A Prospective Trial of Withdrawal and Reinstitution of Ursodeoxycholic Acid in Pediatric Primary Sclerosing Cholangitis.

Saul Karpen, Emory University
DD Black, Pediatrics University of Tennessee Health Science Center Memphis TN.
C Mack, Pediatrics University of Colorado School of Medicine Aurora CO.
N Kerkar, Pediatrics University of Rochester Medical Center Rochester NY.
T Miloh, Pediatrics Baylor College of Medicine Houston TX.
SS Sundaram, Pediatrics University of Colorado School of Medicine Aurora CO.
R Anand, Emmes Corporation Rockville MD.
A Gupta, Emmes Corporation Rockville MD.
E Alonso, Pediatrics Northwestern University College of Medicine Chicago IL.
R Arnon, Pediatrics Mount Sinai Icahn School of Medicine New York NY.

Only first 10 authors above; see publication for full author list.

Journal Title: Hepatol Commun
Volume: Volume 3, Number 11
Publisher: (publisher) | 2019-11, Pages 1482-1495
Type of Work: Article
Publisher DOI: 10.1002/hep4.1421
Permanent URL: https://pid.emory.edu/ark:/25593/v4q8d

Final published version: http://dx.doi.org/10.1002/hep4.1421

Accessed February 13, 2020 2:23 PM EST
Ursodeoxycholic acid (UDCA) is commonly used to treat several liver disorders in adults and children, including primary sclerosing cholangitis (PSC) for which it is not U.S. Food and Drug Administration approved. UDCA treatment has an uncertain impact on disease outcomes and has been reported in high doses to be associated with worse outcome in adults with PSC. In this context, controlled withdrawal and reintroduction of UDCA in children with PSC were studied. Prior to study initiation, participants were required to have alanine aminotransferase (ALT) and gamma-glutamyl transpeptidase (GGT) <2 times the upper limit of normal on stable UDCA dosing. The study included four phases: I (stable dosing), II (50% UDCA reduction), III (UDCA discontinuation), IV (UDCA reintroduction), with a primary endpoint of change in ALT and GGT between phases I and III. We enrolled 27 participants (22 completed) between March 2011 and June 2016. Changes in mean ALT and GGT between phases I and III were ALT, +29.5 IU/L (P = 0.105) and GGT, +60.4 IU/L (P = 0.003). In 7 participants, ALT and GGT ≤29 IU/L (null response group). Eight participants had increases of ALT or GGT >100 IU/L (flare group). None developed elevated bilirubin. All flares responded to UDCA reinstatement. Serum GGT, interleukin-8, and tumor necrosis factor α levels were higher in the flare group at baseline. Liver biochemistries increased in children with PSC during controlled UDCA withdrawal; one third increased above 100 IU/L and one third remained normal during UDCA withdrawal. Conclusion: The impact of prolonged UDCA use in childhood PSC and the significance of a biochemical flare are unclear. Further studies of the natural history and treatment of pediatric PSC and UDCA use are needed.
of portal hypertension developed in one third of children with PSC within 10 years of diagnosis. 

Cholangiocarcinoma, while quite rare in childhood, has been described in adolescents with PSC. 

There are many controversial areas in pediatric PSC, including optimal diagnostic criteria, predictors of outcome, and therapeutic approaches. Diagnostic criteria have not been well-delineated in pediatric PSC but are suggested to include biochemical features of biliary disease (e.g., elevated gamma-glutamyl transpeptidase [GGT]), histologic findings consistent with cholangiocyte-directed injury, and/or cholangiographic evidence of biliary injury. None of these features have been the subject of a rigorous analysis of reliability and reproducibility in children. Overlap of PSC with autoimmune hepatitis (AIH), present in 30%-50% of pediatric patients with PSC, also referred to as autoimmune sclerosing cholangitis, may be an important factor in disease progression and severity. Given the progressive nature of pediatric PSC, there is significant biomedical relevance in identifying prognostic biomarkers of future progression and developing evidence-based therapies that can slow evolution to end-stage liver disease.

Effective therapeutic approaches to the management of pediatric PSC remain uncertain given the lack of reliable treatment data. Immunosuppression is presumed to be important for children with autoimmune overlap. Ursodeoxycholic acid (UDCA) has been a mainstay of therapy in pediatrics, although quality of evidence supporting its long-term use and benefits is lacking. Vancomycin has also been suggested to have therapeutic value in PSC in relatively small studies, and the overall experience with this drug is significantly less than with UDCA. Because of the paucity of data, clinicians, patients, and families are eager to employ any therapeutic regimen that may slow the progression of the disease. In this context, anecdotal approaches are often implemented to emphasize the critical need to avoid potential harm of unproven therapies.

The finding of potential toxicity of “high-dose” UDCA therapy in adults with PSC has led to a vexing problem in recent approaches to disease management. In light of this unexpected finding, the American Association for the Study of Liver Diseases recommended against the use of UDCA in adults with PSC. This recommendation was not extended to children but raised questions about the safety of a therapy in children that seems efficacious, at least from a biochemical perspective. Withdrawal of UDCA therapy in adults with PSC resulted in statistically significant worsening of liver biochemistries 3 months after withdrawal. In that study, consecutive adult patients with PSC were enrolled regardless of their response.
to UDCA therapy. In 2010, shortly after the report of potential UDCA toxicity, the present pediatric UDCA withdrawal study was initiated (Withdrawal of Ursodeoxycholic Acid Therapy in Pediatric Primary Sclerosing Cholangitis [WUPPSC, NCT01088607]). Clinically, study investigators had observed biochemical disease flares in adolescents who were nonadherent to UDCA for PSC; this improved with reinitiation of UDCA treatment. As such, WUPPSC was designed to assess the potential utility of UDCA as a therapy in pediatric PSC. The study focused on participants who were biochemically responsive to UDCA, as reflected by the requirement of alanine aminotransferase (ALT) and GGT less than 2 times the upper limit of normal (ULN) at study entry, thereby potentially enhancing the ability to identify a therapeutic effect. Unlike the adult study, the design involved both a staged withdrawal of therapy and a monitored reinstatement of UDCA in pediatric patients with PSC with mild and stable disease. The results of this novel investigation are reported.

Participants and Methods

PARTICIPANTS

Twenty-seven previously diagnosed male and female participants with PSC, all less than 21 years of age without ethnic or racial restrictions, were recruited from 12 participating pediatric centers between March 2011 and June 2016. Diagnostic criteria for PSC in this study included two of the following three features: 1) serum GGT increased more than 50% above the ULN for age; 2) endoscopic retrograde cholangiopancreatography, percutaneous transhepatic cholangiogram, or magnetic resonance cholangiopancreatography findings of intrahepatic and/or extrahepatic bile duct irregularities consistent with PSC; and 3) liver biopsy abnormalities consistent with chronic biliary injury. Patients with PSC/AIH overlap who met the criteria for PSC plus had liver histologic features of AIH were included.

In addition to the diagnosis of PSC, participants were also required to have the following: 1) biochemically quiescent liver disease defined by an ALT and GGT less than 2.0 times ULN measured on two separate occasions greater than 2 weeks apart; 2) prior and ongoing UDCA therapy at a dose of greater than 13 mg/kg/day or 600 mg/day for more than 6 months by commercially available pill form (compounded liquid preparations were avoided due to bioavailability concerns); 3) ability to swallow pills; and 4) quiescent inflammatory bowel disease (IBD) as reflected by revised Pediatric Ulcerative Colitis Activity Index or Short Pediatric Crohn’s Disease Activity Index scores (see Supporting Materials). Major exclusion criteria included (see also Supporting Materials for details) any other primary liver disease, advanced liver disease and cirrhosis with decompensation, any evidence of portal hypertension with palpable spleen and hypersplenism being used as surrogate markers, previous injury/surgery of the biliary tract, and any systemic disease that secondarily involved the biliary system.

PROTOCOL

The study protocol consisted of four phases (Supporting Fig. S1). During phase I (4 weeks), participants were maintained on UDCA at their baseline clinically prescribed dosing and were monitored for stability. In phase II (4 weeks), the UDCA dose was reduced by 50% and then stopped completely in phase III (8 weeks) with continued monitoring. Dosing was restarted in phase IV (8 weeks) at a standardized dose of as close to but not to exceed 20 mg/kg/day in two divided doses. Baseline and endpoint visits occurred at weeks 0, 16, and 24, with face-to-face surveillance visits at weeks 4, 8, 12, and 20. Participants were maintained on all other clinically prescribed medications for their liver disease and IBD, such as immunosuppressives, during the course of the study. Surveillance visits included review of any new symptoms, study adherence by pill count, adverse events, and collection/review of the following laboratory tests: ALT, aspartate aminotransferase (AST), total bilirubin (TB), direct bilirubin (DB), alkaline phosphatase (ALP), GGT, albumin, and prothrombin time/international normalized ratio (PT/INR). Endpoint evaluations included full history, physical examination, monitoring for adverse events, and collection of plasma and serum for inflammatory biomarkers, bile acid analysis, ALT, AST, TB, DB, ALP, GGT, albumin, PT/INR, platelet count, white blood cell count with differential, and hematocrit/hemoglobin.

Participants were monitored for significant biochemical worsening of their liver disease, referred to as
a significant disease flare (SDF) and defined by a rise in ALT or GGT to greater than 10 times ULN, rise in DB to greater than 5 times entry level, or clinical manifestations of acute liver failure or decompensated cirrhosis. If an SDF was detected, the participant was immediately placed back on UDCA at the dose prescribed by the study protocol as above, moved to phase IV of the study, and the event was reported as a serious adverse event (SAE). Also, if the ALT or GGT was greater than 5 times ULN at any laboratory draw, a 2-week follow-up surveillance blood draw could be ordered at the discretion of the site investigator.

We hypothesized that participants would exhibit a biochemical flare during UDCA withdrawal that would resolve when placed back on the drug. The primary endpoints for the study were the change in ALT and GGT at the end of phase III compared to the beginning of the baseline phase I period (drug withdrawal). The secondary endpoints were the change in ALT and GGT at the end of phase IV compared to the end of phase III (drug reinstitution) and changes in inflammatory cytokines at the end of phase III compared to either baseline (withdrawal) or to the end of phase IV (reinstitution).

LABORATORY TESTING

All blood samples for chemistries, hematology, and coagulation testing were collected at study sites and shipped to a central laboratory (Q² Solutions, Valencia, CA) for automated analysis. Results were reported to centers within 48 hours, then entered into the secure online study database at the Data Coordinating Center (Emmes Corporation, Rockville, MD).

INFLAMMATORY CYTOKINE MEASUREMENT

Baseline and endpoint plasma samples were shipped to the Clinical Laboratory Improvement Amendments-approved Pediatric Clinical and Translational Research Center, University of Colorado, for cytokine measurement using Luminex multi-analyte technology (Invitrogen, Thermo Fisher Scientific, Carlsbad, CA). Analysis included interleukin (IL)-1β, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, IL-15, tumor necrosis factor (TNF)-α, interferon (IFN)-γ, granulocyte-macrophage colony-stimulating factor (GM-CSF), macrophage inflammatory protein (MIP)-1β, and eotaxin. Individual cytokine analyses were performed in duplicate, and the background signal (saline only) was subtracted for each readout. Results are shown as mean of duplicate samples.

BILE ACID DETERMINATION

Bile acid analysis to determine UDCA levels was carried out on baseline and endpoint plasma samples. Whole blood was collected in heparin tubes, and plasma was isolated and stored at −80°C. UDCA was quantified by an ultra-high-performance liquid chromatography–tandem mass spectrometry (UPLC/MS-MS) assay with modifications. Bile acids were first isolated from plasma with a Strata-X C18 solid-phase extraction column (#8B-S100-TAK; Phenomenex). Standard solutions of ursodeoxycholate (C1020, CAS# 128-13-2; Steraloids) were used to generate a standard curve (0-2.5 µM) to determine UDCA concentration. Separation was achieved using UPLC (Waters Corporation) followed by MS-MS (TQD Tandem Mass Spectrometer; Waters Corporation) with a Waters Cortecs C18 + UPLC column (#186007367; Waters). Analyte separation and elution were achieved using a 10 mM ammonium acetate, 0.15% ammonium hydroxide aqueous mobile phase and a methanol/10 mM ammonium acetate 0.15% ammonium hydroxide elution phase. We used d4-ursodeoxycholate (#D3819; CDN, Inc.) as an internal control in each sample.

ETHICAL CONSIDERATIONS

Informed consent in writing was obtained for each participant, and the study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki as reflected by a priori approval by the institutional review committees of each participating center. The study was monitored by a nonconflicted external data and safety monitoring board.

STATISTICAL ANALYSIS

All participants who completed the study were included in the final statistical analyses. All analyses were performed for all such participants for the group as a whole. Further ad hoc analyses included stratification of the participants according to the following three subgroup categories defined based on the UDCA response of ALT and GGT parameters: null, ALT, and GGT persistently ≤29 IU/L (i.e.,
normal); flare, ALT, and/or GGT >100 IU/L during phase II and/or III; indeterminant, ALT, and/or GGT >29 IU/L and <100 IU/L during phase II and/or III.

Participants’ demographics and other baseline characteristics (i.e., disease diagnosis groups, time since diagnosis, UDCA dosing, IBD diagnosis, immunosuppressive medication, all clinical laboratory and cytokine parameters) were summarized in aggregate and by UDCA response subgroup categories using descriptive statistics, including mean and SD for continuous outcomes and frequencies and percentages for categorical outcomes. Differences in participants’ characteristics and laboratory values at baseline among the groups were tested using analysis of variance (ANOVA) for continuous measures and chi-square or Fisher exact test if expected cell sizes were less than five for categorical measures. A pairwise comparison between the groups was done if the overall test comprising all the subgroups showed any significantly different results.

To assess the primary endpoint of the study, ALT and GGT values at the end of phase III were compared with those at the beginning of the baseline phase I in aggregate and by post hoc analysis of UDCA response subgroups using paired $t$ test with no adjustment for multiplicity given the exploratory nature of the analysis. A similar method was used for secondary endpoint analyses to provide statistical comparison between the end of phase III and end of phase IV values. These analyses were replicated for a more conservative estimate of phase III and phase IV data, where maximum of ALT and GGT values across the visits within these phases were considered for the analysis. Spaghetti plots for ALT and GGT by UDCA response subgroups for all the visits were generated to review participants’ responses to UDCA treatment and withdrawal across all the visits and phases of the study.

For all the statistical analyses above, $P < 0.05$ was considered significant. All the analyses were performed in SAS version 9.4 (SAS Institute Inc., Cary, NC).

Results

STUDY PARTICIPANTS

Participant demographics and clinical characteristics in aggregate and by UDCA response group are shown in Table 1. Most of the baseline characteristics were similar among treatment response groups. A majority were male participants (77.3%), with a mean age of 13.9 + 3.3 (mean ± SD) years at entry into the study and 12.0 + 3.7 years at diagnosis. A slight majority (54.5%) had AIH/PSC overlap, 86.4% had IBD (63.2% ulcerative colitis, 15.8% Crohn’s disease, and 21.0% indeterminant), and 40.9% were on immunosuppressive medication at entry to the study. Two diagnostic criteria for PSC were met by 81.8% of participants, and 18.2% met all three (Supporting Table S1). Baseline laboratory values are shown in Table 2. All had quiescent PSC as defined by GGT and ALT < 2.0 times ULN. None of the participants had laboratory evidence of hypersplenism or decompensated liver disease.

STUDY COURSE

Of the total of 27 participants recruited for the study, 22 completed all phases of the study and constituted the primary analysis group (Supporting Fig. S2). There were five early withdrawals: one due to an ulcerative colitis flare, two due to study protocol nonadherence, one with an out of range baseline laboratory value, and one withdrawal reason was not disclosed by the participant. There were three SAEs reported, including two SDF events and one IBD flare. Both participants with SDF were placed back on UDCA and moved to phase IV, with subsequent remission of their flare.

GGT and ALT levels in the three response groups during the course of the study are depicted in Fig. 1. Seven participants had baseline levels of ALT and GGT ≤29 IU/L that did not rise above 29 IU/L during the 24 weeks of the study (null response group, Fig. 1A,B). Eight participants had increases of ALT or GGT >100 IU/L on either reduced or no UDCA therapy (flare response group, Fig. 1C,D). Seven participants had intermediate responses (indeterminant response group, Fig. 1E,F). All flares responded to UDCA reinstitution by the end of the study. One subject flared on UDCA reintroduction (Fig. 1E,F). Follow-up after the study demonstrated resolution of this flare with reduction of ALT to normal and GGT to less than 2 times ULN within 7 weeks with continued UDCA therapy. There were two SDFs (one during phase II and one during phase III) (Fig. 1C,D, red arrows). Both participants were placed back on UDCA and moved to phase IV,
<table>
<thead>
<tr>
<th></th>
<th>UDCA Response Group</th>
<th>Null (n = 7)</th>
<th>Flare (n = 8)</th>
<th>Indeterminant (n = 7)</th>
<th>Total (N = 22)</th>
<th>PValue* (Overall)</th>
<th>PValue†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at diagnosis (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>22</td>
<td></td>
<td>0.101</td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>9.7 (2.2)</td>
<td>12.4 (4.4)</td>
<td>13.9 (3.2)</td>
<td>12.0 (3.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at start of phase I (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.157</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>12.6 (2.1)</td>
<td>13.4 (3.8)</td>
<td>15.9 (3.3)</td>
<td>13.9 (3.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>5 (71.4%)</td>
<td>6 (75.0%)</td>
<td>6 (85.7%)</td>
<td>17 (77.3%)</td>
<td>&gt;0.999</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>2 (28.6%)</td>
<td>2 (25.0%)</td>
<td>1 (14.3%)</td>
<td>5 (22.7%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>1 (14.3%)</td>
<td>1 (12.5%)</td>
<td>1 (14.3%)</td>
<td>3 (13.6%)</td>
<td>0.906</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not Hispanic or Latino</td>
<td>6 (85.7%)</td>
<td>7 (87.5%)</td>
<td>5 (71.4%)</td>
<td>18 (81.8%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not reported</td>
<td>0</td>
<td>0</td>
<td>1 (14.3%)</td>
<td>1 (4.5%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black or African American</td>
<td>0</td>
<td>0</td>
<td>1 (14.3%)</td>
<td>1 (4.5%)</td>
<td>0.299</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White or Caucasian</td>
<td>6 (85.7%)</td>
<td>7 (87.5%)</td>
<td>3 (42.9%)</td>
<td>16 (72.7%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not reported</td>
<td>1 (14.3%)</td>
<td>1 (12.5%)</td>
<td>3 (42.9%)</td>
<td>5 (22.7%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnosis group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSC</td>
<td>2 (28.6%)</td>
<td>3 (37.5%)</td>
<td>5 (71.4%)</td>
<td>10 (45.5%)</td>
<td>0.316</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSQ/AIH overlap</td>
<td>5 (71.4%)</td>
<td>5 (62.5%)</td>
<td>2 (28.6%)</td>
<td>12 (54.5%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBD diagnosis at study entry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present</td>
<td>6 (85.7%)</td>
<td>7 (87.5%)</td>
<td>6 (85.7%)</td>
<td>19 (86.4%)</td>
<td>&gt;0.999</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>1 (14.3%)</td>
<td>1 (12.5%)</td>
<td>1 (14.3%)</td>
<td>3 (13.6%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>On immunosuppressive medication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>3 (42.9%)</td>
<td>4 (50.0%)</td>
<td>2 (28.6%)</td>
<td>9 (40.9%)</td>
<td>0.862</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>4 (57.1%)</td>
<td>4 (50.0%)</td>
<td>5 (71.4%)</td>
<td>13 (59.1%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immunosuppressive medication type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corticosteroids</td>
<td>1 (14.3%)</td>
<td>0</td>
<td>1 (14.3%)</td>
<td>2 (9.1%)</td>
<td>0.644</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>2 (28.6%)</td>
<td>4 (50.0%)</td>
<td>1 (14.3%)</td>
<td>7 (31.8%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N/A</td>
<td>4 (57.1%)</td>
<td>4 (50.0%)</td>
<td>5 (71.4%)</td>
<td>13 (59.1%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Height z score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>21</td>
<td>0.018</td>
<td>0.016 (null vs. flare)</td>
<td>0.008 (flare vs. indeterminant)</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>−0.5 (1.0)</td>
<td>0.6 (0.4)</td>
<td>−0.2 (0.5)</td>
<td>−0.0 (0.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight z score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>21</td>
<td>0.301</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>−0.3 (0.9)</td>
<td>0.4 (0.5)</td>
<td>0.0 (1.0)</td>
<td>0.0 (0.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI z score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>21</td>
<td>0.991</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>0.006 (0.6)</td>
<td>0.022 (0.6)</td>
<td>−0.033 (1.1)</td>
<td>−0.001 (0.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time since diagnosis (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.063</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>3.0 (1.6)</td>
<td>1.3 (1.2)</td>
<td>1.9 (1.2)</td>
<td>2.0 (1.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UDCA dose per kg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>20</td>
<td>0.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>13.0 (2.5)</td>
<td>14.8 (3.4)</td>
<td>11.8 (4.6)</td>
<td>13.2 (3.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
with subsequent flare resolution. None developed elevated bilirubin, and there were no significant increases in ALP levels. At study baseline, height $z$ scores were different in null versus flare participants ($P = 0.016$; Table 1) and flare versus indeterminant groups ($P = 0.008$; Table 1). The diagnosis of isolated PSC versus AIH/PSC overlap, use of immunosuppressives, or presence of IBD did not impact flare or null status (Table 1). Interestingly, the null group had baseline GGT levels significantly lower than the flare and indeterminant groups and albumin levels higher than the flare group (Table 2), suggesting that these two baseline laboratory values may be predictive of a disease flare.

**STUDY ENDPOINT ANALYSIS**

Aggregate clinical laboratory results by study phase for the primary analysis population are summarized in Supporting Table S2. Significant changes occurred in GGT and ALT, as depicted in Fig. 2A, which shows results of the analysis of change in ALT and GGT from study baseline to end of phase III for all participants as well as the three response groups. Although the increase in ALT within the indeterminant group was significant, as shown, the aggregate mean change across all groups did not reach statistical significance. For GGT, the average increases across all response groups and within the null, flare, and indeterminant groups were statistically significant.

The changes in ALT and GGT from the end of phase III to the end of phase IV is shown in Fig. 2C. The mean change in ALT across all response groups as well as within each group did not reach statistical significance. However, in the case of GGT, the mean change among all response groups was significant as was the within-group change for the flare response group, as shown. The changes in ALT and GGT from the maximum value in phases II and III to the minimum value in phase IV are shown in Fig. 2D. The mean change in ALT across all groups and the decreases within the null and flare groups reached statistical significance. For GGT, the average decrease across all response groups as well as within the flare group, as shown, was statistically significant.

**INFLAMMATORY CYTOKINES**

Inflammatory cytokine levels (IL-1β, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, IL-15, TNF-α, IFN-γ, GM-CSF, MIP-1β, and eotaxin) showed no clear pattern of change over time between study phases, although the limited numbers of samples due to loss in storage may have precluded identification of some changes (Supporting Fig. S3). Baseline IL-8 levels were significantly higher in the flare group compared to the indeterminant group (Fig. 3A), and baseline levels of TNF-α were significantly higher in the flare and indeterminant groups compared to the null group (Fig. 3B).
Bile Acid Analysis

Serum UDCA levels in 13 participants at baseline, end of phase III, and end of phase IV are shown in Fig. 4. With UDCA withdrawal, 11 participants exhibited a decline in serum UDCA to very low levels (<0.4 μM), with the exception of 1 individual whose level actually increased from 4.3 μM to 13.1 μM and another who showed a decline from a high baseline level (26.7 μM) to a still elevated level (7.6 μM).
off UDCA. With UDCA reinstitution, there was a variable response, ranging from minimal increases to return to near baseline levels. The same 2 participants who had anomalous responses to withdrawal both unexpectedly exhibited a decline in UDCA from end of phase III to end of phase IV levels. Interestingly, both of these individuals belonged to the null response group.

**FIG. 1.** Serum ALT and GGT levels in study participant response groups. (A,B) Null group; (C,D) flare group; (E,F) indeterminant group. The y axis shows serum levels, and the x axis shows study week and phase. Colored lines depict individual participants. In the flare group (C,D), visits 8S and 12S are off-study visits and red arrows indicate time of early reinstitution of UDCA due to SDF in 2 participants.
**Fig. 2.** Serum ALT and GGT levels for study endpoints. (A) Results of the analysis of change in serum ALT and GGT levels from study baseline to end of phase III for total participants and response groups. (B) Change in serum ALT and GGT levels from baseline to maximum values in phases II and III. (C) Change in serum ALT and GGT levels from the end of phase III to the end of phase IV. (D) Change in serum ALT and GGT levels from the maximum value in phases II and III to the minimum value in phase IV. Data in all panels represent mean + SD. *P* values shown at the top of bars for test of within-group change between the two time points were calculated using paired *t* test.

**Fig. 3.** Baseline plasma levels from 4 null, 8 flare, and 7 indeterminant participants. (A) IL-8 and (B) TNF-α treatment response groups. Null, circle; flare, square; indeterminant, triangle. Each symbol represents an individual participant. Horizontal lines depict mean and SD for each group.
Discussion

UDCA is a hydrophilic bile acid first isolated from *Ursus maritimus* (polar bear). It is a minor bile acid in humans, and oral dosing enriches the bile acid pool up to 40%-60%, depending on the dose used. It is currently U.S. Food and Drug Administration approved for use in primary biliary cholangitis and for cholesterol gallstone dissolution in adults. However, it has been widely used off label in adults and children with other liver diseases, including PSC. Potential therapeutic mechanisms include increase in the hydrophilicity index of the bile acid pool, stimulation of choleresis, cytoprotection, and immunomodulatory and anti-inflammatory effects. The role of UDCA in the treatment of PSC is controversial, both in adults and children. In a prospective, blinded, placebo-controlled trial in adults, high-dose UDCA therapy was shown to result in worse long-term outcomes and increased SAEs compared to controls, despite short-term improvements in liver biochemistries. Similar studies have not been performed in children, although short-term biochemical improvement has been reported. Consequences of UDCA withdrawal in children with PSC are important for informing parents who may wish to discontinue the drug based on the adult study and to inform clinical trial design of UDCA studies with a control arm or any new PSC medications (e.g., vancomycin, farnesoid X receptor agonists) that would optimally be performed off UDCA.

During our study, we found that average liver biochemistries increased modestly overall with controlled UDCA withdrawal in children with quiescent PSC. Further analysis after stratification into treatment response groups showed that one third had normal ALT and GGT that remained normal in the absence of UDCA therapy (null group), one third increased above 100 IU/L (flare group), and the rest had intermediate responses (indeterminant group). None developed elevated bilirubin or significant increases in ALP levels, and all flares responded to the reinstatement of UDCA therapy. Diagnosis of PSC versus AIH/PSC overlap, use of immunosuppressive medications, or presence of IBD did not impact response status.

In this study, we observed that baseline levels of plasma IL-8 and TNF-α were higher in the flare group compared to null and/or indeterminant groups. There have been few studies of serum biomarkers reported in adult patients with PSC and none previously reported in children with PSC. A recent study on bile and serum biomarkers in adults with PSC revealed that IL-8 along with other biomarkers was significantly increased in PSC compared to controls. Moreover, serum IL-8 levels independently predicted transplant-free survival. Another adult study analyzed cytokine levels before and after UDCA treatment in patients with PSC. Similar to our findings, elevation in TNF-α and IL-8 were reported but none were significantly affected by UDCA therapy. Due to the small number of patient samples available for cytokine analyses, broad conclusions regarding cytokines as potential biomarkers of disease severity cannot be made. However, the findings do suggest that a large multicenter study of immunophenotyping in pediatric PSC is warranted to provide insight into biomarkers of disease severity, immunopathogenesis, and potential therapeutic targets.

There are several potential reasons for the variability we observed in response to UDCA withdrawal. Eight weeks off of UDCA may not be sufficient for some patients to generate a biliary inflammatory response sufficient to elevate serum biochemistries. Higher mean serum GGT and lower serum albumin at baseline were present in the flare group compared to the null group and may be predictive of a flare. Also, despite remaining in the “normal” range throughout the study, GGT levels rose significantly in the null group during or after UDCA withdrawal. These observations suggest that even when these tests are within the normal limits, patients may have low-grade indolent inflammation and be primed for a flare.
following UDCA withdrawal. This is also supported by the baseline cytokine findings for IL-8 and TNF-α in our study. However, in order to determine if these cytokines will be useful as biomarkers for predicting a flare during or after UDCA withdrawal, future investigations are necessary.\(^{(29)}\)

While we measured serum UDCA levels, our methodology did not assess for UDCA conjugates, leading to potential underestimation of the total serum UDCA levels. However, our data do support that there was withdrawal of medication during phase III in the majority of participants. The variable increase in UDCA levels with reinstitution of the drug may be explained by variability in the time required to re-establish prewithdrawal levels as well as our measurement excluding the conjugated species. Alternatively, some parents may have decided not to restart the medication, although all flare participants responded to UDCA reinstitution. Interestingly, only 1 participant in the indeterminant group flared during phase IV. The high withdrawal levels in the 2 participants noted with anomalous levels probably, at least partially, explains the lack of a flare and inclusion in the null response group.

Limitations of our study included a relatively small sample size as recruitment was difficult for several reasons. In children, PSC is less common than in adults, with an incidence of 0.23 cases per 100,000 person-years.\(^{(30)}\) Because of safety considerations in a vulnerable population, inclusion and exclusion criteria were relatively strict and may have introduced bias for selection of participants more responsive to UDCA or with milder disease. Recruitment was also challenging because the study involved withdrawing a drug that had resulted in an initial improvement in liver tests and parents are more likely to participate in a trial testing a new drug. Study retention was also problematic with 5 of 27 participants lost from the study due to early withdrawal for various reasons (Supporting Fig. S2). Finally, the period of time our participants were on a reduced dose or off UDCA was relatively brief (12 weeks).

The results of this study may have important implications for pediatric PSC therapy. The fact that biochemical flares occurred when the drug was reduced or withdrawn strongly suggests that UDCA was having at least a biochemical therapeutic effect in a significant number of study participants. Long-term “hard” endpoint trials for drug efficacy are extremely difficult to perform in children. In adults, ALP has been used as a biomarker for disease severity for most PSC trials, with data showing that patients who respond to UDCA therapy, based on significant improvement or normalization of ALP, have a better prognosis and disease outcome.\(^{(31)}\) Thus, there may be a subgroup of adult patients with PSC who are UDCA responders and perhaps should be treated.\(^{(32)}\) In children, a significant contribution from bone to serum ALP precludes its use as a biliary biomarker, and GGT is therefore used, as in our study, despite its own inherent limitations.\(^{(33)}\) A recent retrospective analysis of a large international cohort of pediatric patients with PSC found that 46% of patients treated with UDCA had a complete normalization of GGT during the first year after diagnosis, and this subset had a more favorable 5-year outcome with 99% survival with their native liver.\(^{(34)}\) If confirmed in a prospective controlled trial, these findings may suggest that UDCA responders should continue the medication whereas nonresponders could discontinue UDCA therapy for lack of efficacy.

There are several differences between features of PSC in children compared to adults, and there is speculation that pediatric PSC may represent an earlier more immunologically active stage of the disease or even a different disease.\(^{(11,35)}\) Overlap of AIH with PSC is present in a much higher percentage of children (up to 50%) compared to adults (6%-8%). However, there is no universal agreement on how to make the diagnosis of overlap in children, especially based on histologic features, or consensus on optimal treatment.\(^{(8)}\) It is theoretically possible that the immunomodulatory and anti-inflammatory properties of UDCA may have a positive impact on the AIH component of overlap disease as well. In the series reported by Miloh et al.,\(^{(7)}\) children with PSC and positive autoimmune markers without histologic features of AIH responded to UDCA therapy similar to patients with PSC but without autoimmune markers. In the present study, although presence of overlap did not predict response group, 5/8 participants in the flare group had overlap. It is tempting to speculate that UDCA may have a therapeutic impact on AIH/PSC overlap in children and should be studied.

In summary, children with PSC and AIH/PSC overlap on UDCA treatment with biochemically quiescent disease demonstrate a range of biochemical responses to a stepped withdrawal of the drug. Responses range from no change to severe flares with
ALT and/or GGT increases to above 10 times ULN, and all appear to be responsive to restarting UDCA therapy. Although normal reference intervals for GGT levels have been defined, our study suggests that “normal” GGT levels may not in fact be biologically normal in patients with PSC in light of the significant increase noted in the null group with UDCA withdrawal and reinstitution that did not rise above the normal cutoff (Fig. 2B) as well as significantly higher mean GGT baseline levels noted in the flare group. Factors, such as baseline levels of GGT and albumin as well as IL-8 and TNF-α, may be helpful in predicting which patients are most at risk for a significant disease flare off UDCA. More studies are needed to define the long-term benefit-to-risk ratio for the use of UDCA in children with PSC as well as AIH/PSC overlap, which is significantly more prevalent in children than in adults with PSC. As in adults, more reliable biomarkers of disease activity and progression along with well-designed prospective clinical trials are needed for this progressive disease with no proven effective therapy beyond liver transplantation.

Acknowledgment: The editorial assistance of Dr. Amanda M. Preston, Scientific Editor, Children's Foundation Research Institute, is greatly appreciated.

REFERENCES


32) Tabibian JH, Lindor KD. Ursodeoxycholic acid in primary sclerosing cholangitis: if withdrawal is bad, then administration is good (right?). Hepatology 2014;60:785-788.


Supporting Information

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep4.1421/suppinfo.