Supplemental Online Content

Williams DJ, Creech CB, Walter EB, et al; DMID 14-0079 Study Team. Short- vs standard-course outpatient antibiotic therapy for community-acquired pneumonia in children: the SCOUT-CAP randomized clinical trial. *JAMA Pediatr*. Published online January 18, 2022. doi:10.1001/jamapediatrics.2021.5547

eMethods

eResults

This supplemental material has been provided by the authors to give readers additional information about their work.

eMethods

S1. Study Population

Potentially eligible participants included otherwise healthy children 6-71 months of age diagnosed with uncomplicated CAP and prescribed either amoxicillin/amoxicillin-clavulanate or cefdinir in an ambulatory care setting at one of eight study sites (University of Alabama at Birmingham, University of Arkansas, Cincinnati Children's Hospital Medical Center, Duke University Health System, Children's Hospital of Philadelphia, UPMC Children's Hospital of Pittsburgh, Vanderbilt University Medical Center [coordinating site], and Washington University). Patient screening began on 02 December 2016. New enrollments ceased when our enrollment goal was met, and the last subject's final visit occurred on 16 December 2019.

Children with severe pneumonia, defined as: need for hospitalization; radiographic evidence of clinically significant parapneumonic effusion, empyema, lung abscess, or pneumatocele; or microbiologically confirmed *Staphylococcus aureus* or *Streptococcus pyogenes* pneumonia were excluded. In addition, children receiving either parenteral or combination antibiotic therapy, those undergoing surgery or invasive airway procedures requiring general anesthesia or hospitalization within 7 days prior to the diagnosis of CAP, those with known beta-lactam allergy, those with concomitant bacterial infection necessitating >5 days of antibiotic therapy (e.g., otitis media), and those with a provider diagnosis of aspiration pneumonia, bronchiolitis, bronchitis, or acute asthma exacerbation were also excluded. Finally, we excluded children with underlying chronic medical conditions (except children with asthma who had no bronchodilator or inhaled corticosteroid use in the preceding 6 months), those with a history of pneumonia within the prior 6 months, and those previously enrolled in the trial or concurrently enrolled in another trial of an investigational agent.

Enrollment Procedures

At enrollment, study staff interviewed participants' caregivers, performed a physical assessment to collect demographics, medical history, and concomitant medications; performed a physical assessment, including general appearance and hydration status, temperature, pulse, and respiratory rate, work of breathing, and presence of skin rash; dispensed study medication; and provided caregivers with a memory aid to record administration of study drug and concomitant medications, temperature, antibiotic adverse effects, and presence of cough.

Outcome assessment visit 1 (OAV1) occurred on Study Day 6-10, and outcome assessment visit 2 (OAV2) occurred on Study Day 19-25. At each visit, staff collected interval medical history and concomitant medications, performed a physical assessment, reviewed the memory aid and study product administration, and assessed clinical response, resolution of symptoms, and antibiotic adverse effects.

Randomization and Blinding

Enrollment of subjects was performed online using AdvantageEDC. Randomization to short vs. standard course therapy occurred at a 1:1 ratio. Subjects were stratified by age group <24 months vs. 24-71 months), type of initial antimicrobial therapy, and initial treatment in an ED or outpatient clinic/urgent care center. The list of randomized treatment assignments was prepared by statisticians at The Emmes Corporation and included in The Emmes Corporation's Internet Data Entry System (IDES). IDES assigned each volunteer a treatment code and a designated individual at each site was provided with the key.

The study subjects and their parents/guardians, investigators, and study team staff were blinded to study treatment assignment throughout the study. The subjects and their families, investigators, and study team staff were not blinded to which of the three antibiotics (amoxicillin, amoxicillin-clavulanate, cefdinir) the subject was initially prescribed. The study products and placebo was prepared by the unblinded site Research Pharmacist. Only the preparing pharmacist was aware of the study product bottle assignments. For subjects randomized to standard course therapy, the pharmacy provided the same medication prescribed initially. For subjects randomized to short course therapy, the pharmacy provided a placebo that resembles the appearance (color and texture), flavor, and consistency of the active study product. All study products were packaged with an identical appearance.

The study product was labeled with a numerical code that masks site investigators, site staff, parent(s)/guardian(s) and children to the formulation. The unblinded site Research Pharmacist was the only person to perform the unmasking, if needed.

Desirability of Outcome Ranking (DOOR) Definition

An 8-level ordinal outcome was created from three components: adequate clinical response, resolution of symptoms, and the presence and severity of solicited events. Adequate clinical response was defined as the absence of a medically attended visit (ED, outpatient clinic, or hospitalization), surgical procedure (e.g., pleural drainage), or receipt of one or more doses of a non-study antibiotics for persistent or worsening pneumonia that occurred after randomization and receipt of at least one dose of study drug. Resolution of symptoms was defined as the absence of all of the following: fever in the preceding 24 hours, elevated respiratory rate, and moderate or severe cough. Antibiotic adverse effects were defined as irritability, vomiting, diarrhea, allergic reaction, stomatitis, or candidiasis, and were graded as mild, moderate, or severe.

Resistome Substudy

DNA was extracted from throat swabs obtained from children who consented to have them collected at OAV2 using the PureLink™ Microbiome DNA Purification Kit (Invitrogen Carlsbad, CA). Shotgun sequencing libraries were prepared using the NEBNext® Ultra II FS DNA Library Prep Kit with Sample Purification Beads for Illumina® following the protocol for inputs ≤ 100 ng. Adaptor–ligated DNA was by enriched using eight PCR cycles and the NEBNext® Multiplex Oligos for Illumina. Individual libraries were pooled at equal ng amounts and submitted for sequencing to the Yale Center of Genome Analysis on the Illumina HiSeq 4000 using a paired-end 150 bp protocol. Btrim software was used to remove adaptors and low-quality sequence regions from the shotgun metagenomic sequence reads.[1] The Antibiotic Resistance Genes (ARG) Online Analysis Pipeline v2.0 with the expanded Structured Antibiotic Resistance Genes Database was used to classify and quantify ARGs.[2] Detected ARGs were normalized against prokaryotic cell numbers and ARG abundances provided as resistance genes per prokaryotic cell (RGPC).

Statistical Analysis

The primary efficacy analysis compared the two treatment strategies using a Mann-Whitney U Test. An estimate of the probability of a more desirable outcome in the short course arm (plus one-half the probability of tied desirability), and its 95% confidence interval (CI) were obtained through the inversion of the F-test. This approach was also carried out for the secondary efficacy endpoints of RADAR at OAV2 and DOOR at OAV1 and OAV2. Each of these analyses were carried out under the Intention-to-Treat (ITT) principle with analyses performed per randomized treatment assignment. Randomized participants who became ineligible before Study Day 1 were excluded from ITT analyses. Subjects randomized but not treated for reasons other than ineligibility were included in the ITT analyses. In ITT analyses, DOOR was imputed for subjects that were not treated, terminated early from the study, or completed the visit but were missing adequate clinical response and its components. Whenever possible, data collected post-randomization for a subject with missing DOOR in the ITT analyses to assist in imputing the missing DOOR. Multiple imputation with linear models was used to impute values using all available information, assuming a missing at random model.

Efficacy analyses for DOOR and RADAR at OAV1 and OAV2 were repeated in *a priori*-defined sensitivity analyses including complete case (CC), according-to-protocol (ATP), and worst-case analysis sets. For the CC analyses, participants were analyzed as randomized, but participants with missing data that prevented assignment of an unambiguous value to the primary endpoint, and those not receiving at least one dose of study product were excluded. For the ATP analyses, participants were analyzed as treated and included participants from the CC analyses who had no major protocol deviations and who received at least one daily dose of study drug on Study Days 1-5. In the worst-case analysis, all imputations of missing data used the worst case (i.e., the least desirable possible DOOR given available information) for subjects in the 5-day arm and best case for subjects in the 10-day arm. For the analysis of the effect of increasing the minimum difference in the duration of antibiotic use, the probability of a more desirable RADAR due to assignment to the short antibiotic course was computed along with its 95% CI for different values of k.

Finally, using the CC analysis sets for OAV1 and OAV2, lack of resolution of symptoms, occurrence of medically attended visits or surgical procedures, receipt of non-study antibiotics, and risk of mild, moderate, or severe solicited events were compared between the two treatment strategies. Risk differences were compared using Fisher's exact tests. The Newcombe method with continuity correction was used to compute 95% confidence intervals for risk differences.

Sample size was based on a superiority test. The null hypothesis was that the probability of a more desirable RADAR plus one-half the probability of a tied RADAR was 50%. The alternative hypothesis was that the

probability of a more desirable RADAR plus one-half the probability of a tied RADAR was 60%. A sample size of 360 (180 per strategy) provided 90% power using a 2-sided α=0.05. The sample size was inflated by 10% to account for loss to follow-up resulting in a total enrollment target of 400 participants.

Interim Analysis of Efficacy, Futility, and Safety

One interim analysis was performed and reported to the DSMB after at least 30% of the targeted subjects completed the study to inform DSMB decisions regarding stopping early for efficacy, futility, or safety. For the interim analysis, a snapshot of the study database was unblinded and used to conduct analyses as follows. An ITT analysis including all enrolled subjects in the snapshot of the study database was performed testing the primary aim's null hypothesis using the Haybittle-Peto boundary ($p<0.001$) to determine statistical significance. A 95% confidence interval for the probability that a randomly selected subject will have a better DOOR if assigned to the 5-day strategy (vs. the standard strategy) was also estimated but was not used to inform DSMB decisions about stopping early for efficacy. Predicted interval plots were constructed to provide the DSMB with a prediction of the trial results were the trial to continue as planned under varying assumptions regarding future data (e.g., current trend continues, null hypothesis is true, alternative hypothesis is true).

Resistome Analyses

The endpoint for the resistome analyses was the number of RGPC in throat swabs collected at OAV2. We compared 1) beta-lactamase RGPC, and 2) the total RGPC, which was calculated as the sum of RGPC across all 24 antibiotic classes per sample. The null hypothesis was that the distribution of the number of RGPC in the two study arms would be the same. The alternative hypothesis was that children randomized to the 5-day arm would have fewer resistance genes detected. For the ITT analysis set, a one-sided Wilcoxon Rank Sum test was conducted to assess statistically significant differences (alpha level <0.05) between the number of beta-lactamase RGPC as well as the total number of RGPC between treatment groups. Fisher's exact test was used to evaluate differences in baseline characteristics used to stratify participants for randomization and/or known to impact the microbiota (e.g., age group, sex, antibiotic, and initial site of treatment).

Sample size calculations were based on preliminary data on the distribution of beta-lactamase genes and total antibiotic resistance genes from 57 throat samples that were blinded to treatment group assignment. Using a one-sided α=0.05, a sample size of 200 total subjects (100 in each group) would achieve 75% power to detect a difference between the two groups.

References

1. Kong Y. Btrim: a fast, lightweight adapter and quality trimming program for next-generation sequencing technologies. Genomics 2011;98:152-3.
2. Yin X, Jiang XT, Chai B, et al. ARGs-OAP v2.0 with an expanded SARG database and Hidden Markov Models for enhancement characterization and quantification of antibiotic resistance genes in environmental metagenomes. Bioinformatics 2018;34:2263-70.

eResults

S2. Results of Sensitivity Analyses

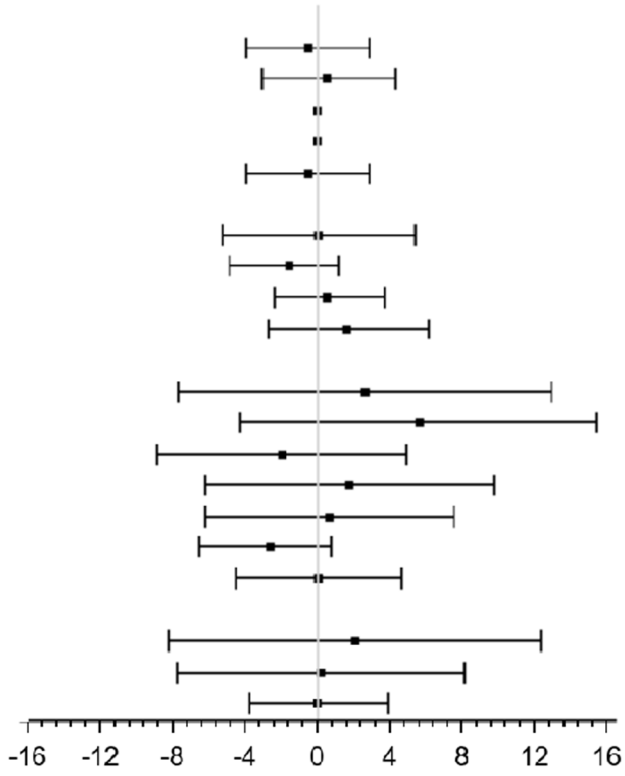
Sensitivity analyses using CC and ATP analysis populations, and a worst-case analysis where the ITT population had missing data imputed in the most favorable way for standard course subjects and the least favorable way for short course subjects, were consistent with the primary analysis for OAV1 (range of probabilities of a more desirable RADAR for short course strategy at OAV1: 0.62-0.69; $p<.001$ for all). Results for OAV2 were also similar for the CC and ATP analyses (range of probabilities of a more desirable RADAR for short course strategy at OAV1: 0.63-0.65; $p<.001$ for both). However, the results from the worst-case analysis at OAV2 did not show evidence to conclude that the short course subjects had a probability of a more desirable RADAR (0.53; $p=0.392$).

Analysis of the effect of increasing the minimum difference in the duration of antibiotic use ($k=2,3,4$, or 5, or infinity) before a favorable response is given to the subject with shorter duration of antibiotic use were consistent to those from the primary analysis when $k=1,2,3,4,5$ (range of probabilities of a more desirable RADAR for short course strategy at OAV1: 0.68-0.69; $p<.001$ for all). Note that the RADAR analysis when $k=$infinity is equivalent to the analysis of DOOR without regard for number of days of antibiotic use.

S3. Inadequate Clinical Response, Persistent Symptoms, and Antibiotic Adverse Effects, ITT Population at OAV2

Frequencies (%) by treatment strategy and risk differences (95% confidence intervals) for each component of the Desirability of Outcome Ranking (DOOR). Risk differences were compared using Fisher's exact tests. The Newcombe method with continuity correction was used to compute 95% confidence intervals for risk differences.

	Short Course (N=189) n (%)	Standard Course (N=191) n (%)
Inadequate Clinical Response		
Any	2 (1)	3 (2)
ED or Clinic Visit	4 (2)	3 (2)
Hospitalization	-	-
Surgical Procedure	-	-
Receipt of Non-Study Antibiotic	2 (1)	3 (2)
Persistent symptoms		
Any	11 (6)	11 (6)
Fever	0 (0)	3 (2)
Elevated Respiratory Rate	2 (1)	1 (<1)
Cough	8 (4)	5 (3)
Antibiotic Adverse Effects		
Any	96 (51)	92 (48)
Irritability	72 (38)	62 (32)
Vomiting	20 (11)	24 (13)
Diarrhea	34 (18)	31 (16)
Allergic reaction	22 (12)	21 (11)
Stomatitis	1 (<1)	6 (3)
Candidiasis	7 (4)	7 (4)
Severity of Antibiotic Adverse Effects		
Mild	91 (48)	88 (46)
Moderate	32 (17)	32 (17)
Severe	4 (2)	4 (2)



S4. Desirability of Outcome Ranking (DOOR), ITT Population at OAV2

DOOR (Rank) Description	Short Course (N=189) n (%)	Standard Course (N=191) n (%)	Cumulative Risk Difference (95% CI)
(1) ACR with No Antibiotic Adverse Effects	73 (39)	81 (42)	-3.8 (-13.9, 6.4)
(2) ACR with Mild Antibiotic Adverse Effects	53 (28)	51 (27)	-2.4 (-12.1, 7.3)
(3) ACR with Moderate Antibiotic Adverse Effects	25 (13)	20 (10)	0.3 (-8.2, 8.8)
(4) ACR with Severe Antibiotic Adverse Effects	3 (2)	4 (2)	-0.2 (-8.4, 8.0)
(5) ACR with Persistent Symptoms	9 (5)	8 (4)	0.4 (-7.0, 7.8)
(6) No ACR with ED or clinic visit	2 (1)	3 (2)	-0.1 (-7.3, 7.0)
(7) No ACR with hospitalization	0	0	-0.1 (-7.3, 7.0)
(8) Death[3]	0	0	-0.1 (-7.3, 7.0)

DOOR Rank (1) represents the best possible outcome and DOOR Rank (8) the worst; Cumulative risk differences were calculated as follows: for $i \; \Sigma$ [1-8], the difference in proportions of participants between treatment strategies with DOOR rank ≤ i was calculated and 95% confidence intervals were estimated using the Newcombe method with continuity correction. Abbreviations: OAV, Outcome Assessment Visit; CI, Confidence Interval; ACR, Adequate Clinical Response; ED, Emergency Department

S5. Resistome Profiles (Antibiotic Resistance Genes per Prokaryotic Cell) by Treatment Strategy

A one-sided Wilcoxon Rank Sum test was used to assess statistically significant differences (alpha level <0.05) between the number of beta-lactam and total RGPC in participants assigned to a short course or standard course strategy. Log_2 normalized RGPC data were used for visualization. Abbreviations: OAV, Outcome Assessment Visit