Published in final edited form as:

J Clin Pharmacol. 2018 July ; 58(7): 885–890. doi:10.1002/jcph.1109.

Primary Endpoints in Pediatric Efficacy Trials Submitted to the US FDA

Dionna J. Green, MD¹, Janelle M. Burnham, MD¹, Paul Schuette, PhD², Xiaomei I. Liu, PharmD³, Brian M. Maas, PharmD⁴, Lynne Yao, MD⁵, Susan K. McCune, MD⁶, Joseph Chen, PharmD⁷, John N. van den Anker, MD, PhD, FCP⁴, and Gilbert J. Burckart, PharmD, FCP¹

¹Office of Clinical Pharmacology, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD, USA

²Office of Biostatistics, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD, USA

³Children's National Medical Center, Washington, DC, USA

⁴University of North Carolina Eshelman School of Pharmacy, Chapel Hill, NC, USA

⁵Office of New Drugs, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD, USA

⁶Office of Pediatric Therapeutics, Commissioner's Office, US Food and Drug Administration, Silver Spring, MD, USA

⁷University of California San Diego, Skaggs School of Pharmacy, San Diego, CA, USA

Abstract

The selection of appropriate endpoints in pediatric drug development trials is a critical aspect of trial design. Given the high pediatric trial failure rate, selecting optimal trial design elements, such as the primary efficacy endpoint, is essential to ensuring increased potential for trial success. The objectives of this study were to identify the primary efficacy endpoints measured in pediatric drug development trials submitted to the US Food and Drug Administration and to relate endpoint attributes to trial and label outcome. The analysis included pediatric pivotal efficacy studies submitted from September 2007 to July 2016 for which there was a corresponding adult trial for the same indication. Two hundred and thirty-four efficacy trials on 138 unique products studied in pediatric patients were assessed. The adult and pediatric endpoints were the same in 141 of the 234 trials (60.3%), and these trials succeeded in meeting their primary endpoint more often (122 of 141 [86.5%]) than when the adult and pediatric endpoints differed (57 of 93 [61.3%]; odds

Corresponding Author: Gilbert J. Burckart, PharmD, FCP, US Food and Drug Administration, Bldg. 51, Room 3184, 10903 New Hampshire Ave., Silver Spring, MD 20903 Gilbert.Burckart@fda.hhs.gov.

Declaration of Conflicting Interests

The authors declare no potential conflicts of interest with respect to the research, authorship, and publication of this article.

Funding

The authors received no financial support for the research, authorship, and publication of this article.

Disclaimer

The opinions expressed in this article are those of the authors and should not be interpreted as the position of the U.S. Food and Drug Administration.

ratio, 4.03; 95%CI, 2.10–7.80). Trials that included both pediatric and adult patients succeeded more frequently than those trials that did not combine pediatric and adult patients (85 of 95 versus 94 of 139, respectively; odds ratio, 4.05; 95%CI, 1.94–9.31). No differences were observed in pediatric trial success between those using subjective and objective endpoints. Using the same endpoint in the pediatric trial as was measured in the corresponding adult trial and enrolling pediatric and adult patients in the same trial were attributes associated with trial success.

Keywords

pediatrics; regulatory/scientific affairs; clinical trials; endpoints; US FDA

There continues to be advancement in efforts to ensure that medications used in children are safe and effective, but pediatric drug development still represents a relatively new science. With the passage of the Best Pharmaceuticals for Children Act (BPCA) in 2002 followed by the passage of the Pediatric Research Equity Act in 2003, a substantial increase in the number of pediatric drug studies submitted to the US Food and Drug Administration (FDA) occurred.^{1,2} However, there still remain a significant number of trials that fail to result in an FDA-approved indication for pediatric use.³ In 1 prior assessment, up to 42% of studies performed under BPCA failed to achieve an indication for use in pediatric patients.²

To obtain marketing approval in the United States, drug manufacturers must demonstrate the effectiveness and safety of their products, usually by conducting adequate and well-controlled trials.^{4,5} Although a large number of factors go into the design and ultimately determine the success of a pediatric trial, the selection of a primary efficacy endpoint is vital to a well-designed study and serves as the basis for robust evaluation of the clinical impact of the therapeutic intervention. The use of appropriate endpoints that are well defined, reliable, interpretable, and directly applicable to the disease or condition being studied is critical.

Given its influence on trial outcomes, the selection of efficacy endpoints in pediatric trials should involve a rigorous process that is dependent on scientific and practical considerations. The objectives of this systematic review were to describe the attributes of primary efficacy endpoints measured in pivotal pediatric efficacy trials submitted to the FDA, to compare endpoints measured in pediatric trials with those measured in pivotal adult trials, and to assess the relationship between attributes of pediatric endpoints and trial/labeling outcomes.

Methods

Pediatric efficacy trials from drug development programs with data submitted to the FDA under the Food and Drug Administration Amendments Act (September 2007-June 2012) and the Food and Drug Administration Safety and Innovation Act (July 2012-July 2016) were surveyed. Only trials that were deemed pivotal in supporting a labeling claim for efficacy were reviewed as part of this survey, and only pediatric trials for which there was a corresponding adult trial for the same indication studied were included in the analysis. FDA-authored reviews of each trial were used to extract the following data: drug name, indication

studied in children, therapeutic area, age group studied, primary efficacy endpoint, time of measurement of primary efficacy endpoint, trial outcome, and drug approval status. For a given indication studied in children, information regarding the primary efficacy endpoint measured in the corresponding adult trial was extracted from the FDA-approved drug label obtained from the drugs@FDA database.

For the purposes of this analysis, primary endpoints measured in pediatric trials were classified into the following categories: subjective versus objective vs. both, and same versus different than the adult endpoint. Measures that are highly based on personal opinion, interpretation, emotion, or judgment were categorized as subjective endpoints and those that were not were categorized as objective endpoints. An endpoint measured in a pediatric trial was considered different from an endpoint measured in the corresponding adult trial when the endpoint itself was different and/or when the time of measurement differed between the 2 populations. Trials that enrolled both patients younger than and older than 18 years of age were considered combined adult and pediatric trials.

Trial outcome was categorized as a “success” or “failure” based on whether the trial successfully met its primary efficacy endpoint. Drugs that were approved only in an age subset of the pediatric patients studied in the trial were considered approved.

Statistical evaluation was performed using Blaker’s exact method,^{6–8} which was chosen as a method suitable for the range of observed values. Computations were performed using R v3.4.3 (www.r-project.org) and employed the R package’s Blaker CI and exact 2×2 . Odds ratio (OR) and the 95% confidence interval (CI) were calculated as appropriate. No adjustment for multiple testing was made, as this analysis was exploratory.

Results

Two hundred thirty-four pivotal pediatric efficacy trials were submitted to the FDA during this period. The trials evaluated 138 unique drug products in children, with some products being studied in more than one trial. Fifty-two different indications and 171 distinct endpoints (Supplemental Table S1) are represented, and the most frequently studied areas were the pulmonary, antiviral, and allergy therapeutic areas (Table 1). Of the 234 trials, 95 (40.6%) were combined trials in which both adult and pediatric patients were enrolled in the same trial. Forty-four products were evaluated in combined adult and pediatric trials, and these trials were most frequently seen in the therapeutic areas of allergy, dermatology, and pulmonology.

Of the 234 trials, 179 (76.5%) successfully met their primary efficacy endpoint, whereas 55 (23.5%) failed (Figure 1). Consequently, 107 of the 138 unique drug products (77.5%) were granted FDA approval for pediatric use, whereas 31 (22.5%) were not approved for use in the pediatric population studied. The therapeutic areas with the highest nonapproval rates included ophthalmology (2 of 2 [100%]), oncology (7 of 10 [70%]), hematology (4 of 8 [50%]), cardiology (3 of 7 [43%]), and analgesia/anesthesia (4 of 10 [40%]); see Table 1.

Trial outcome was improved when pediatric patients could be included in the adult trials compared with when pediatric trials were conducted separately (89.5% success [85 of 95] vs 67.6% success [94 of 139]; OR, 4.05; 95%CI, 1.94–9.31; Figure 1).

In 141 of the trials (60.3%), the primary efficacy endpoint measured in the pediatric trial was the same as the endpoint measured in the corresponding adult trial, whereas in 93 of the cases (39.7%), the pediatric and adult endpoints were different (OR for success/failure, 4.03; 95%CI, 2.10–7.80; Figure 2). When the primary efficacy endpoint measured in the adult and pediatric trials differed, the reasons were the time of measurement of the endpoint was different (35 of 93 [37.6%]), the endpoint measure itself was different (32 of 93 [34.4%]), or both the endpoint measure and timing of measurement were different (26 of 93 [28.0%]); see Figure 3. For the comparison of timing differences and both differences, the OR for success/failure was 8.61 (95%CI, 2.54–32.04). For the comparison of endpoint differences versus both endpoint and time of measurement differences as it related to trial outcome, the OR was 4.18 (95%CI, 1.32–13.53). Of the 114 trials that measured objective endpoints, 23 failed (20.2%), whereas 31 of the 99 trials that measured subjective endpoints failed (31.3%). Only 1 trial of 21 failed (4.8%) when a composite endpoint composed of both objective and subjective measures was used. Subjective endpoints had a success rate that was less than the composite endpoint (OR, 0.11; 95%CI, 0.0052–0.68); see Figure 4.

Discussion

The results of this analysis of pediatric primary endpoints in efficacy trials submitted to the US FDA demonstrate that endpoint selection does have an important impact on the outcome of a drug development program. Two attributes of endpoints were identified as being different between those trials that met their primary endpoint and were successful in obtaining pediatric labeling for the indication pursued compared to those that did not. Those successful attributes included: (1) using an endpoint in a pediatric trial that was the same as the endpoint measure in the corresponding adult trial, and (2) having pediatric and adult patients enrolled in the same trial. Of particular interest was the observation that incorporating unique endpoints into pediatric trials that differed from those measured in adult trials represented a risk factor for trial failure.

Efficacy endpoints measured in adult trials for a given indication may not always be suitable for use in pediatric trials.⁹ This may be because of differences between the 2 populations in the natural history (eg, migraine),¹⁰ in the pathophysiology, or in the symptomatology (eg, heart failure)¹¹ of the disease or condition under study. Furthermore, the adult endpoint may not be suitable to measure in certain pediatric age subgroups because of differences in performance capacity. For example, the primary efficacy endpoint commonly measured in interventional pulmonary arterial hypertension trials in adults is the 6-minute walk test (6MWT), which is a measure of exercise capacity.^{12–14} However, the 6MWT would not be a feasible endpoint to measure in infants and young children because they are not developmentally or physically capable of performing the test.^{15,16} In other instances, the relatively small pediatric patient population available or the infrequency of endpoint events often make it unfeasible to incorporate the adult endpoint into a pediatric trial because of the

sample size that would be required to achieve adequate statistical power to detect a difference.

However, pediatric trials in our cohort failed more frequently when they were using endpoints that were different than the endpoint measured in the preceding adult trial. When trials fail, it is not always clear whether it is because of the endpoint selected or other factors. For example, the drug may not work in pediatric patients even if it works in adults for a disease that is considered to be similar. The lack of experience with an endpoint is another possible reason for trial failure. This leads to the inclusion of endpoints that are not validated for the purpose for which they are being used in the pediatric drug development program. Validation and acceptance of trial endpoints have been recognized as problems in drug development since the Critical Path document of 2004.¹⁷ The FDA does not require that an endpoint be formally validated prior to its use in a trial. In fact, most biomarkers are eventually accepted as useful and informative through prior use in a successful drug development program, so the longterm experience with developing drugs for use in adults makes the endpoints used more likely to be accepted in the context of drug development.

Although subjective outcome measures can provide valuable insight into the therapeutic effect of a drug, they have the potential to introduce bias into the trial, particularly in the case of an open-label trial. Even the precise definition of a “subjective” outcome varies considerably among trials.¹⁸ Some authors have suggested that subjective assessment should be minimized in a randomized clinical trial,¹⁹ but that is not possible in all therapeutic indications (eg, pain, major depressive disorder, allergic rhinitis). Also, the use of clinical outcome assessments and patient-reported outcomes in pediatric studies may be important considerations in the future for pediatric drug development.²⁰

The finding that the use of a composite endpoint including both objective and subjective components within a trial resulted in few study failures could lead to the conclusion that this is an attractive approach, but the use of composite endpoints in adult drug development has been questioned.²¹ When using a composite endpoint, each of the individual components is of equal importance in the analysis and characterization of the overall treatment effect of the drug. However, the interpretation of the results obtained when measuring a composite endpoint may be misleading if the clinical importance of the various components are markedly different, as the component with the least importance could be driving the favorable treatment effect or masking that the component with the greatest importance is adversely affected by the drug. Therefore, careful consideration should be given to the choice of the components of a composite endpoint, and each of the individual components should have reasonably similar clinical importance.²² For the most part, composite endpoints have been used very successfully in adult drug development.^{23–25}

The enrollment of both adult and pediatric patients into a single trial may represent a reasonable approach when the available safety data permit and the disease in the 2 populations are considered similar. We found that trial outcome was better when pediatric patients could be included in the adult trial compared with when the pediatric trial was conducted separately. This was likely because of an increased understanding of the disease being studied, prior experience with the endpoint, and the similarity of the adult and

pediatric diseases. For example, in our cohort, asthma, acne, and allergic rhinitis were the diseases for which combined adult and pediatric trials were commonly conducted. This approach also helps to address the problem of having a small sample size in pediatric trials, which is a common problem. Patient selection criteria and enrichment of the pediatric study population are additional considerations for trial success, and a recent review of pediatric enrichment has been published.²⁶

Although our cohort contained more than 200 trials, this should be considered an exploratory analysis because of the relatively small numbers of trials in some therapeutic areas. It can also be difficult to determine any one isolated reason why a trial may have failed, and there are often multiple factors contributing to trial outcome.³ Therefore, solely attributing trial failure for any given indication to endpoint selection may be incorrect. As such, caution should be used in applying these findings.

In conclusion, endpoint selection is a key factor influencing the success of pediatric drug development trials. Two attributes of endpoints were identified as differing between those trials that were successful and those that failed in obtaining pediatric labeling for the indication pursued: (1) measuring the same endpoint in the pediatric trial as that measured in the corresponding adult trial, and (2) enrolling pediatric and adult patients in the same trial. When the endpoint chosen has not been used in adult studies, careful consideration should be given to its validation in this context of use. Studying certain therapeutic areas remains problematic because of the percentage of failed pediatric trials in those areas. Understanding the pediatric disease process and understanding the endpoint, as well as careful consideration of trial designs, are all integral components of obtaining success in pediatric drug development.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Sachs AN, Avant D, Lee CS, Rodriguiz W, Murphy MD. Pediatric information in drug product labeling. *JAMA*. 2012;307:1914–1915. [PubMed: 22570457]
2. Wharton GT, Murphy MD, Avant D, et al. Impact of pediatric exclusivity on drug labeling and demonstrations of efficacy. *Pediatrics*. 2014;134:e512–518. [PubMed: 25022732]
3. Momper JD, Mulugeta Y, Burckart GJ. Failed pediatric drug development trials. *Clin Pharmacol Ther*. 2015;98:245–251. [PubMed: 25963725]
4. Guidance for Industry: Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products. US FDA; 1998.
5. Downing NS, Aminawung JA, Shah ND, Krumholz HM, Ross JS. Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005–2012. *JAMA*. 2014;311:368–377. [PubMed: 24449315]
6. Blaker H Confidence curves and improved exact confidence intervals for discrete distributions. *Can J Stat*. 2000;28:783–798.
7. Fay MP. Confidence intervals that match Fisher's exact or Blaker's exact tests. *Biostatistics*. 2010;11:373–374. [PubMed: 19948745]
8. Fay MP. Two-sided exact tests and matching confidence intervals for discrete data. *R Journal*. 2010;2:53–58.

9. Wang S, Laitinen-Parkkonen P. Efficacy assessment in paediatric studies. *Handb Exp Pharmacol*. 2011;205:149–168. [PubMed: 21882110]
10. Sun H, Bastings E, Temeck J, et al. Migraine therapeutics in adolescents: a systematic analysis and historic perspectives of triptan trials in adolescents. *JAMA Pediatr*. 2013;167:243–249. [PubMed: 23359002]
11. Shaddy RE, Olsen SL, Bristow MR, et al. Efficacy and safety of metoprolol in the treatment of doxorubicin-induced cardiomyopathy in pediatric patients. *Am Heart J*. 1995;129:197–199. [PubMed: 7817916]
12. Demir R, Kucukoglu MS. Six-minute walk test in pulmonary arterial hypertension. *Anatol J Cardiol*. 2015;15:249–254. [PubMed: 25880178]
13. Farber HW, Miller DP, Mcgoon MD, Frost AE, Benton WW, Benza RL. Predicting outcomes in pulmonary arterial hypertension based on the 6-minute walk distance. *J Heart Lung Transplant*. 2015;34:362–368. [PubMed: 25312386]
14. Fritz JS, Blair C, Oudiz RJ, Dufton C, Olschewski H, Despain D, et al. Baseline and follow-up 6-min walk distance and brain natriuretic peptide predict 2-year mortality in pulmonary arterial hypertension. *Chest*. 2013;143:315–323. [PubMed: 22814814]
15. Cappelleri JC, Hwang LJ, Mardekian J, Mychaskiw MA. Assessment of measurement properties of peak VO₂ in children with pulmonary arterial hypertension. *BMC Pulm Med*. 2012;12:54. [PubMed: 22963001]
16. Zijlstra WM, Ploegstra MJ, Vissia-Kazemier T, et al. Physical activity in pediatric pulmonary arterial hypertension measured by accelerometry: a candidate clinical endpoint. *Am J Respir Crit Care Med* 2017;196:220–227. [PubMed: 28178426]
17. Food US and Drug Administration: Challenge and Opportunity on the Critical Path to New Medical Products. <https://www.fda.gov/downloads/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/ucm113411.pdf>.
18. Moustgaard H, Bello S, Miller FG, Hrobjartsson A. Subjective and objective outcomes in randomized clinical trials: definitions differed in methods publications and were often absent from trial reports. *J Clin Epidemiol*. 2014;67:1327–1334. [PubMed: 25263546]
19. Kahan BC, Dore CJ, Murphy MF, Jairath V. Bias was reduced in an open-label trial through the removal of subjective elements from the outcome definition. *J Clin Epidemiol*. 2016;77:38–43. [PubMed: 27262238]
20. Kovacs SM, Turner-Bowker DM, Calarco G, Mulberg AE, Paty J. Practical considerations for the use of clinical outcome assessments (COAs) in pediatric clinical research: examples from pediatric gastroenterology. *Ther Innov Regul Sci*. 2016;50:37–43.
21. Tomlinson G, Detsky AS. Composite end points in randomized trials: there is no free lunch. *JAMA*. 2010;303:267–268. [PubMed: 20085955]
22. US Food and Drug Administration: Multiple Endpoints in Clinical Trials: Guidance for Industry. <https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM536750.pdf>. Accessed June 18, 2017.
23. Anker SD, Schroeder S, Atar D, Bax JJ, Ceconi C, Cowie MR, et al. Traditional and new composite endpoints in heart failure clinical trials: facilitating comprehensive efficacy assessments and improving trial efficiency. *Eur J Heart Fail*. 2016;18:482–489. [PubMed: 27071916]
24. Del Prato S, Fleck P, Wilson C, Chaudhari P. Comparison of alogliptin and glipizide for composite endpoint of glycated haemoglobin reduction, no hypoglycaemia and no weight gain in type 2 diabetes mellitus. *Diabetes Obes Metab*. 2016;18:623–627. [PubMed: 26865535]
25. Bulger EM, May A, Dankner W, Maislin G, Robinson B, Shirvan A. Validation of a clinical trial composite endpoint for patients with necrotizing soft tissue infections. *J Trauma Acute Care Surg*. 2017;83:622–627. [PubMed: 28538644]
26. Green DJ, Liu XI, Hua T, et al. Enrichment strategies in pediatric drug development: an analysis of trials submitted to the US Food and Drug Administration [published online ahead of print 2017]. *Clin Pharmacol Ther*. <https://doi.org/10.1002/cpt.971>

Figure 1.

Success and failure of pediatric trials when the trials were combined with adult efficacy trials. Shown are the number of pediatric trials that were combined (yes) and were not combined (no) with adult trials, the upper and lower bounds of the 95% confidence intervals (CIs), and a graphical depiction of the confidence intervals.

Figure 2.

Success and failure of pediatric trials when the trials used the same endpoint as the adult trials. Shown are the number of pediatric trials that used the same endpoint (same) and used a different endpoint (different) as the adult trials, the upper and lower bounds of the 95% confidence intervals (CIs), and a graphical depiction of the confidence intervals.

Figure 3.

Success and failure of pediatric trials that used a different endpoint than the adult endpoint, broken down to differences in the endpoint or differences in the timing of the evaluation. Shown are the number of pediatric trials that were different based on endpoint, different based on timing of the evaluation, and different based on both endpoint and timing (both); the upper and lower bounds of the 95% confidence intervals (CIs); and a graphical depiction of the confidence intervals.

Figure 4.

Success and failure of pediatric trials when the trial had a subjective endpoint, an objective endpoint, or used a composite endpoint of both a subjective and an objective endpoint (both). Shown are the number of pediatric trials that were subjective, objective, or both, the upper and lower bounds of the 95% confidence intervals (CIs), and a graphical depiction of the confidence intervals.

Table 1.**Breakdown of Results by Therapeutic Area**

	Label Outcome ^a		Endpoint Type			Same Endpoint ^b	
	Approved	Not Approved	Objective	Subjective	Both	Yes	No
Allergy	24	3	0	27	0	25	2
Analgesia/Anesthesia	6	4	3	7	0	6	4
Anti-infectives	15	1	9	1	6	16	0
Antivirals	33	0	32	0	1	11	22
Cardiology/renal	4	3	7	0	0	2	5
Dermatology	14	0	4	1	9	14	0
Gastrointestinal	12	5	5	10	2	8	9
Hematology	4	4	8	0	0	3	5
Inborn errors of metabolism	1	0	1	0	0	0	1
Metabolic/endocrine	6	0	6	0	0	2	4
Neurology	17	2	4	15	0	12	7
Oncology	3	7	10	0	0	10	0
Ophthalmology	0	2	1	1	0	1	1
Psychiatry	17	8	0	25	0	3	22
Pulmonary	28	9	24	12	1	28	9
Rheumatology	2	0	0	0	2	0	2
Total	186	48	114	99	21	141	93

^aDrugs that were approved only in a subset of the pediatric patients studied were considered approved.

^bPediatric versus adult trials for the same indication.