NORD, the most opaque in terms of understanding whether the quantum of efficacy evidence would have been sufficient for approval even if these disorders were prevalent and not rare and if these approvals required any exercise of FDA administrative or case-by-case flexibility. NORD includes this one only to illustrate the value of the insights provided by the FDA official. In this case, the FDA official explained that while NORD's catalogue of the evidence is accurate, the conclusion is wrong because, to the FDA official, even if one million Americans had the condition, this quantum of evidence would have been adequate for approval. The official explained that this therapy simply replaces a protein that is missing and that replacing the missing protein by giving one unit of this product will predictably raise blood levels of that protein by a certain amount and it is well established in hematology what the blood levels of that protein should be for surgery, for satisfactory hemostasis and for other situations. (See also #36 Mononine and #37 Alphanine, which are plasma-derived and for which this same paradigm applies.) However, there are two issues to note: one, Benefix was the first recombinant Factor IX, and therefore, there were safety issues that needed to be addressed such as immunogenicity concerns, and two, this approval was in 1997 and FDA likely would hold a new Factor IX product today to a more demanding efficacy requirement that includes a demonstration of the drug's effect in surgery and on the treatment of bleeding generally. But, as of 1997, the quantum of efficacy evidence provided with this application was not only sufficient for approval for this orphan condition but would have been adequate even if the condition had been prevalent. Therefore, the classification here is "administrative flexibil-

<sup>18</sup> It is to be noted that the lack of medical and statistical reviews is nearly unique to this drug in this analysis of non-cancer orphan drug approvals. Also, it is worth noting that this drug was designated as an orphan in September 1984 but had been approved in July 1983.

<sup>19</sup> While NORD did not see any statement in the FDA review documents that in this orphan condition, the serum calcium levels never fall spontaneously, if this is nevertheless the case in this condition, then this drug may be classified as having been approved under the May 1998 Guidance conditions and could then be classified as "administrative" flexibility, but again, still flexibility.

ity” and not “case-by-case flexibility.”<sup>20</sup>

### 38. Coagulation Factor VIIa (Recombinant) - NovoSeven

In this March 1999 approval for treating bleeding episodes in hemophilia A or B patients with inhibitors to Factor VIII or Factor IX, the situation was wholly different than, for comparison, for Benefix (Factor IX, #35 above) because with Benefix, the protein was simply providing that which was missing in that individual, whereas the scientific basis for NovoSeven was mainly to provide Factor VII in order to bypass the cascade of Factors VIII or IX since these patients had inhibitors to these other two factors. (Note—there are persons who are deficient in Factor VII but that is not the approved indication here.) In other words, this was not the case of simply supplying exogenously that which was missing endogenously, but this was more akin to a more typical pharmacotherapeutic intervention that relies upon pharmacological intervention to achieve its therapeutic benefit.

Therefore, in this case, the standard expectation of clinical evidence of effectiveness would be expected. However, in this case, there were compassionate use, open-label studies of NovoSeven but, as described in the FDA summary basis of approval (SBA) of March 22, 1999 for these, the “clinical data from [these] were insufficient to evaluate the safety and efficacy of the product by statistical methods.” (SBA, 7). There was also one double-blind, randomized trial comparing two doses of NovoSeven for which FDA states that: “No comparisons between NovoSeven and other coagulation products have been made; therefore, no conclusions regarding the comparative safety or efficacy of NovoSeven can be made.” (SBA, 7). In consulting with an FDA hematologist, there was, of course, an adequate and sufficient scientific basis for this product approval based on the therapy, the condition, the compassionate use data, the dose ranging study, pharmacodynamic and pharmacokinetic studies and animal data; however, this quantum of evidence would not have been sufficient had this therapy been proposed for a prevalent use. Most importantly, FDA here was exercising “case-by-case flexibility.”

### 41. Corticorelin Ovine Triflutate - Acthrel

In this May 1996 approval for differentiating between pituitary and ectopic production of ACTH in patients with ACTH-dependent Cushing’s syndrome, the medical officer’s review of April 5, 1981 notes that the NDA rests upon two “pivotal” bioequivalence trials comparing the sponsor’s ovine corticorelin releasing hormone (oCRH) and the NIH preparation in order to rely upon all the NIH published data to support the efficacy (and safety) of this product. However, the medical reviewer notes that the corticorelin releasing hormone

(CRH) in the published studies were different formulations and sometimes human and not oCRH was used. Moreover, the hormonal response to the oCRH was not defined in either the two “pivotal” bioequivalence trials (which were in 20 “normal” subjects and 10 “normal” subjects) or in the published literature. As for the published literature the medical reviewer notes that all were submitted under the heading of “well-controlled studies without case report forms” (Medical Review, 4) and: (1) that the oCRH formulations differed from study to study and in some, human CRH was used (and with respect to these, the medical reviewer states that “these reports...using [human CRH] do not support any claims on oCRH”); (2) “that efficacy is defined differently from report to report”; (3) that the overall quality of the publications differ widely; and (4) “that the Agency does not have access to the original data to support or discuss the sponsor’s claims” (Medical Review, 5).

### 42. Cysteamine Bitartrate - Cystagon

In this August 1994 approval for treating nephropathic cystinosis, the medical reviewer (in the medical review of October 27, 1993) relied upon 3 open-label multicenter studies: the National Collaborative Cystinosis Study (NCCS), a so-called “Long Term Study” and a UK retrospective study. (Medical Review, 44-45). However, the medical reviewer states that the UK retrospective study only provides “supportive evidence of efficacy” because “a minimal deterioration in renal function [the study’s primary endpoint] occurred in the treated group in the UK study” (Medical Review, 45), and therefore, if any inference can be made about the efficacy of the compound in this disease from this study, it would be against, and not for, the drug’s efficacy.

As for the “Long Term Study,” the “primary endpoints of death and renal death (need for dialysis or transplant) were compared to historical controls represented by 205 unselected, unassociated cystinotic patients analyzed in a retrospective European study. The comparison between the two groups was not prespecified in the study protocol. Inferential statistical testing of the differences was not done because numerical values for each data point were not available for the untreated controls.” (Medical Review, 45, emphasis added). Consistent with this, the statistical review characterized this “Long Term Study” as not well-controlled. Therefore, it is difficult to regard this as a positive trial.

As for the NCCS, the comparison group was noted by the medical reviewer as “a group of patients treated with placebo in a previous double-blind study of ascorbic acid for the treatment of cystinosis.” (Medical Review, 4). The statistical review of December 13, 1993 stated that, “there were statistically significant differences between the cysteamine and placebo groups in terms of age at diagnosis, age at entry, height and renal function for evaluating patients. Due to historically controlled study and insufficient sample size for placebo (n=17) it is very difficult to have meaningful inference between treatment comparisons.” (Statistical Review, 16). It is difficult as

20 This is the only product approved for which NORD provides this kind of explanatory text, because it illuminates the rationale and reasoning behind the classifications for several other blood disorder biologic orphans as being approved with an exercise of “administrative flexibility” (see also #9, 10, 11, 36 & 37). Including this explanation of Benefix also provides an opportunity to illustrate how efficacy standards evolve over time and evidence sufficient at some earlier point in time may no longer be prognostic for FDA action on a later similar product.

well, therefore, to consider the NCCS as a positive adequate and well-controlled trial. However, the role of historical controls in this setting likely provides the basis for the finding of efficacy here.

#### 46. Deferasirox - Exjade

In this November 2005 approval for treating chronic iron overload in patients with transfusion-dependent anemia, the Subpart H/Fast Track approval was based essentially on one single, non-inferiority trial on an unvalidated surrogate primary endpoint. At the pre-NDA meeting, the Division had generally agreed to the statistical methods but had "indicated that the efficacy of DFO [deferoxamine mesylate, the active control] would have to be established and that the margin of -15% would have to be justified in the NDA." (Medical Review, 39 [October 26, 2005]).

DFO had been approved prior to 1982 and is the only FDA approved drug for this use. (Medical Review, 1). The reviewer stated the following with respect to the primary endpoint results of this study: "Exjade [deferasirox] was to be declared non-inferior to DFO if the lower limit of the 2 sided 95% CI for the difference in the percentage of treatment success between Exjade and DFO in the [primary population] was above -15%. For the entire primary population this goal was not achieved. [The success in Exjade was 52.9% and in DFO was 66.4% and the lower limit of 95% CI for difference in percentage of treatment success was -21.6%.] This led the sponsor to segment the primary population into multiple subcategories to determine whether or not non-inferiority could be achieved for any subgroup." (Medical Review, 47). The review then shows the results for eight post-hoc subgroup analysis; it is unclear if this is an exhaustive list of all subgroups analyzed post-hoc. The reviewer then comments that: "These results are problematic. Analysis should be prespecified, not retrospective. The identification of a subgroup in which efficacy is demonstrated can be used for hypothesis generation, but not to provide support for efficacy to gain approval of the drug. Subgroup analysis should lead to a prospective study to establish efficacy in that subgroup. However, the sponsor's argument has merit even though the sponsor's predicament is of its own devise." (Medical Review, 49).

#### 47. Dexrazoxane HCl - Zinocard

This drug was approved in May 1995 for preventing cardiomyopathy associated with doxorubicin. While it appears from three randomized, placebo-controlled trials that the drug is able to prevent and/or reduce the incidence and severity of doxorubicin-induced cardiomyopathy, FDA also notes that in the largest of these 3 trials, which was in breast cancer patients, the patients on the doxorubicin arm with dexrazoxane, "had a lower response rate (48% vs. 63%,  $p = 0.007$ ) and a shorter time to progression than those who received [doxorubicin without dexrazoxane], although survival of patients who did or did not receive [dexrazoxane] was similar." (Label, 13 [comments by the medical reviewer], April 28, 1995).

#### 48. Diethylenetriamine pentaacetic acid - DTPA

In this August 2004 approval for treating patients with known or suspected contamination with plutonium, americium or curium to increase the rates of elimination, FDA had announced in a September 15, 2003 Federal Register notice (that was prior to submission of this NDA) that FDA had already concluded that the drug would be effective based on FDA's review of the Federal government's "database on 646 patients who received one or more doses...during the past 40 years.... In these patients, administration of [this drug] increased the rate of radiation elimination in the urine on average of 39-fold." (60 Fed. Reg. 53,984, 53,986). FDA had established in 2002 a regulation for assessing the safety and efficacy of drugs to deal with the radiation that may be emitted from a "dirty bomb" or other bioterrorism agents. (21 C.F.R. Part 314, Subpart I (so-called "animal efficacy rule"). However, FDA did not require the sponsor to conduct such animal studies either pre- or post-approval.<sup>21</sup>

#### 55. Ethanolamine Oleate - Ethamolin

In this December 1988 approval for treating patients to prevent rebleeding in esophageal varices that have recently bled, it appears from the FDA approved labeling at the time of approval that the demonstration of efficacy was based upon the clinical pharmacology of the drug that causes "fibrosis and occlusion of the vein" when injected intravenously. "The time course of these findings [from human autopsy studies] suggests that sclerosis of esophageal varices will be a delayed rather than an immediate effect of the drug."<sup>22</sup> (Label, 1).

#### 57. Fomepizole - Antizol

In this December 1997 approval for the treatment of methanol or ethylene glycol poisoning, the medical reviewer concludes that:

[T]here seem only two courses possible for this application at this time: (1) [Not Approvable] the entire application so that the company may then perform some decent studies...as a sole therapy employed; (2) approve the NDA for use...only as an adjunct to use of hemodialysis and require the studies under (1) immediately above as phase IV trials. It is true that even if this preparation were completely unavailable at this time...there would be no great hardship or loss. Ethanol, even though it may be more difficult to use, is still an adequate therapy." (Medical Review, Nov. 13, 1987,13). Earlier in his review the medical officer stated the following conclusion on efficacy: "Efficacy when fomepizole is given as a single...agent has not been demonstrated in any sort of controlled study (even historical control).

(Medical Review, 10). The NDA had 2 studies submitted and some historical control data dating back to 1946. The statis-

21 If one regards the experiences of the 646 persons over 40 years who received at least one dose of this drug as having had their results compared to historical controls, then this classification may move from "case-by-case flexibility" to "administrative flexibility".

22 Even if the clinical pharmacology of the drug is a very good surrogate, two positive adequate and well-controlled studies with that surrogate would be needed to satisfy the conventional showing of evidence. If any reclassification were to be considered here on the basis of the use of the clinical pharmacology of the drug, the alternate classification would be under the May 1998 Guidance and therefore, "administrative flexibility".

tical reviewer observes that, "from 1965 to the present, the administration of ethanol as an antidote and the use of renal dialysis have been the treatments of choice." (Statistical Review, Oct. 16, 1997, 4). As for the two studies, the statistical reviewer states that "interpretation of the efficacy results are confounded by the use of ethanol...and the use of hemodialysis in both studies." (Statistical Review, 3). As for the historical data, the statistical reviewer concludes that, "this reviewer does not believe that the historical data is helpful in establishing the efficacy of the drug." (Statistical Review, 5). The statistical reviewer's overall conclusion is that "the sponsor's efficacy database consists solely of data from open-label uncontrolled studies; therefore, there are no statistical issues [because there are no data to analyze statistically]. The descriptive data does not clearly delineate the effects of [fomeprazole] alone since the majority of the patients were treated... in combination with ethanol and/or hemodialysis." (Statistical Review, 4, emphasis in original). Consistent with the statistical reviewer's report, the medical reviewer recommended either not approval of the NDA or approval as an adjunctive indication and the NDA was approved as first line therapy.

#### 58. Fosphenytoin Sodium - Cerebyx

In this August 1996 approval for the acute treatment of patients with status epileptic of the grand mal type, the drug approval is the prodrug use of phenytoin and the medical reviewer notes that it is, "rapidly and completely converted to phenytoin in vivo." (Medical Review, Feb. 1, 1996, 18). The medical reviewer further comments:

This NDA is unique in many ways. First, there are no controlled trials to support the efficacy of [this drug]. The 'controlled trials' submitted were really not designed to show a difference between treatment groups on a protocol specified efficacy outcome. The majority of patients studied were not having seizures, but were only at risk for seizures.... Secondly, the bioequivalence data... really only applies to the isolated instance of IV loading. To my knowledge, no bioequivalence data for IV maintenance dosing, IM loading, or IM maintenance dosing has been submitted.

(Medical Review, 19).

#### 59. Gallium Nitrate - Ganite

In this January 1991 approval for treating clearly symptomatic cancer-related hypercalcemia that has not responded to adequate hydration, the Division Director expressed serious concerns about the nature of the efficacy evidence, specifically, "the participation of only one principal investigation... in the pivotal clinical trials, the performance of the clinical studies, essentially in only one clinical center (Sloan Kettering) [and] Sloan Kettering holds the use patent on the drug." (Division Director's memo, Sept. 28, 1990, 1).

While the single randomized trial comparing gallium to calcitonin reported "achieving normocalcemia in much higher percentage of gallium treated patients than calcitonin treated patients, the overall survival of patients in both treatment groups was poor (median survival time was 29 days for gallium and 35 days for the [calcitonin] group." (NDA Review, Dec. 11, 1989, 4). With respect to the treatment effect of more patients

on gallium achieving normocalcemia, "the treatment effect would not be significant if the expected percentage (60%) of calcitonin patients [had] achieved normocalcemia."<sup>23</sup> (Statistical Review, 8, September 20, 1989).

#### 60. Ganciclovir Sodium - Cytovene

The June 1989 approval for treating cytomegalovirus (CMV) retinitis in immunocompromised patients with AIDS presents a very complicated regulatory history. FDA urged, and the sponsor had agreed, to conduct a prospective, randomized, no-treatment controlled study; however, the NDA ended up being approved on a post-hoc, retrospective review of a case series of subjects treated by one physician at the Johns Hopkins University, and it is that "study" and only that study whose results are shown in the FDA approved labeling.<sup>24</sup>

#### 64. Hemin - Panhematin

In this July 1983 approval for ameliorating recurrent attacks of acute intermittent porphyria (AIP) and similar symptoms in other patients with AIP, porphyria variegata and hereditary coproporphyria, the SBA notes that, "sorbitol serves as a useful stabilizer" in this drug product (SBA, 1), but the SBA later lists five published open studies that were conducted with a formulation without sorbitol and only one, "progress report" of an open-label study with a formulation containing sorbitol. (SBA, 6-8). These six reports together with a couple single dose case reports totaled 125 subjects, of whom over 85% experienced symptom relief on this drug. (SBA, 6). Of the five studies of the formulation without sorbitol, study 1 administered the drug to seven subjects for three to 13 days, study 2 treated 28 subjects for one to six days, study 3 treated 11 patients for three to 13 days, study 4 treated 57 patients for "an unspecified time period" and study 5 treated eight subjects for three to five days, and despite the short duration of these treatments a total of 13 of these 111 subjects died. (SBA, 6-8).

In the single "progress report" that administered the drug in a formulation with sorbitol, these seven patients received drug for two to five days and none were reported to have died. (SBA, 8).

Overall, there was no concurrent control in any study and no reference to any historical control. Moreover, if these studies relied upon each patient's prior clinical experience as his/her own control, there were no reports of the previous patient experiences without the drug. Given the design of these very short duration, open-label, uncontrolled studies for which no mention was made whether line listings, case report forms or even protocols were ever made available to FDA, this may be a case in which FDA relied upon historical controls that were not well documented.

#### 70. Imiglucerase - Cerezyme

In this May 1994 approval for treating Type 1 Gaucher's Dis-

<sup>23</sup> Achieving normocalcemia may be an appropriate surrogate and putting aside the concerns expressed about the single investigator at a single site that has a financial interest in the outcome of the study and the overall poor survival results, then this drug's classification may change to "administrative flexibility".

<sup>24</sup> If one regards that the Hopkins case series was reviewed as though it was compared to a historical control, then this approval may be "administrative flexibility".



ease, the single “pivotal” study compared the 1981 human placenta-derived form of this drug (Ceredase) to the proposed recombinant version, imiglucerase (Cerezyme). The statistical review of September 2, 1983 reports that the primary endpoints were an increase in hemoglobin concentration of at least 1 g/dl, an increase in platelet count and a decrease in liver and spleen volumes over the 6 month study duration in the 15 imiglucerase and 15 Ceredase subjects. (Statistical Review, 2).

According to the statistical review, “The sponsor failed to detect a statistically significant difference with regard to the proportion of patients ([Cerezyme] 11/15, Ceredase 12/15) who achieved an increase of at least 1 g/dl in hemoglobin concentration from baseline to conclusion of 6 months of double-blind treatment. A 95% confidence interval for the between-treatment group difference ([Cerezyme]-Ceredase)... is (-61.5%, 48.1%) which of course is extremely wide.” (Statistical Review, 3). Similar nonsignificant results with wide confidence intervals are seen in the other primary endpoints as well. The statistician was concerned over these wide confidence intervals and to illustrate this noted that, “the 95% confidence interval... indicates that it is statistically conceivable that the Ceredase increase [in hemoglobin concentration] may be as much as 0.52 g/dl greater than the [Cerezyme hemoglobin] increase.” (Statistical Review, 3).

Applying non-inferiority margins to the first efficacy parameter (preserving at least 50% of the benefit seen in the approved active control) would lead to the following:

(1) mean hemoglobin concentration was increased by 1.53 g/dl over baseline in the Ceredase arm and therefore, to be non-inferior the Cerezyme would need to have a lower 95% confidence interval that was greater than +0.765 g/dl, when the lower CI for Cerezyme was -0.52; and

(2) mean platelet count in the Cerezyme arm was increased by  $16.13 \times 10^{-3}/\text{mm}^{-3}$  which means that the lower 95% CI in the Cerezyme arm needed to be greater than  $+8.065 \times 10^{-3}/\text{mm}^{-3}$ , but it was -8.11.<sup>25</sup>

#### 71. Interferon Beta-1a - Avonex

In this May 1996 approval for treating patients with relapsing forms of multiple sclerosis, evidence of efficacy at the time of initial approval came from one randomized, double-blind placebo-controlled study in 301 subjects. The primary endpoint, time to progression, was statistically significant at a p value of 0.02 and the secondary clinical endpoints were generally significant: change in Expanded Disability Status Scale ( $p = 0.006$ ), number of exacerbations ( $p = 0.03$ ), percentage exacerbation free ( $p = 0.10$ , not significant) and annual exacerbation rate ( $p = 0.04$ ). The secondary MRI endpoints were number of lesions at end of year 1 ( $p = 0.02$ ), at end of year 2 ( $p = 0.05$ ), T2 lesion volume at end of year 1 ( $p = 0.02$ ) and at end of year 2 ( $p = 0.36$ ). (See Label, 8-9).

<sup>25</sup> If the comparison between the investigational and the active control is not a non-inferiority comparison, but rather the investigational arm's results are being compared to historical control, then this classification may change to “administrative flexibility”.

While FDA had previously approved another interferon beta compound for MS in 1993, FDA had determined that, for orphan drug purposes, these two were different drugs. Accordingly, Avonex was approved on the basis of this single study (and without reliance on the efficacy results for the previously approved interferon beta drug for MS) in which the primary endpoint results are not “very persuasive” (that is, not less than a p of 0.01).<sup>26</sup>

#### 74. Ferric Hexacyanoferrate (II) - Prussian Blue, Radiogardase

In this October 2003 approval to treat patients with known or suspected internal contamination with radioactive cesium or thallium, there were no prospectively randomized controlled clinical trials and the “best human data on the efficacy of Prussian Blue will come from retrospective analysis of data on accidentally contaminated patients... treated with Prussian Blue.” Such studies cannot, of course, be powered to achieve statistical significance [and] no formal statistical analysis has been performed on this data.” (Medical Review, Sept. 15, 2003, 19).

Nevertheless, the medical reviewer concluded that, “although these publications all describe retrospective studies and the number of patients is small compared to a typical Phase 3 clinical trial, the evidence [of effectiveness] is compelling.” (Medical Review, 42). The reviewer explains that, “in this retrospective study each patient served as his/her own control. For each patient the half-life during treatment was compared to the half-life after treatment had stopped, which was assumed to be equal to the half-life if no treatment had been given.” (Medical Review, 42). Also, the reviewer pointed to animal efficacy data including that, “Prussian Blue has been shown to consistently decrease the half-life of 137 CS [radioactive Cesium] in dogs, rats and farm animals.” (Medical Review, 43).

#### 76. Laronidase - Aldurazyme

In this April 2003 approval for treating patients with mucopolysaccharidosis-I (MPS-I), efficacy was established on the basis of a single randomized, placebo-controlled trial of 45 MPS-I patients. The medical review of April 25, 2003 explained, that, “the study would be considered statistically significant if both primary endpoints of forced vital capacity and 6 minute walk meet or exceed the critical p-value of 0.05 in the difference between the treatment groups.” (Medical Review, 20, emphasis in original). While the forced vital capacity between treatment group difference was statistically significant ( $p = 0.02$ ), the 6 minute walk between group difference was not statistically significant ( $p = 0.07$ ). (See Label, 1). Moreover, there were four prespecified secondary endpoints and only one was statistically significant: apnea ( $p = 0.14$ ), liver volume ( $p = 0.001$ ), Disability Index ( $p = 0.99$ ), and shoulder

<sup>26</sup> This approval can be alternatively read as consistent with the Evidence Guidance if that single study example can be read broadly so as to regard the multiple positive secondary endpoint results and MRI lesion results, in addition to a novel primary endpoint, as fulfilling the May 1998 Guidance for a single study, and in that case, this approval would be considered “administrative flexibility”.

flex ( $p = 0.99$ ). However, the first tertiary endpoint of urinary glycosaminoglycans (GAG) was statistically significant ( $p = 0.001$ ). The medical reviewer concludes that “two markers of in vivo enzyme activity were associated with significant reductions during the 26 weeks of [the pivotal trial]: liver size reductions and urinary GAG concentration. The response of these markers to laronidase has been consistently shown also in the pre-clinical experiments and in the Phase 1 clinical trial, as well as in the placebo-treated subjects switched to laronidase treatment during the open-label extension.” (Medical Review, 113). Therefore, the reviewer concluded, “given the lack of alternative treatments in a rare disease with severe or fatal consequences, this reviewer recommends approval of laronidase, supported by the evidence of efficacy in the co-primary endpoints and favorable trends in subsets of MPS-I in secondary endpoints.” (Medical Review, 113).

### 77. Lenalidomide - Revlimid

In this December 2005 approval for treating patients with transfusion-dependent anemia due to myelodysplastic syndromes (MDS), the approval was based primarily on the results of one single arm, non-randomized, not controlled study. The demonstration of clinical benefit was RBC transfusion independence, defined as having had any rolling 56 day period without need for any RBC transfusion during a treatment duration of up to 672 days. The reviewer commented that, “in MDS, which is a heterogeneous disease, single arm studies using patients as their own controls are generally not acceptable. The sponsor definition of transfusion independence with a rolling duration as defined here is problematic in an unblinded study. In an end-of-Phase 1 meeting...FDA recommended a randomized controlled trial using an endpoint with a longer duration of response.” (Medical Review, April 7, 2005, 65). FDA noted a randomized controlled trial with a longer duration of responses was ongoing at the time of approval. (Medical Review, 135). The first question put to the Oncologic Drugs Advisory Committee (ODAC) for a vote at its September 14, 2005 meeting on this drug was:

Randomized controlled trials allow for direct comparisons of treatment effects and safety between treatment arms. A single arm study has been submitted using an 8-week run-in period to serve as baseline for each patient's transfusion requirements. A comparison is subsequently made to a follow up 8-week period on [lenalidomide] to compare transfusion requirements. Does the study design allow adequate characterization of [lenalidomide's] treatment effect in the population described in the proposed indication?

The ODAC voted yes = 11, no = 5. (Medical Review, 130). However, the ODAC may not have been aware that the comparison was not between periods of equal duration, that is, the comparison was not between the percentage of subjects who were transfusion independent in the run-in eight-week period to the first on-drug eight-week period, but instead the comparison was between the eight week run-in period and any rolling 56 day period of on-drug transfusion independence over a total of up to 672 days (that is, day 1 to day 56, day 2 to day 57, day 3 to day 58). The comparison was between each

subject's transfusion independence over a single 56 day run-in period compared to up to as many as 671 rolling 56-day periods. In addition, to be included in the trial a subject had to have received “at least 2 or more units of RCBs within 8 weeks of study treatment.” So, to be enrolled, a subject had to have had a run-in period with a transfusion of 2 or more units. (See, e.g., study inclusion criteria at Medical Review, 25. See also Medical Review, 43). Therefore, by definition, the “comparator” run-in eight-week period had to have had no subject who was transfusion independent, and there is no mention in the comprehensive 152 page Medical Review to that comparison between each subject's transfusion requirements during the run-in period and during the treatment phase except that FDA notes that 4.7% of the study subjects had only one transfusion in the eight-week run-in period (Medical Review, 44, but these were excluded from FDA's primary analysis of estimating the percentage that were transfusion independent in the treatment phase as protocol violators) and, “the statistical reviewer noted that there was a correlation in the number of pre-treatment RBC transfusions and the transfusion response. It is more likely for those patients with less than or equal to 5 pre-treatment transfusions to develop a transfusion independence response.” (Medical Review, 65).

### 78. Lepirudin - Refludan

In this March 1998 approval for “anticoagulation in patients with heparin-induced thrombocytopenia and associated thromboembolic disease in order to prevent further thromboembolic complications,” the efficacy evidence came from two non-randomized, open-label multi-center (all sites in Germany) trials using a historical control comparator group. However, as noted in the FDA approved labeling, “the key criteria of efficacy ...[was] platelet recovery...[but] comparable rates for the historical control group cannot be given, because [...] platelet counts were not monitored as closely as in the Refluden group.”

Reliance upon a historical control group is fraught with uncertainty generally for many reasons which have been well-articulated elsewhere. However, FDA has relied upon such comparators in the case of rare conditions where the ability to have sufficient subjects to randomize to both the investigational and a concurrent control arm is limited, if not non-existent. (See Label; see also FDA approvals of Myozyme #101 and Ceprotin #99 for infantile-onset Pompe disease).

### 79. Levomethadyl Acetate HCl - Orlaam

In this July 1993 approval for treating heroin addicts suitable for maintenance on opiate agonists, active control (methadone) Phase 3 trials established that response to treatment for levomethadyl acetate was similar to that for methadone. However, there was no formal non-inferiority testing and, although an Advisory Committee indicated it was willing to accept a placebo-control in this patient population, there were no Phase 3 methadone-controlled studies that also included a placebo arm to establish the assay sensitivity of that study design and conduct. Because of the lack of formal statistical

comparisons of the treatment effects of levomethadyl acetate to methadone and the lack of a concurrent placebo arm in any of the Phase 3 trials, the classification is "case-by-case flexibility."

#### 81. Mecasermin Rinfabate Recombinant - Iplex

In this December 2005 approval for treating growth hormone insensitivity syndrome, FDA accepted the sponsor's arguments that "the need for a concurrent control group was obviated by obtaining a documented pre-treatment height velocity in each subject for comparison to on-treatment height velocity [and]...that it was furthermore unnecessary due to the well-known natural history of the condition, in which the poor height velocity is not expected to improve spontaneously." (Statistical Review, Aug. 28, 2005, 7). In a single prospective, open-label multicenter study 36 prepubertal subjects received either 1 mg/kg or 2 mg/kg daily and on the primary endpoint of height velocity, the 1 mg/kg pretreatment values were 3.4 cm/year compared to on-treatment values of 7.4 cm/year ( $p < 0.0001$ ) and the 2 mg/kg cohort had pretreatment height velocities of 2.2 cm/year and on-treatment values of 8.8 cm/year ( $p < 0.0001$ ). The statistician observed that "efficacy is supported by the fact that [the 2 mg/kg cohort] with a higher dose level had a larger growth velocity than [1 mg/kg cohort]," but the statistician also noted that because there was no randomization here, these differences have to be viewed with caution. (Statistical Review, 3).

#### 82. Mecasermin Recombinant - Increlex

In this August 2005 approval for treating growth hormone insensitivity syndrome, FDA permitted the sponsor to pool post-hoc five small clinical trials (four open-label and one double-blind placebo-controlled) to permit a global efficacy analysis relying upon all 71 treated pediatric subjects from these five trials. The primary efficacy analysis was of the 58 subjects for which adequate pretreatment height velocity data were available so that paired t-tests could compare the pretreatment height velocities of the same subjects completing each year of treatment, and the pretreatment height velocity was 2.8 cm/year for these 58 subjects compared to 8.0 cm/year in the first year of treatment ( $p < 0.0001$ ). Without FDA's exercise of scientific judgment in permitting this post-hoc pooling, the pairing of each of these five small trials separately for signs of efficacy would have been problematic.

#### 84. Midodrine HCl - Proamatine

In the September 1996 approval for treating symptomatic orthostatic hypotension, three studies were submitted with the NDA, two with the original NDA and a third added later with respect to the two in the original NDA. The statistical reviewer stated the following conclusion:

The first one...was supposedly a multicenter study, but only one site collected data, and only for 7 patients. This is too few data for the results to be useful. Because of the other difficulties with the study...this reviewer feels that the medical reviewer's (negative) conclusion for the study should be heeded. [Note: it is unusual for the medical reviewer to be an outside consultant as it was in this case: Dr. Joel Morganroth.] The other study... randomized 97 pa-

tients... The analyses...by the sponsor and...by this reviewer did not show midodrine treatment effect... There was no midodrine treatment effect compared to placebo as measured by the syncopal symptoms endpoint.

(Statistical Review, March 13, 1996, 9). With respect to the third study, the statistical reviewer said that:

[T]his study demonstrates that midodrine treatment has a significant effect on systolic blood pressure, and appears to affect standing time and dizziness in this highly selective group of patients. This study is unable by design to show that this temporary effect can be sustained over long-term use. The study contributes very little toward establishing that midodrine is an effective treatment for orthostatic hypotension. The study was too short (seven hours), involved only one dose at the upper level of the dosing range and a three hour dosing interval, was compromised by potential unblinding, and was limited to an enriched population of patients known to respond to midodrine treatment.

(Statistical Review, Sept. 2, 1993, 6).

#### 85. Miglustat - Zavesca

In this July 2003 approval for treating mild to moderate type I Gaucher's disease patients for whom enzyme replacement therapy (ERT) is not an option, the NDA was supported by two Phase 1/2 studies and one Phase 2 study with extension studies to each. In the two open-label uncontrolled monotherapy Phase 1/2 studies, there were four primary endpoints: reductions from baseline in liver and spleen volumes and increases from baseline in platelet counts and hemoglobin. According to the Label, "In study 1...the results showed significant...reductions...in liver volume of 12% and spleen volume of 19%, a non-significant increase from baseline in...hemoglobin... and a [non-significant]...increase in platelet counts... In study 2...the results showed significant...reductions...in liver volume of 6% and spleen volume of 5%. There was a non-significant...decrease...in hemoglobin...and a non-significant increase...in platelet counts." (Label, 5). The statistical reviewer stated that, "Study 004 was an open label, randomized, comparative study with Cerezyme monotherapy as the control group." (Statistical Review, April 27, 2002,3). "The primary objective for the comparative study was to assess the tolerability of [miglustat]... The efficacy analysis of liver volume was exploratory since no clinically meaningful difference was hypothesized and no sample size was determined." (Statistical Review, 27). As for the overall results of these trials and then applications for switching patients from ERT to miglustat, the medical reviewer concluded that: "These results suggest that switching to [miglustat] monotherapy may have a detrimental effect in 'well-controlled' patients with smaller liver and spleen volumes, and higher hemoglobin and platelet counts at baseline who had been receiving ERT." (Medical Review, May 2, 2002, ii).

#### 87. Monoctanoin - Moctanin

In the October 1985 approval of this compound made from medium chain fatty acids derived from coconut oil to dissolve cholesterol gallstones retained in the common bile duct, FDA had issued a Federal Register notice on December 10,

1982 inviting submission of an NDA. In addition to published clinical data, FDA relied on the existence of 4 animal studies already reviewed by FDA, as well as one additional animal (dog) study proposed by FDA whose design is described in the Federal Register notice, as well as in vitro data showing dissolution of gallstones in this compound. (See SBA, 3). Also, "in her memo of September 1982, Dr. Finkel reviewed the reports of clinical trials...published throughout 1981... {the medical reviewer} added reviews of 7 reports which have been published since that time. Results published in the literature support the claim that infusion of monoctanoin into the biliary tract is effective in dissolution of cholesterol stones..."["in about 1/3 of the patients" from Medical Review, 6]. The treatment is attended with a high incidence of adverse effects." (Medical Review, Nov. 26, 1984, 2). In a multicenter study of 377 patients, 32% of the subjects were considered to have had a complete response (Medical Review, 20), however, there was not only no concurrent control but no comparison to historical controls or to using each patient as his/her own control and no formal established analysis of success versus any control arm. (Medical Review, 2-3).

### 88. Galsulfase - Naglazyme

In this May 2005 approval for treating patients with mucopolysaccharidosis IV (MPS IV), the evidence of efficacy was derived essentially from a single, randomized, double-blind, placebo-controlled trial of 39 subjects for 48 weeks. The primary endpoint of 12 minute walk test had a p value of 0.025 (which is not the usual standard for single study in FDA's May 1998 Guidance in that it would not appear to be "a statistically very persuasive finding.") The two secondary endpoints of improvement in rate of stair climbing and urinary GAG levels have p values of 0.053 and less than 0.001, respectively. Also, "among patients who had been randomized initially to placebo [for the double-blind 24 week phase of the trial], the increases after 24 weeks of Naglazyme treatment compared to the start of the open-label period were [comparable in magnitude to the improvements seen in cohort initially randomized to Naglazyme for the 24 week double-blind phase]." (Label, 1). In sum, the primary endpoint of this single pivotal study was less than a p value of 0.05 but greater than 0.01 (that is, not a "statistically very persuasive finding") and one of the two prespecified secondary endpoints was not statistically significant.<sup>27</sup>

### 92. Oprelvekin - Neumega

In this November 1997 approval for preventing severe thrombocytopenia and relieving the need for platelet transfusions following thrombocytopenia chemotherapy in patients at high risk of thrombocytopenia, two randomized, double-blind, placebo-controlled Phase 3 trials formed the basis of the efficacy evidence. In one study of those who had recovered from an episode of chemotherapy-induced thrombocytopenia, the primary endpoint of whether the patient needed one or more

platelet transfusions in the next course of chemotherapy was met with a p value of 0.04. The second study evaluated whether platelet transfusions were needed in either of the next two chemotherapy cycles in patients who had not previously experienced chemotherapy-induced thrombocytopenia. In this study the primary endpoint trended in favor of drug but was not statistically significant. The FDA approved labeling cited one additional positive analysis which, "in an unblinded, retrospective analysis of the 2 placebo-controlled studies, 19 of 69 patients (28%) receiving [oprelvekin] and 34 of 67 patients (51%) receiving placebo reported at least one hemorrhagic event which involved bleeding." (Label, 1).

### 93. Pegademase Bovine - PEG-ADA, Adagen

In this March 1990 approval of this enzyme replacement therapy for ADA deficiency in patients with severe combined immunodeficiency (SCID), the clinical evidence of this drug's efficacy comes from its use in 6 patients with ADA-deficiency SCID. The medical reviewer summarized his review this way:

[I]n view of the rarity of the disease, insufficient cases to study, the orphan status of the disease, the potential lethality of the disease and the non-toxicity of PEG-ADA, the weak data provided might be enough evidence of efficacy in this case.... The strongest support of efficacy is the dramatic biochemical and in vitro immunological modulation by PEG-ADA in these patients, the trend of decreased infections in these patients, and the non-toxicity of PEG-ADA.... School attendance, hospitalizations, bouts of pneumonia, and growth data were inconclusive.

(Medical Review, Addendum IV, Jan. 12, 1989, 2).

### 99. Protein C Concentrate - Ceprotrin

In this March 2010 approval for "patients with severe congenital Protein C deficiency for the prevention and treatment of venous thrombosis and purpura fulminans," there was a single, 18 subject, open-label, non-randomized Phase 2/3 trial with a historical control, as well as a retrospective analysis of 11 other subjects who had been on drug. As described in the case above for Refluden (see #78), a historical control comparator was appropriate here, but it is unlikely that if this condition were prevalent and there was no lack of subjects to enroll in a study, that this showing of efficacy would have been sufficient.

### 100. Rasburicase - Elitek

In this July 2002 approval for treating malignancy-associated or chemotherapy-induced hyperuricemia, the primary clinical efficacy evidence came from a single open-label, randomized, active control (allopurinol) Phase 3 study and two Phase 2 studies.

The Phase 3 study randomized 27 patients to rasburicase and 25 to allopurinol. The primary endpoint was a measure of plasma uric acid levels, and rasburicase was robustly statistically superior to allopurinol, p value of < 0.001. (Statistical Review, Nov. 28, 2000, 6). On each of the three prespecified secondary endpoints, rasburicase was also statistically superior to allopurinol. (Statistical Review, 8). The two Phase 2

27 In his clinical team leader's memo, Dr. Hyde notes that at the January 15, 2003 Advisory Committee on this drug, "some [panel] members expressed a sentiment for liberalizing p-value criteria in diseases as rare and difficult, but important, to study as this." Clinical Team Leader's memo of May 27, 2005 at page 13.

studies were both open-label single arm trials with a total of 238 subjects in both studies combined. The response rate was 99% and 95% in these two studies with uric acid levels reduced by 88% in these studies. (Statistical Review, 2).

If the reviews had indicated the uric acid levels do not spontaneously return to normal, then the implied historical control would have converted the two Phase 2 trials into "adequate and well-controlled" trials and there would be no exercise of judgment in this approval since the efficacy evidence would be straightforward. Similarly, if the review had indicated that it would be unethical to replicate the Phase 3 trial, then the data from the single Phase 3 study would satisfy the single study policy articulated at Section C.3 of FDA's Evidence Guidance. However, neither of those conditions apply and therefore this approval demonstrates the exercise of some scientific judgment and warrants a "case-by-case flexibility" classification.

### 101. Alglucosidase ALFA - Myozyme

In this April 2006 approval for treating Pompe disease patients, the clinical efficacy evidence is derived from a single open-label historically-controlled trial in infantile-onset Pompe disease patients. The study enrolled 18 patients on Myozyme and compared their one year performance on Myozyme against a historic control group of 62 untreated patients with a primary endpoint of invasive ventilator-free survival and proportion of patients alive. The statistical review of April 27, 2006 summarized its conclusion this way:

The historical control subgroup contains data from subjects with birthdates over 20 years. The applicant's analysis points to the potential for improved outcome over time due to more aggressive therapy and better availability of the therapies in more diverse geographic regions. The result from the [historical control] cohort, however, support the contention that the long-term survival of patients with infantile-onset Pompe disease, which are not treated with Myozyme, is poor. The comparison of data between the historical control subgroup and the Myozyme-treated subjects does suggest a treatment effect. This observation is not based on statistical conclusions, per se, but more on the visual inspection of the results in the Myozyme-treated subjects compared with results in the historical control subgroup. The qualification of the treatment difference is almost impossible. Not only are there the issues of improved outcomes, however slight they may be, over time among the untreated subjects, but there remains the issue of selection bias among the Myozyme-treated subjects.

(Statistical Review, 32).

### 102. Recombinant Human Antithrombin - ATryn

In this February 2009 approval for the "prevention of peri-operative and peri-partum thromboembolic events in hereditary antithrombin deficient patients," the efficacy data came from combining one Phase 2, single arm, open-label trial (n=13 evaluable) with one Phase 3, single arm, open-label trial (n=18) to achieve a pooled cohort of 31 subjects on ATryn. The comparison was to those treated with plasma antithrombin and their data for comparison were collected from a prospectively-designed concurrently conducted retrospec-

tive chart review of 35 subjects. If this condition were not so rare, it is likely that a more substantial quantum of efficacy information would have been needed than the non-inferiority comparison based on a pooled comparison of 31 subjects on investigational therapy to a retrospective comparator arm of 35 subjects.<sup>28</sup>

### 103. Respiratory Syncytial Virus Immune Globulin (Human) - Respigam

In this January 1996 approval for prophylaxis of respiratory syncytial virus (RSV) lower respiratory tract infections in infants and young children at high risk of RSV disease, the principal efficacy evidence was from a randomized, double-blind placebo-controlled study in children under 24 months of age and at high risk of RSV disease. In this trial of 510 subjects, the primary endpoint was "the reduction of the incidence of RSV hospitalization (p = 0.047)." (Medical Review, April 30, 1998, 5). Almost all the secondary endpoints also showed a statistically significant separation between placebo and drug arms. (Medical Review, 5). There were two other key trials reviewed: the Cardiac trial and the NIAID trial. According to the medical reviewer:

The Cardiac trial was a...randomized, non-placebo controlled, single-blind study conducted in 429 children with congenital heart disease of less than 48 months of age at enrollment. A 31% reduction in the primary endpoint (RSV hospitalization) was noted in the treatment group compared to the control group (p = 0.164). Not statistically significant reductions were observed in the treatment group of RSV ICU stay, RSV-associated mechanical ventilation and supplemental oxygen use... Adverse events were more severe in the [drug] group (64 children had severe AE compared to 44 control group children).

(Medical Review, 6).

"The NIAID trial was reviewed in detail at the December 2, 1993 meeting of the Blood Products Advisory Committee. At this meeting it was pointed out that the trial conduct was flawed (unblinded, local randomization at a major site)," and "The NIAID trial and the Cardiac trial did not demonstrate efficacy in infants with congenital heart disease." (Clinical Review, 6).

### 104. Rho (D) Immune Globulin Intravenous (IGIV) (Human) - WinRho

In this March 1995 approval for treatment of chronic and acute immune thrombocytopenic purpura, the set of four clinical trials described in the FDA SBA included three small (n of 24, 24 and 63), open-label, single arm trials together with one trial in which 38 subjects were randomized to WinRho and others were randomized to prednisone with either high or low dose IGIV. There were no statistically significant differences found among treatment groups in any of the efficacy vari-

<sup>28</sup> Also, while not affecting the quantum of efficacy evidence directly, it may be of interest to note that the drug was "manufactured" (made by!) genetically-altered cloned goats with the drug expressed in and purified from goat's milk. This is the first (and to date, only) FDA approved use of a cloned genetically-altered animal for drug production.

ables including response rate, peak platelet counts and times to achieve predefined platelet counts. In consultation with an FDA hematologist, such a quantum of evidence would not have been sufficient if this condition were prevalent and the number of subjects capable of being enrolled in trials was not a consideration.

#### 106. Rifapentine - Priftin

This January 1998 approval to treat pulmonary tuberculosis "is based on the 6 month follow-up treatment outcome observed in the controlled trial as a surrogate for the 2 year follow-up accepted as evidence of efficacy in the treatment of pulmonary tuberculosis." (Label, 9). The primary endpoint for the single trial on which this drug was approved was "clinical equivalence on success rate to be no more than 10% worse than [the active approved control] rifampin with two-sided 95% confidence," which was met. (Statistical Review, July 27, 1997, 4). However, the statistical reviewer noted that "the two most important conclusions from this study are the following: 1. The cure rates are comparable between the rifampin (83%) and rifapentine (88%) arms [the primary endpoint].... and 2. There is a statistically [significant] difference between the arms in the chance of a relapse...the risk is 5% for rifampin ...and 11% for rifapentine.... Rifapentine appears to be an effective drug in producing conversion to TB negative sputum...[but] [i]t is less effective than rifampin in preventing later relapse." (Statistical Review, 22-23). Furthermore, the medical reviewer noted that the CDC made a closed door presentation to the Advisory Committee which caused concern within the Committee over this drug's use in HIV positive patients because of "a study presented by the CDC where rifapentine resistance developed in the HIV-positive patients, and the potential for rifapentine to significantly reduce the AUC of the protease inhibitor, Indinavir." (Medical Review, June 19, 1998, 61).<sup>29</sup>

#### 107. Rilonecept - Arcalyst

In this February 2008 approval for Cryopyrin-Associated Periodic Syndrome (CAPS), there was a single double-blind placebo-controlled study, but because of the rarity of this condition, FDA permitted there to be two segmented parts of the study, Parts A (n=47) and B (n=45), with separate randomizations for each part. Both Part A and Part B of the trial met their primary endpoints (p of less than 0.001 for each). Also, while the drug was designated as a Fast Track drug, it received a full approval without the need for a confirmatory Phase 4 study.

#### 108. Riluzole - Rilutek

In this December 1995 approval for treating amyotrophic lateral sclerosis (ALS or Lou Gehrig's disease), the approval rested on the studies, both of which failed to hit their primary endpoints of time to tracheostomy or death according to the prespecified analysis in these placebo-controlled, randomized

29 Query: Does this approval mean that one positive non-inferiority trial versus an approved active control is the equivalent level of evidence as two positive placebo-controlled superiority trials? If so, would one adequate and well-controlled positive superiority trial over an approved active control also be considered the equivalent of two positive superiority trials versus placebo?

trials. The primary endpoint results by the prespecified analysis in these two trials were p values of 0.076 and 0.12. (Label, 2). In both cases, FDA salvaged each trial by permitting a post-hoc analysis that in each case yielded a p value of exactly 0.05 in each trial, not less than 0.05. (Label, 2). In addition, there had been one interim analysis in study 301 with an alpha "cost" of 0.001 so that the hypothesis was being treated to determine not if it were less than a p value of 0.05 but less than 0.049. It is also noteworthy that both trials had numerous secondary endpoints of muscle strength and neurological indices and not only did these not show any statistically significant separation between placebo and drug arms, there was hardly any numerical difference between the groups on these indices. Finally, in both studies, "there was no statistical significance in mortality at the end of the study." (Label, 2). The FDA medical reviewer notes that the apparent improvement in survival occurs early in each study period and the Kaplan-Meier curves came nearly together at the end of the study period, so that the FDA medical reviewer further observes: "Of course, the unanswered questions are whether the Kaplan-Meier curves eventually meet and follow a common path therefore [or] potentially the curves could cross with cumulative survival being worse on drug after 2-3 years." (Medical Review, 22 [Aug. 18, 1985]).

#### 110. Rufinamide - Banzel

In this November 2008 approval for treating Lennox-Gastaut Syndrome, there was a single placebo-controlled randomized study (n=138) which was robustly statistically positive on all three co-primary endpoints of seizure activity (p values of 0.0015, <0.0001, 0.0041). However, Institutional Review Boards may not have found it unethical for a second study to be conducted. Therefore, the "statistically very persuasive finding" in this one trial may not have satisfied the strict application of FDA "single study" policy in its Evidence Guidance (See Section II.C.3). However, this sponsor also conducted 2 large studies of this drug in a prevalent disorder, "partial seizures," and while FDA did not find the efficacy evidence in these 2 "partial seizure" trials adequate to warrant the drug's approval for that prevalent indication, FDA found that the efficacy evidence in these 2 studies provided "additional support"<sup>30</sup> for the orphan indication as noted in the final sentence of the conclusionary paragraph by the medical reviewer on the efficacy evidence for the Lennox-Gastaut use: "The agent is additionally supported by the evidence from the partial seizure trials which indicate anticonvulsant activity." (Medical Review, Oct. 1, 2008, 77).<sup>31</sup>

#### 111. Sacrosidase - Sucraid

30 Query: Does this constitute "confirmatory evidence" under FDAMA 115?

31 The Evidence Guidance explains that studies in a closely related disease can essentially supply the "second" study necessary for approval and, coincidentally, FDA in this same 1998 Guidance cites an eerily nearly identical earlier precedent when FDA observes that: "The recent approval of lamotrigine for treatment of Lennox-Gastaut Syndrome (a rare, largely pediatric, seizure disorder) was based on a single adequate and well-controlled study, due in part to related data showing efficacy of the drug in partial-onset seizures in adults." (Evidence Guidance, 10). However, it is critical to note the difference between lamotrigine and rufinamide is that FDA viewed that lamotrigine had established proven efficacy in partial-onset seizures in adults.

In this April 1998 approval for treating congenital sucrase-isomaltase deficiency, there were two identically designed key trials: one of which was negative and one positive. However, FDA approval was based on a single double-blind, randomized, placebo-controlled, dose-response positive trial in face of a conflicting negative trial because the positive trial not only met its primary, but almost all of its secondary, endpoints which showed not only clinical improvements (e.g., fever, watery stools, more solid stools), but mechanistically showed that better results were observed in those who had higher enzyme (that is, drug) levels. In addition, there was a dose-response and subjects responded well to sucrose challenge. (Medical Review, Aug. 14, 1997, 82-84). However, the statistical reviewer concluded by recommending yet a third trial be conducted prior to approval. (Statistical Review, 19-20 [Sept. 15, 1997]).

### 112. Sapropterin Dihydrochloride - Kuvan

In this December 2007 approval for reducing blood phenylalanine (Phe) levels in patients with BH4-responsive phenylketonuria (PKU), there were four efficacy studies. The primary so-called "efficacy study" was a randomized, double-blind, placebo-controlled study (n=88) with a primary endpoint of mean change in Phe at week six ( $p < 0.001$ ). The FDA medical review of December 7, 2007 concluded that this finding was both clinically meaningful and statistically significant, as well as noted important secondary endpoint results of clinically meaningful decreases in blood Phe levels at weeks one, two and four, which supported the primary endpoint finding. (Medical Review, 12). Other findings in a "Diet Study" and an "Extension Study" provided additional confirmatory evidence of efficacy. (Medical Review, 14). For instance, while the medical reviewer did not find the statistically significant primary endpoint results in Part II of the Diet Study to be clinically meaningful, the reviewer noted that a "secondary efficacy finding {in Part II of the Diet Study which was mean change in blood Phe from baseline to week 3} supports the primary efficacy finding of the Efficacy Study." (Medical Review, 13).

### 119. Sterile Talc Powder – Sclerosol

This December 1997 approval for treating malignant pleural effusions was based solely on published literature. The statistical review of January 5, 1996 notes that: "Talc has been used for years to treat patients with malignant pleural effusions, but talc has never been approved by the FDA for this purpose. It was felt that if approval were granted, there would be more control over the mechanism by which patients are treated with talc. For example, one concern is the asbestos which some talc contains." (Statistical Review, 2). In determining that substantial evidence of efficacy was provided in this NDA, FDA overcame concerns with both the quantum and quality of evidence as seen in the following comments about the published studies:

Each study was sponsored by an investigator and there was no control body coordinating these research activities. Consequently, the studies use different study designs, different doses of talc, dif-

ferent routes of administration, different control groups, different definitions of response, and different lengths of follow-up. No CRFs are available, so it is impossible to determine exactly how the patients were treated and exactly how they responded. The quality of the safety data and prognostic factors for efficacy variables is then compromised.

(Statistical Review, 2). The statistician viewed five of the published studies as being of more reliable design and/or quality. Of these five studies, the statistical reviewer noted that in the intent to treat analysis, only one of these five had a statistically significant higher response rate in the talc group than in the control group. The other four of five studies had a statistically significant response rate in the evaluable population, but this analysis has a "potential bias" in that in three of the five studies the talc group, "was associated with a higher incidence of premature death than the control group." (Statistical Review, 10-14).

### 122. Tetrabenazine - Xenazine

This August 2008 approval for treating chorea associated with Huntington's disease relied upon one 12 week randomized, double-blind, placebo-controlled trial (n=84) and one five-day randomized, double-blind, placebo-controlled, staggered withdrawal study (n=30). In the larger efficacy trial, the primary endpoint of change from baseline in the Chorea Score (a subset of the Motor Assessment Scale of the Unified Huntington's Disease Rating Scale) for the average of weeks 9 and 12 was statistically significant ( $p = 0.0001$ ); however, the primary endpoint of the smaller, staggered withdrawal study had only a trend suggestive of efficacy, but was not statistically significant for its primary endpoint.

### 123. Thalidomide - Thalomid

This July 1998 approval to treat erythema nodosum leprosum (ENL or leprosy) relied upon "primary data demonstrating the efficacy of thalidomide...[that] are from the published medical literature and from a retrospective study of 102 patients treated by the U.S. Public Health Service (PHS)." (Label, 7). With respect to the PHS study, the statistical review of August 7, 1997 stated:

These [102] patients were treated from 1973 to 1997, which is a long period of time. Hence, the data generated from these medical records is of varying quality and completeness. No analytical protocol was available... no comparative drug or therapy was used, subjects were not randomized to treatment groups, and there is no fixed dose or duration of dose, no rules of titration up and down. This data set is of inferior quality as compared to the data from an adequate and well-controlled clinical trial...therefore the statistical analysis of this review will not contain any p values.

(Statistical Review, 1-2). Subsequently, the statistical reviewer stated that, "this data set is not from an adequate and well-controlled study." (Statistical Review, 36).

### 124. Tiopronin - Thiola

This August 1988 approval to prevent cystine nephrolithiasis in patients with homozygous cystinuria has an unusual regulatory history. The medical reviewer states,

In 1979, the sponsor of this NDA was approached by FDA to consider obtaining an IND for Thiola and organizing a multiclinic trial with this drug.... The sponsor was advised that two other investigators had declined to undertake this task. A specific guideline for the preparation of the IND was provided to the sponsor [by the FDA]. A requirement of the inclusion of the placebo control group for the multiclinical trial was also deleted by the FDA, when potential co-investigators refused to conduct [a] randomized trial for bioethical reasons.... On December 5, 1985, the FDA invited the sponsor to submit a new drug application.

(Medical Review, July 25, 1988, 2-3). The medical officer concluded by finding efficacy on the basis of the sponsor's report of 57 patients treated with this drug, using each patient as his/her own control. (Medical Review, 24-25).

### 125. Tranexamic acid – Cyklokapron

In this December 1986 approval for treating hemophilia patients “to reduce or prevent hemorrhage and reduce the need for replacement therapy during and following tooth extraction,” all the efficacy evidence came from 6 published literature studies that were all conducted “in the late sixties and early seventies” (that is, more than a decade and a half before the approval) and only one of these studies was placebo-controlled, randomized and double-blind, with two others open and retrospective and the remaining 3 uncontrolled. (Medical Review, Nov. 6, 1985, 18).

### 126. Treprostinil sodium - Remodulin

This May 2002 approval to treat pulmonary arterial hypertension was based on the results of two concurrently run, identically designed trials, both of which were double-blind, randomized, placebo-controlled with a primary endpoint of the 6 minute walk test of exercise capacity. The sponsor and FDA had agreed in advance that a positive result would be either: (a) both trials having a p value of <0.05 on the primary endpoint; or (b) one trial having a p of <0.05 plus the pooled result having a p value of <0.01. The primary endpoint results of each of the two trials were p values of 0.0607 and 0.0550, while the pooled result was 0.0064.

### 127. Trientine - Syprine

This November 1985 approval for treating Wilson's Disease was based on a summary of results obtained by two different investigators in a total of 41 subjects, in which there were no concurrent controls. Particularly, there was no placebo control as the FDA medical reviewer observed that: “The sponsor did not initiate and/or subsidize the [two] clinical trials reported herein. They were carried out independently by two recognized experts in the field. The sponsor was able to obtain the detailed records of the cases and to transfer the data to case report forms for inclusion in this NDA. Placebo-controlled studies were not done because they would be flagrantly unethical in this disease.” (Medical Review, April 9, 1984, 2).

### 128. Trimetrexate Glucuronate – TMTX, Neutrexin

This December 1993 approval to treat pneumocystis carinii pneumonia (PCP) in AIDS was based on a single randomized, active-control (trimethoprim/sulfamethoxazole or TMP/

SMX) trial. According to the medical review:

The stated objective of this study was to attempt to show that TMTX was superior to TMP/SMX with respect to survival of the PCP episode, as assessed at day 56. Clearly the data do not support such conclusion [because the risk of death in the TMTX group was roughly twice that in the TMP/SMX group]. However, from the regulatory perspective, this was not the appropriate objective. From a scientific and regulatory perspective, the objective should have been to attempt to show that TMTX was ‘equivalent’ to the approved therapy, TMP/SMX. The treatment groups were equivalent with respect...to the percentage of successful responders [which] was 50% for each...group.

(Medical Review, Aug. 9, 1993, 38). Further, “The reasons for failing to respond to therapy were, however, different for the two treatments. TMTX patients were more likely to fail due to lack of efficacy, while TMP/SMX patients were more likely to be failures due to treatment limiting toxicity.” (Medical Review, 38).

### 129. Vaccinia Immune Globulin (Human) Intravenous - VIGIV

This February 2005 approval to treat severe complications from the smallpox vaccine was based on two studies in healthy volunteers, and without any controlled studies showing benefits such as decreased mortality or severity of smallpox. One study was an open-label safety study in 33 healthy volunteers and the sole evidence of efficacy was an open-label study in 78 healthy volunteers in whom the sponsor showed serum neutralizing antibodies for vaccinia 5 days after drug, which “were not less than those expected following a similar dose of” an approved therapy. (Label, 6).

### 131. Vigabatrin - Sabril

This August 2009 approval for treating infantile spasms was based on “studies...that are principally derived from published reports.” (Cross-discipline Team Leader Review, [July 20, 2009]). There were three controlled studies submitted: Study FR03, of which the cross-discipline leader stated, “would not normally meet the criteria as a pivotal trial” (Cross-discipline Leader Review, 11); Study 1A “does not meet the normal standards for the FDA for reasons described above (e.g. lack of a predefined protocol, interim statistical plan, questions regarding the completeness of the blinding...)...nevertheless the primary endpoint analysis would suggest a positive effect” (Cross-discipline Team Leader Review, 10); and Study WO19 whose prespecified primary endpoint was change in average spasm frequency as measured over a 2-hour window ( $p = 0.562$ ). However, “this endpoint was generally considered inadequate by Dr. Sheridan as it provided a very small sampling of seizures and therefore was likely to result in a larger variance...this combined with the small size of the study was unlikely to provide adequate power to detect a treatment effect. One of the secondary endpoints in Study WO19 included a 24-hour...observation window. When this is examined a large and statistically significant ( $p = 0.03$ ) difference is observed with a 68.9% reduction in the vigabatrin group and a 17% [reduction] in the placebo group. Thus, while the primary end-



point of this study was negative, the endpoints, which were also not optimal, suggested an effect.” (Cross-discipline Team Leader Review, 11).

**132. von Willebrand Factor/Coagulation Factor VIII Complex - Wilate**

In this December 2009 approval for treating “spontaneous and trauma-induced bleeding episodes in patients with severe von Willebrand disease”, the results of four open-label, non-randomized, non-controlled trials in a total of 70 subjects were pooled for analysis (and several subjects participated in more than one of these trials, raising also issues of patient selection bias). It is observed that at the time of these trials there were two other FDA-approved therapies for this condition, Alphanate and Humate-P, and therefore, the possibility of a non-inferiority trial without exposing subjects to the risk of randomization to a placebo arm was a possibility. However, the FDA statistical review of this application stated that these “efficacy data of Wilate are considered as secondary and are derived from [4 studies] which were open-labeled and uncontrolled” and therefore a “PK study...is the pivotal study for the basis of the product approval.” (Statistical Review, 16). This quantum of efficacy evidence, while entirely appropriate for this orphan condition, illustrates an FDA exercise of judgment in its review of therapies for rare conditions.

**133. Zalcitabine - Hivid**

In this June 1992 approval for treating MDS, FDA relied upon “2 small studies. The first was a Phase 1/2, open-label, dose-ranging study...the second study was a randomized Phase 2 study designed to evaluate the virologic and immunologic effects of the combined administration of two nucleoside analogues (zidovudine combined with either [zalcitabine] or didanosine.) Both studies used an experimental regimen of zidovudine...and neither was designed to assess the clinical efficacy of the combination.” (Label, 3).

APPENDIX 2: SUBPART H AND FDAMA 115

In its May 1998 Evidence Guidance, FDA describes nine different circumstances in which a single trial may provide the statutorily-required effectiveness evidence. Often this guidance has been misread to mean that only the last of the nine circumstances represents a situation in which a “single” study may be adequate. The last circumstance is a situation in which a highly persuasive statistical finding (a p value of less than 0.01 and often even “more persuasive” than that) in a single trial with some other indicia of the study’s reliability (e.g., multicenter with no center driving the results) out of a pot-pourri of possible factors that may provide such additional credibility to the primary endpoint finding and where it is likely unethical to conduct a second study.

However, it is critical to observe that FDA lays out eight other circumstances in this same guidance in which a single study may be adequate for meeting the statutory standard. However, of the other eight circumstances of “single study” approvals described in the May 1998 Guidance, only one is relevant to a new chemical entity. Therefore, for purposes of this analysis of orphan drugs approved as new chemical entities, there are only two circumstances for a single study approval applicable to a new chemical entity described in the May 1998 Evidence Guidance.

At the same time that FDA was developing what later became its May 1998 guidance, Congress was enacting an amendment to that 1962 effectiveness standard that created a new alternative statutory standard for establishing a drug’s effectiveness. This new alternative statutory standard is: “one adequate and well-controlled study and confirmatory evidence.” This provision of the law is referred to as FDAMA 115 (after the section in the FDAMA that inserted this statutory standard into the law.)

The nine types of circumstances that FDA described whereby a single study may be sufficient to prove a drug’s treatment benefit had been based by FDA on its 36 years of collective experience and set forth in its May 1998 guidance. These nine types of circumstances can be seen as ways for implementing the FDAMA 115 “one adequate and well controlled study and confirmatory evidence” alternative statutory standard. In this way, the May 1998 Evidence Guidance and FDAMA 115 can be seen as fundamentally similar policies that were fortuitously issued almost simultaneously. One must, however, guard against a commonly-held misconception which is that the ninth of those nine circumstances in the May 1998 Guidance is the sole method for approving a drug based on a single trial. There are eight other circumstances described in the May 1998 Guidance itself. Moreover, the breadth of the FDAMA 115 “one adequate and well-controlled study and confirmatory evidence” statutory standard extends beyond these nine circumstances described in the May 1998 Guidance. For instance, Dr. Russell Katz of FDA at an FDA orphan drug conference in October 2010 presented the approval of

tetrabenazine for Huntington’s disease as an example of FDA employing the FDAMA 115 standard in approving this orphan drug.



**NORD**

National Organization for Rare Disorders

**Connecticut Office**

55 Kenosia Avenue  
PO Box 1968  
Danbury, CT 06813-1968  
(203) 744.0100

**Washington DC Office**

1779 Massachusetts Avenue NW  
Suite 500  
Washington, DC 20036  
(202) 588.5700

[rarediseases.org](http://rarediseases.org)