Comparing Two Ways to Manage Asthma in African American Children—The ASIST Study

Kaharu Sumino, MD, MPH; Leonard Bacharier, MD; Juanita Taylor, RRT; Kelley Chadwick-Mansker, RRT; Vanessa Curtis, RRT; Alison Nash, MD; Shawni Jackson-Triggs, PhD; Joseph Moen, MS; Ken Schechtman, PhD; Jane Garbutt, MD; Mario Castro, MD, MPH

AFFILIATIONS:

1Department of Medicine, Division of Pulmonary and Critical Care Medicine, Washington University School of Medicine, St. Louis, Missouri
2Department of Pediatrics, Division of Allergy, Immunology, and Pulmonary Medicine, Washington University School of Medicine, St. Louis, Missouri
3Department of Medicine, Division of Biostatistics, Washington University School of Medicine, St. Louis, Missouri

Institution Receiving the PCORI Award: Washington University School of Medicine
Original Project Title: Preference and Effectiveness of Symptom-Based Adjustment of Inhaled Corticosteroid Therapy in African American Children
PCORI ID: AS-1307-05588
HSRProj ID: HSRP20143422
Clinicaltrials.gov ID: NCT02298205

# TABLE OF CONTENTS

**ABSTRACT** .............................................................................................................................................................................. 4

**BACKGROUND** ........................................................................................................................................................................ 6

Intermittent, Symptom-based Use of Inhaled Corticosteroid ............................................................................................................ 6

Figure 1. Time to first treatment failure in BASALT (Best Adjustment Strategy for Asthma in Long Term) studya ...................................................................................................................... 7

Figure 2. Annualized rate of severe exacerbation and time to first severe exacerbationa ...................................................................................................................... 9

Rationale for Choosing SBA of ICS as the Intervention for Our Study .......................................................................................... 9

Potential Benefit of SBA in African American Children ................................................................................................................. 10

Project Structure and Aims ............................................................................................................................................................... 11

**PATIENT AND STAKEHOLDER ENGAGEMENT** ............................................................................................................................. 12

Table 1. Summary of Stakeholder Engagement ................................................................................................................................. 12

Study Advisory Board ........................................................................................................................................................................... 12

Engagement of Asthma Coaches ........................................................................................................................................................... 14

Engagement of Participating Providers ............................................................................................................................................. 15

Engagement of Patients and Caregivers ............................................................................................................................................. 15

Washington University Community Engagement Research Community Advisory Board ............................................................................. 15

**AIM 1** .......................................................................................................................................................................................... 16

Methods ........................................................................................................................................................................................... 16

Results .............................................................................................................................................................................................. 17

Table 2. Characteristics of Caregivers and Their Children in the Focus Group .................................................................................. 18

Table 3. Provider Characteristics ......................................................................................................................................................... 20

Caregiver and Provider Perception of SBA ........................................................................................................................................ 21

**AIM 2** .......................................................................................................................................................................................... 22

Methods ........................................................................................................................................................................................... 22

Table 4. Self-reported Adherence and Sense of Self-managementa .................................................................................................. 30

Results .............................................................................................................................................................................................. 36

Figure 3. Flow of study participants .................................................................................................................................................... 37

Table 5. Demographics and Characteristics of Study Participants at Entry ...................................................................................... 38

Figure 4. ACT and cACT scores at each visit ..................................................................................................................................... 40

Table 6. ACT/cACT Change From Baseline to 12 Months ............................................................................................................ 42

Table 7. Proportion of Exacerbation in Each Arm .......................................................................................................................... 42
ABSTRACT

**Background:** African American children experience a disproportionate burden in asthma. Symptom-based, intermittent inhaled corticosteroid (ICS) adjustment, or SBA (symptom-based adjustment: use of ICS together with rescue short-acting bronchodilator or inhaler), is a patient-centered approach in which patients adjust their ICS dose on a daily basis guided by symptoms and albuterol need. In controlled trials in highly selected, adherent adults and children with mild asthma, SBA was as effective as provider-driven, guideline-based management, or PBA (provider-based adjustment: daily ICS with subsequent ICS dose adjustment by provider).

**Objective:** We evaluated whether the use of SBA would result in similar asthma control compared with PBA in African American children in a real-world, primary care setting.

**Methods:** First we conducted a qualitative study that included focus group meetings with children with asthma and their caregivers as well as interviews with primary care providers to evaluate perceived barriers to and perspectives on SBA and PBA. In the main study, we conducted a randomized, open-label, parallel group, pragmatic trial in St. Louis, Missouri. African American children (aged 6-17 years) with mild asthma cared for by 12 PCPs were randomly assigned to receive either SBA (ie, 2 puffs of beclomethasone 40 mcg every time albuterol was used for rescue) or PBA (ie, fixed daily dose of beclomethasone 40 mcg, 2 puffs/day for 6- to 11-year-olds, 4 puffs/day for 12- to 17-year-olds with subsequent guideline-guided adjustment by their PCP) for 12 months. The PCP implemented treatment assignment in both arms, and blinded study staff measured the outcomes every 3 months. The primary outcome was change in asthma control at 12 months (measured by Asthma Control Test [ACT] and childhood ACT [cACT]; ACT score range = 5 to 25 and cACT = 0 to 27; lower scores indicate poor control) from baseline. Secondary outcomes included asthma exacerbation, dose of ICS taken, and quality-of-life measures.

**Measurements and Results:** We randomly assigned 206 children (SBA n = 103; PBA n = 103), and 87% completed 12-month follow-up. The average age was 10.2 years and 60% were receiving Medicaid. Household income was less than $35 000 in 57% of the participants, and 42% had at least 1 asthma exacerbation in the past year. Baseline ACT score (n = 71) was 21.5 ± 2.9 and baseline cACT score (n = 135) was 21.7 ± 2.8. The mean changes in ACT and cACT scores from baseline to 12 months (ACT mean [95% CI] were as follows: PBA: 1.55 [0.68-2.42], SBA: 0.67 [–0.33 to 1.66]; cACT mean [95% CI] change: PBA: 1.43 [0.45-2.41], SBA: 0.69 [–0.26 to 1.65]). In a modeled analysis of transformed combined ACT/cACT (controlling for test, season, gender, and clinic), the difference was not significant between arms (p = 0.10). Secondary outcomes—including change of asthma control at 6 months, proportion of those with more than one exacerbation, changes of quality-of-life measures, and lung function at 12 months—were similar between groups. Average monthly beclomethasone use was as follows: PBA 1961.3 mcg (95% CI, 1681.5-2241.1 mcg) and SBA 526.2 mcg (95% CI, 412.8-639.5 mcg) per month, p < 0.0001. More children and/or their caregivers in the SBA group felt that they, rather than their PCP, were managing their asthma (SBA: 63% vs PBA: 40% ; p = 0.001).
**Conclusions:** Control of mild asthma in African American children using symptom-based, intermittent ICS with a rescue bronchodilator (SBA) was similar to guideline-based asthma management implemented by the PCP (PBA) and supported by parent education. SBA management resulted in lower ICS exposure and a higher sense of self-management.

**Limitations:** Limitations included lack of a placebo group, lack of valid information regarding adherence to the study interventions, and use of an open-label design.
BACKGROUND

Intermittent, Symptom-based Use of Inhaled Corticosteroid

The National Asthma Education and Prevention Program (NAEPP) guideline recommends daily use of asthma controller medication and rescue use of short-acting beta-agonist (SABA) to treat symptoms for persistent asthma. In this guideline, daily use of inhaled corticosteroid (ICS) with subsequent ICS dose adjustment by the provider, or PBA (provider-based adjustment), is recommended for management of persistent asthma. However, since the publication of the NAEPP guideline in 2007, many studies have demonstrated that alternative, nondaily intermittent ICS therapy may be effective, especially in mild asthma. Intermittent, symptom-based ICS treatment, or SBA (symptom-based adjustment), is a nondaily approach in which patients adjust their ICS dose on a day-to-day basis guided by symptoms and rescue inhaler (ie, rescue short-acting bronchodilator) needs. In SBA, daily ICS is not used; rather, ICS is delivered whenever a rescue inhaler (SABA or albuterol in the United States) is used to treat symptoms. Previous studies have demonstrated that SBA of ICS is equally as effective as daily use of ICS for controlling asthma and preventing exacerbations. A study in patients with mild asthma demonstrated that symptom-based use of beclomethasone as rescue treatment together with SABA was as effective as regular use of twice-daily beclomethasone and resulted in less growth impairment. The Best Adjustment Strategy for Asthma in Long Term (BASALT) study took this evidence for symptom-based intermittent therapy one step further and investigated its use as a treatment adjustment strategy to test the following options:

1. Treatment is stepped up or down by the physician according to a clinical assessment following the NAEPP guideline (PBA in our study: physician assessment–based adjustment [PABA] in BASALT).

2. Adjustment decisions are guided by levels of biomarkers (such as exhaled nitric oxide or biomarker-based adjustment [BBA]) and by symptom-based intermittent ICS adjustment.
The BASALT study, a 9-month randomized controlled trial (RCT), demonstrated that time to moderate loss of asthma control (treatment failure) did not differ significantly among treatment strategies (9-month treatment failure rates: PBA 22% [97.5% CI, 14%-33%]; BBA 20% [97.5% CI, 13%-30%]; and SBA 15% [97.5% CI, 9%-25%]; see Figure 1) or in secondary outcomes, including exacerbation rate and lung function. Mean monthly beclomethasone use was roughly twice as high in PBA and BBA as in SBA. Moreover, in African American (AA) patients in BASALT, SBA was also effective compared with PBA, even though the treatment failure rate was higher in AA participants than in white participants.

Figure 1. Time to first treatment failure in BASALT (Best Adjustment Strategy for Asthma in Long Term) study

---

In addition to SBA with rescue inhaler, another strategy of as-needed ICS use is to implement the combination of budesonide/formoterol (Symbicort), given the fast-acting property of formoterol. Recently, as-needed use of budesonide/formoterol was investigated in many patients with mild asthma \( n = 3849 \) in SYGMA 1 clinical trial, \(^5\) \( n = 4215 \) in SYGMA 2 clinical trial.\(^6\) In both studies, as-needed use of budesonide/formoterol provided a similar rate of asthma exacerbation as daily use of maintenance budesonide (Figure 2) with less ICS; this suggests that intermittent, as-needed use of budesonide/formoterol can be considered an alternative to daily maintenance ICS in mild asthma when accounting for the patient’s preference and attitude toward daily maintenance ICS. Finally, budesonide/formoterol maintenance and inhaler therapy is an established method shown to reduce exacerbations and achieve asthma control in moderate to severe asthma.\(^7,8\) This approach has been included as one of the options for step 3 therapy in the Global Strategy of Asthma Management and Treatment guideline\(^9\) and is an approved therapy for asthma in Europe. These new methods of nondaily, symptom-triggered use of ICS are emerging as alternative patient-centered, ICS-sparing ways of taking controller medication.

However, most of these previous studies were conducted in a clinical trial setting that is quite different from usual clinical practice. The study populations were highly selective, including mostly adherent and motivated patients. Moreover, participants in these studies were required to use blinded daily and as-needed ICS inhalers to keep them blinded from study assignments. Thus, the effectiveness of SBA in a real-world, primary care setting has not been established.
Rationale for Choosing SBA of ICS as the Intervention for Our Study

African American children carry a disproportionate burden of mortality and morbidity in asthma. Even after adjusting for the higher prevalence of asthma in AA children compared with white children, the emergency department (ED) visit rate is 2.6 times higher in AA children; hospitalization rate is 2.0 times higher, and death rate is 4.9 times higher.\textsuperscript{10} We surveyed 1119 parents receiving care from more than 60 primary care providers (PCPs) in St. Louis to assess their health concerns for their children. Respondents were from every zip code in the St. Louis metropolitan area representing the local population (23% African American, 68% white, 1% Asian, 2% Hispanic, and 6% others). Asthma was the third most common parental concern, ranked as a large or medium problem by 64% of respondents.\textsuperscript{11} However, guideline-based asthma care has consistently been shown to be underdelivered, despite attempts to improve guideline adherence in both providers and patients,\textsuperscript{12-14} especially in the AA population. Thus, intermittent, as-needed use of ICS may be a more acceptable alternative approach to treating asthma in this population. Among many factors that the health community has identified as
being associated with the racial disparity in asthma treatment, lack of adherence to clinical visits and asthma therapy is one of the most frequently noted. Adherence to asthma medication is low in all populations\textsuperscript{15,16} and the adherence rate is even lower in the AA population.\textsuperscript{17} Only 40\% of minority children with asthma attend recommended follow-up appointments with PCPs, \textsuperscript{18} and many use the ED for usual asthma care.\textsuperscript{15} These patients often self-discontinue their asthma controller medications,\textsuperscript{19} resulting in poorly controlled asthma.\textsuperscript{20} SBA, on the other hand, is a patient-centered strategy in which the patient can initiate his or her own therapy (with a caregiver’s assistance if necessary) at the onset of symptoms and does not require a visit to the PCP. In addition, SBA may overcome the negative beliefs and concerns of patients and caregivers about asthma medication. Studies have shown that asthmatic patients and caregivers have health beliefs that may conflict with the treatment recommended by the provider,\textsuperscript{21,22} and this is especially so in AA populations.\textsuperscript{23,24} Studies have consistently shown that patients using SBA were exposed to lower doses of ICS compared with traditional daily therapy\textsuperscript{2-4} and had less growth impairment.\textsuperscript{3} Therefore, an SBA treatment strategy may be better accepted in a minority population, resulting in better adherence and lowered morbidity.

**Potential Benefit of SBA in African American Children**

If SBA is similar to PBA in controlling asthma and preventing exacerbations, there is a strong potential to benefit many more high-risk patients with asthma than in previous educational interventions, which commonly reach only limited populations. Our study included an intervention that may result in a fundamental shift in the way PCPs adjust the dose of ICS in a minority asthma population. Studies have already shown in multiple RCTs that SBA is similar to PBA in preventing exacerbation and controlling asthma in patients with mild to moderate asthma; therefore, in our project we first conducted a qualitative study to identify the barriers to SBA and PBA in AA children with asthma, their caregivers, and PCPs. Then, we conducted a randomized, pragmatic trial in the AA pediatric population, in which the idea of SBA has been most accepted. We recruited participants from pediatricians’ offices in the St. Louis AA community.
Project Structure and Aims

Our project consisted of 2 studies:

1. **Qualitative study:** Focus group study of parents and older children with mild to moderate asthma and provider interview
   
   - *Aim 1:* To determine expectations and barriers to SBA and PBA in AA children with mild to moderate asthma, their caregivers, and PCPs.

2. **Main study:** The Asthma Symptom-based adjustment of Inhaled Steroid Therapy in African American children (ASIST) study
   
   - *Aim 2:* To evaluate if SBA is equivalent to PBA in improving asthma control in AA children with mild asthma.

   **Primary hypothesis:** SBA of ICS is equivalent to PBA in improving asthma control in AA pediatric patients with mild asthma.

   **Secondary hypotheses:** The monthly cumulative dose of ICS is lower in participants with SBA compared with PBA in AA pediatric patients with mild asthma.

   The rate of asthma exacerbation over 12 months is equivalent in participants with SBA of ICS compared with PBA in AA pediatric patients with mild asthma.

   Satisfaction and adherence at 12 months is higher in participants with SBA of ICS compared with PBA in AA pediatric patients with mild asthma.

To obtain perspectives from caregivers and older children for whom SBA may be appropriate, we included participants with mild to moderate asthma in the qualitative study. The main study (ASIST) included children with mild asthma, as previous studies of SBA in children were conducted in those with mild asthma. Mild asthma also reflects most of the asthma population in the primary care setting.25
PATIENT AND STAKEHOLDER ENGAGEMENT

During the study, we engaged several groups of stakeholders, including (1) a study advisory board, (2) asthma coaches, (3) PCPs, and (4) the Washington University Community Engagement Research Community Advisory Board. The summary of the engagement process is listed in Table 1 and the specific description of each stakeholder is described below.

Table 1. Summary of Stakeholder Engagement

<table>
<thead>
<tr>
<th>Community From Which We Identified Stakeholders</th>
<th>Stakeholders Engaged</th>
<th>Methods Used to Engage Stakeholders</th>
<th>Process in Which Engagement Occurred</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Previous research participants</td>
<td>• Parents and caregivers of asthmatic child</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Word of mouth</td>
<td>• Adolescents with asthma</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Coordinator’s personal connection</td>
<td>• Primary care providers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Pulmonary clinics, nurses</td>
<td>• Asthma coaches</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Existing community advisory board at the university</td>
<td>• School nurse</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• American Lung Association</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Community board at the university</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Asthma researchers</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Focus groups with parents and children</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Provider interviews</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Study advisory board meetings</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Meetings with research group</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Presentation at research and community meetings</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Identifying participating providers</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Defining outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Defining study structure and protocol</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Soliciting study questionnaire feedback</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Soliciting recruitment input</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Troubleshooting various issues (eg, no-shows, nonadherence)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Study Advisory Board

Our study advisory board consisted of a PCP, a parent who is also a church pastor, a school nurse, a nurse practitioner affiliated with an asthma van that serves the target community and schools, local American Lung Association personnel, a parent who is also an asthma coach, a previous public school official who is now involved in community engagement research and acts as a local asthma association board member, and a youth patient with asthma. We identified the advisory board members by using personal referrals from asthma patients, clinicians, researchers, and community members. We asked for referrals for motivated individuals interested in community research in childhood asthma. We explained the role and
required time commitment to all potential advisory board members who were referred to us, and we invited any interested stakeholders to join. We engaged the study advisory board by conducting in-person meetings every 2 months and email communications between meetings. We sought input from the advisory board throughout the planning and conduct of ASIST, and the advisory board played an essential role in implementing ASIST. We engaged the board for all stages of the study, including planning (study design and protocol development) and implementation (structure of study visits and intervention, recruitment and retention, and dissemination). We made many changes from the original plan for ASIST based on advisory board input, and the board played a significant role in addressing issues related to recruitment and retention during the study. Below are some specific examples.

Changes Made Regarding Recruitment

1. **Ways to decrease no-shows for the first visit:** Board members suggested (1) confirming that the parent knows about transportation support and flexibility in scheduling visits, (2) scheduling the first visit at the time of the initial screening call, and (3) creating a postcard with fun study incentives. Per this suggestion, we (1) created a postcard with a picture of children and study incentives, (2) refined the direct scheduling process for the screeners, and (3) notified the parent regarding availability of evening visits and transportation support.

2. **Change in inclusion criteria:** Due to challenges in recruitment and the need for a more diverse study population, we identified the need to change the inclusion criteria and discussed 2 options for this change. First, we asked the board members’ opinion regarding the enrollment of participants with uncontrolled asthma symptoms (with indication for maintenance asthma therapy) who were not currently using ICS (required in the current inclusion criteria) and including them with participants currently using ICS. Some members of the advisory board were opposed to this change and mentioned that PCPs should be the ones deciding whether a child should take maintenance medication rather than the researchers doing so per the study protocol. Based on this input, we chose not to implement this change. Second, we asked the board’s opinion regarding expanding our criteria to include participants with more symptoms. The board members felt that it was important to diversify our study population and that this was an acceptable change. We implemented these changes, as detailed under [Changes to the Original Study Protocol](#).
Changes Made Regarding Retention

1. **Input on items with study logo:** Members recommended items that they thought would enhance retention (water bottle, umbrella, lunch bag, and backpack). Per these suggestions, we developed and distributed these items at each follow-up visit.

2. **Additional reminder phone call by coaches at 6 months:** The board members suggested that the asthma coaches give feedback to the participants to enhance retention, as feedback could be part of the asthma education already built into the study. Based on this input, we added an additional phone contact from the coaches after the 6-month visit. This additional contact was well received by the parents and likely enhanced adherence to the study arm.

Engagement of Asthma Coaches

We engaged 2 asthma coaches, who are also parents of a child with asthma, as part of the research team. We identified them through a recommendation from other asthma researchers who had worked with them previously. The coaches delivered education by phone about symptom recognition and albuterol use; this coaching was essential for conducting our study and intervention arms, and provided many other benefits toward implementing ASIST. We believe the coaches’ involvement also contributed significantly to the retention of study participants in ASIST. The coaches connected very effectively with the parents and children during their coaching phone call to deliver asthma education, and we received positive feedback from the participants and parents. In fact, one of the coaches reestablished contact with a participant who was in danger of being lost to follow-up. We featured the coaches in our newsletter in response to a parent’s request to know more about the asthma coaches.

The coaches provided input regarding various issues, including (1) development of the coaching manual, reflecting their suggestions in the content and wording; and (2) development and revision of our asthma action plan for to improve the likelihood that parents/participants would accept the plan.
Engagement of Participating Providers

Twelve primary care pediatricians and 3 nurse practitioners in St. Louis who manage AA children with asthma acted as participating providers for ASIST. We identified them through an existing practice-based research network consortium of community pediatric providers (Washington University Pediatric and Adolescent Ambulatory Research Consortium) as well as through personal referrals from advisory board members and other stakeholders. Many of them participated in the provider interview and provided input on the initial planning of ASIST, including outcome, visit structure and logistics, study intervention, and the process of study-related education. Per study protocol, the providers administered the study intervention to those patients who participated in ASIST. They addressed issues the children/parents faced during participation and provided assistance to improve adherence to study arms. They also played an essential role in retaining participants/parents in ASIST, working with the research team to maintain contact with the study participants. In the latter part of the study, we engaged a provider to help us establish the process for transitioning the care back to the PCP when the child completes the study. Per this provider’s suggestion, we developed a formal process and checklist for the final visit.

Engagement of Patients and Caregivers

We conducted patient and caregiver engagement mainly through a focus group study, which is described in the Methods and Results sections. We also included youth patients with asthma and parents of asthmatic children on our advisory board.

Washington University Community Engagement Research Community Advisory Board

We held an initial meeting in early 2014 to introduce the study and presented an update and results of the focus group study on June 4, 2015, to the Community Engagement Research Advisory Board. Community members, including staff of federally funded clinics, school officials, nurses, educators, and members of organizations, attended the regularly scheduled meeting. We received recommendations for possible participating providers and a dissemination plan.
AIM 1

Aim 1: To determine expectations and barriers to symptom-based adjustment (SBA) and provider-based adjustment (PBA) in African American (AA) children with mild to moderate asthma, their caregivers, and primary care providers (PCPs).

Methods

Focus Group Interviews

We recruited our target population of caregivers of AA children and adolescents with asthma using flyers posted in community spaces, including health fairs, bulletin boards, and pediatricians’ offices. We included caregivers of AA children aged 6-17 with asthma who met the criteria of (1) self-reported, physician-diagnosed asthma; (2) prescribed treatment with low-dose ICS, leukotriene receptor antagonist (LTRA), or low-dose ICS plus LTRA; (3) symptoms fewer than 5 times a week per parent report; and (4) no severe asthma (ie, not taking high-dose ICS or combination therapy and no history of intensive care unit [ICU] admission). We also invited adolescents (aged 12-17) who met the above criteria to join the focus group interviews. We used the same inclusion criteria as the main study so that we could conduct the focus group in the target population for the main study. We conducted focus group interviews with AA caregivers of children and adolescents with asthma to identify the barriers to and challenges of treating asthma, address goals for asthma care, and explore their perceptions of SBA and the main study. We conducted 6 focus group meetings with parents and 2 meetings with adolescents with asthma from May 14, 2014, to July 18, 2014, using a discussion guide developed with our study advisory board and an experienced qualitative researcher at Washington University. We held an adolescent focus group independently without parental presence. We continued the focus group meetings until no new themes emerged.

Provider Interviews

We invited PCPs for AA children with asthma in the city of St. Louis to participate. We conducted face-to-face interviews with a purposive sample of 14 pediatric PCPs until no new themes emerged. The interviews followed a discussion guide to gather providers’ opinions
about the barriers in their routine asthma care, perspectives regarding SBA, and input into logistics and design of the main study.

**Analysis**

We transcribed and analyzed interview data using inductive thematic-based coding to assess caregiver and provider perception. Two external facilitators who conducted the focus group interviews also conducted the qualitative analysis using Dedoose (https://www.dedoose.com/), a cross-platform software. The qualitative analysts employed a mixed content analysis approach that included (1) development of a codebook and (2) identification of themes. The detailed method for analyzing the perception of SBA by caregivers and providers is described in Dy et al.²⁶ This manuscript also describes caregiver/provider input into planning the ASIST study.

**Results**

**Focus Group Participants**

Twenty-six parents and 10 adolescents participated. Table 2 displays the sociodemographic characteristics of caregivers and children. All participants were African American and cared for children with mild to moderate asthma; 42% were fully employed, and 54% had an annual household income of less than $20,000.
Table 2. Characteristics of Caregivers and Their Children in the Focus Group

<table>
<thead>
<tr>
<th>Caregiver demographics (n = 26)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Relationship to child (%): parent</td>
<td>100</td>
</tr>
<tr>
<td>Mean age in years at visit (SD)</td>
<td>37.32 (8.6)</td>
</tr>
<tr>
<td>AA race (%)</td>
<td>100</td>
</tr>
<tr>
<td>Married (%)</td>
<td>35</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Caregiver employment status (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Full-time employment</td>
<td>42</td>
</tr>
<tr>
<td>Part-time employment</td>
<td>19</td>
</tr>
<tr>
<td>Unemployed</td>
<td>31</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Caregiver annual income (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than $20 000</td>
<td>54</td>
</tr>
<tr>
<td>$20 000-$35 000</td>
<td>19</td>
</tr>
<tr>
<td>$35 000-$50 000</td>
<td>15</td>
</tr>
<tr>
<td>$50 000-$70 000</td>
<td>4</td>
</tr>
<tr>
<td>$70 000-$100 000</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level of education (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Some high school</td>
<td>35</td>
</tr>
<tr>
<td>Completed high school</td>
<td>31</td>
</tr>
<tr>
<td>Below high school</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Characteristics of children (n = 26)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent-reported severity of asthma (%)</td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>35</td>
</tr>
<tr>
<td>Moderate</td>
<td>46</td>
</tr>
<tr>
<td>Severe</td>
<td>12</td>
</tr>
</tbody>
</table>

| Mean number of ED or urgent care visits in the past year (SD) | 1.46 (1.84) |

<table>
<thead>
<tr>
<th>Frequency of controller medication use (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Every day</td>
<td>46</td>
</tr>
<tr>
<td>Only when asthma is bad</td>
<td>42</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency of rescue medications in the past 4 weeks (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>More than once daily</td>
<td>15</td>
</tr>
<tr>
<td>Few times a week</td>
<td>46</td>
</tr>
<tr>
<td>Few times a month</td>
<td>15</td>
</tr>
<tr>
<td>None in the past 4 weeks</td>
<td>24</td>
</tr>
</tbody>
</table>

Abbreviations: AA, African American; ED, emergency department; SD, standard deviation.
Input From Parents and Older Children About the Main Study. The participants provided input regarding important goals of asthma care, their routine asthma care and barriers, our SBA intervention, and study recruitment strategies. Many parents said that “having a normal life,” “no limitation in activity,” and “less symptoms” were important goals of their child’s asthma care. This justified our initial plan for using asthma control (measured by Asthma Control Test score) as the primary outcome for ASIST. Many parents and some older children also listed “taking less medicine” as an important goal, even ranking it above some of the other common goals, such as “less asthma attacks.” Therefore, we moved the cumulative monthly dose of ICS outcome to the top of the secondary outcomes and also added plans to measure the ICS dose by the dose counter. The participants also provided valuable input regarding their perception of the alternative treatment approach, SBA. At the initial screening for study participation, we included a detailed explanation of SBA’s efficacy and safety from previous trials. We also gained input about the recruitment strategies, including appropriate amount of reimbursement, how providing free study drugs would motivate parents/children to participate, and the importance of the PCP’s involvement regarding children’s participation in this clinical study.

Provider Interviews

We conducted face-to-face interviews with 14 PCPs (13 pediatricians and 1 nurse practitioner) serving AA children with asthma. Table 3 displays the sociodemographic characteristics of providers (n = 14). For most providers, the ethnicity of their patient population was more than 50% African American.
Table 3. Provider Characteristics

<table>
<thead>
<tr>
<th>Providers</th>
<th>n = 14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean number of years in practice</td>
<td>24 ± 7</td>
</tr>
<tr>
<td>Solo practice structure</td>
<td>35%</td>
</tr>
<tr>
<td>Number of patients seen daily</td>
<td></td>
</tr>
<tr>
<td>15-30</td>
<td>93%</td>
</tr>
<tr>
<td>31-50</td>
<td>7%</td>
</tr>
<tr>
<td>More than 50% AA patients in practice</td>
<td>71%</td>
</tr>
<tr>
<td>Number of patients with asthma a</td>
<td></td>
</tr>
<tr>
<td>1-100</td>
<td>7%</td>
</tr>
<tr>
<td>101-300</td>
<td>7%</td>
</tr>
<tr>
<td>301-500</td>
<td>43%</td>
</tr>
<tr>
<td>501-1000</td>
<td>21%</td>
</tr>
<tr>
<td>More than 1000</td>
<td>21%</td>
</tr>
</tbody>
</table>

Abbreviation: AA, African American.

a Self-reported.

Providers’ Input Into the Study Protocol. The providers offered their insights on (1) the current NAEPP asthma guideline; (2) barriers to and advantages of SBA; (3) study structure, including their involvement, randomization, and follow-up visits; (4) screening and retention; and (5) feasibility of the ASIST study. Although all providers felt the current asthma guideline was useful and used it routinely in their practice, they recognized that patients’ adherence to the guideline-recommended therapy was a constant challenge. Most providers said that it was difficult to adjust patients’ asthma medication in a timely fashion, as patients frequently miss their follow-up visits and seek medical attention only when—or after—having problems.

Regarding SBA, the concept was generally accepted with nearly half of providers saying their patients had been engaged in some form of intermittent therapy. The providers had concerns about potential confusion associated with patients/parents changing the strategy of management as well as recognition of symptoms that would trigger albuterol use. The providers agreed with our plan to provide education to the participants at the beginning of our
study. They also made suggestions regarding the importance of patient reimbursement, free medication and free parking, and asthma education within the trial to retain participants, all of which we included in the protocol.

Caregiver and Provider Perception of SBA

We analyzed caregiver and provider perception of SBA and published the results.26 Caregivers and PCPs viewed SBA favorably, stating its similarity to their current practice and its potential for decreased cost and use of ICS. They agreed that SBA could be a patient-centered strategy empowering caregivers/patients to be actively involved in their medication management. Some caregivers expressed concern regarding an abrupt transition to intermittent dosing from daily therapy and the risk of worsening asthma symptoms. Caregivers and PCPs both suggested that education on symptom recognition and close physician–patient communication would be essential to facilitate SBA implementation.
AIM 2

**Aim 2:** To evaluate whether symptom-based adjustment (SBA) is equivalent to provider-based adjustment (PBA) in improving asthma control in African American (AA) children with mild asthma.

Methods

**Study Overview**

The overall goal of aim 2 was to identify an acceptable, pragmatic, and effective approach to asthma management in high-risk AA children. The ASIST study was a randomized, open-label, 2-arm, parallel, pragmatic trial to evaluate whether symptom-based, intermittent, ICS adjustment (ie, SBA)—an alternative, patient-centered approach to asthma therapy—is as effective as provider-driven guideline-based management (ie, PBA) in improving pediatric asthma outcomes.

**Study Design and Setting**

ASIST was a randomized, open-label, parallel group, pragmatic trial in St. Louis, Missouri. AA children (aged 6-17; N = 206) with mild asthma from the practices of 12 PCPs were randomly assigned to receive either SBA (n = 103) or PBA (n = 103) for 12 months. Study intervention was implemented by the PCP following routine practice in AA children with mild asthma. We chose the randomized, open-label, pragmatic design to evaluate the effectiveness of SBA in a real-world setting.

**Participants**

**Target Population.** AA children with mild asthma who were under care of the participating PCPs in St. Louis, Missouri, were the target population. The total target sample size was 200 (see Analysis section for sample size calculation).

**Recruitment.** With the approval of the Institutional Review Board, the participating providers used the electronic patient care record to prescreen asthma patients they thought
might meet the eligibility criteria and provided the study team with a list of potential participants (including the parents’ contact information) for recruitment purposes. Parents of all children identified as potentially eligible received a written invitation letter and study participation opt-in/-out postcard signed by the provider and the principal investigator from the research team. Following a 1-week waiting period, the study team contacted the potential participants’ parents by phone. In this screening call, the study staff reviewed study participation requirements and the contents of the consent document; assessed eligibility, including the asthma evaluation questionnaire; and reviewed the child’s medical history. For those who met the inclusion criteria, the first visit was scheduled within 4 weeks. We did not collect reasons for declining to participate. Parents or guardians provided written consent and the participant provided verbal or written assent at the first visit. We informed participants that the study’s purpose was to compare 2 asthma treatment strategies, SBA and PBA.

**Inclusion and Exclusion Criteria**

**Inclusion Criteria**

1. Patients aged 6-17 with physician-diagnosed asthma for at least 6 months
2. Self- or parent-reported AA race or mixed race with at least 1 grandparent with AA race
3. Receiving asthma care by the participating primary care pediatrician
4. Prescribed low-dose ICS monotherapy (beclomethasone up to 160 mcg per day for 6- to 11-year-olds, 240 mcg per day for those 12-17); leukotriene receptor antagonist (LTRA); or low-dose ICS plus long-acting beta-agonist (LABA; for those older than 12 years) for at least the past 12 weeks, regardless of adherence
5. Asthma Evaluation Questionnaire (AEQ)\(^4\) score 0 to 2 out of 3, in 2 of the 3 questions with a total score of less than 5 out of a possible 9 (see following section for explanation of AEQ; the AEQ questionnaire is provided in Appendix 1)
6. Pre-bronchodilator forced expiratory volume in the first second of expiration (FEV\(_1\)) ≥ 80% of predicted value at entry
7. No history of or current smoking
8. Ability to provide baseline information at phone screening and randomization visit

9. Ability and willingness to provide informed consent

   If item 5 was not met at the time of screening due to recent asthma exacerbation (within 4 weeks), rescreening was allowed once.

**Justification for Using AEQ Score as an Inclusion Criterion.** The AEQ included 3 questions asking about the participant’s frequency of symptoms, albuterol use, and nighttime awakening due to symptoms. The AEQ uses a 0-3 Likert-type scale and a score range of 0 to 9, with higher scores indicating more symptoms. This questionnaire was successfully used in the BASALT trial in adult asthmatics to safely select those with mild asthma who would be appropriate candidates for SBA.4 Since the AEQ was used in an adult population, we lowered the threshold on the second score category (score 1) on the first two questions from 3 to 5 times a week (as the original questionnaire stated) to 3 to 4 times a week. We expected that patients with a score of 0 or 1 on all 3 questions would have mild asthma, and that SBA treatment would be appropriate for them if randomized.

**Exclusion Criteria**

1. Pre-bronchodilator FEV₁ < 80% of predicted value at entry

2. Asthma requiring daily combination therapy with medium- to high-dose ICS with LABA

3. History of intubation, noninvasive ventilation, or ICU admission for asthma exacerbation

4. Chronic oral corticosteroid therapy

5. Chronic disease that in the opinion of the principal investigator/PCP would prevent participation in the trial

6. No landline telephone or cell phone to communicate with study staff

7. Non–English speaker

8. Another participant of ASIST in the same household
Randomization

After eligibility was confirmed at the first visit and consent obtained, participants entered a 2- to 4-week run-in period during which they were placed on a fixed dose of twice daily beclomethasone (beclomethasone 40 mcg 1 puff a day for children aged 6-11 and beclomethasone 40 mcg 2 puffs twice a day for children aged 12-17). They also received telephone-based education (2-4 sessions) from an asthma coach that focused on symptom recognition and appropriate use of albuterol. We required completion of the coaching call to proceed with randomization. The asthma coach made an additional reminder call at 6 months to review the initial education and asthma action plan given at randomization. For those participants who met the randomization criteria, randomization was performed online with the REDCap data management system using computer-generated random numbers stratified by practice. To ensure there was no temporal bias in arm assignment, we employed a blocked randomization scheme in which the study team or provider was blinded to the scheme. Following the run-in period, participants were randomly assigned to SBA or PBA in a 1:1 ratio. Unblinded study staff performed the randomization. Study staff notified the participants and their PCP of their assignment after it was determined.

Intervention and Control

The PCP received a training session from the research team to review each strategy (SBA and PBA) and the asthma action plan before enrolling any participant from his or her practice. All participants received a written asthma action plan, specifically developed for each study arm by the study team, which they were instructed to keep throughout the study. The participants implemented their assigned asthma management strategy for the 12-month study.

**Intervention.** The intervention consisted of symptom-based, intermittent ICS adjustment (SBA: use of ICS together with rescue short-acting bronchodilator) in which participants were instructed to take beclomethasone 40 mcg 2 puffs (80 mcg total) each time they needed to take the rescue medication (albuterol) because of asthma symptoms, as instructed by the asthma coach (due to shortness of breath, coughing, wheezing, inability to play sports, increase in work of breathing, etc). When the participant was feeling well and not
using albuterol he or she did not need to take beclomethasone. Participants were instructed to carry the albuterol and beclomethasone together.

**Control.** The control consisted of provider-driven guideline-based management (PBA) in which the provider adjusted the dose of ICS per routine practice (following the NAEPP guideline), guided by his or her clinical assessment of asthma control, including albuterol use, frequency of symptoms, Asthma Control Test (ACT) score, nocturnal symptoms, and history of exacerbation(s) since the last PCP visit.

After randomization, participants assigned to PBA began taking beclomethasone 40 mcg 1 puff twice a day for children aged 6 to 11 and beclomethasone 40 mcg 2 puffs twice a day for children aged 12 to 17, with subsequent adjustments made by the provider. We chose the PBA arm as a control to reflect the current standard of care for these participants with mild to moderate asthma.

**Study Outcomes**

The blinded study coordinator measured study outcomes every 3 months at follow-up study visits. This coordinator was blinded to study interventions and did not discuss issues regarding study assignments with the participants. Primary and secondary outcomes were prespecified.

**Primary Outcome Measure (Aim 2).** Our primary outcome measure was the change in ACT (used in 12- to 17-year-olds) or children’s ACT (cACT; used in 6- to 11-year-olds) score from baseline to 12 months in the SBA and PBA arms.

**Rationale for Using ACT as Primary Outcome.** The ACT is a validated, 5-item (7-item for cACT) questionnaire that assesses asthma control (on a scale of 5-25) over the previous 4 weeks. The ACT is included in the current NAEPP guideline as one method to assess asthma control; it is the recommended outcome measurement for NIH asthma clinical trials. A minimally important clinical difference of 3 has been determined for the ACT. We chose this measure of asthma control over alternative outcomes, such as exacerbations or health care
utilization, as the parents and older patients in this population identified “having less symptoms” and “be[ing] able to do physical activity without any limitation” as their most important goals of asthma care. In addition, because exacerbation rates are low in this patient population with mild to moderate asthma, measured reduction in exacerbation rates may be too small to be clinically meaningful or easily detected with our sample size. Finally, both measurements demonstrate robustness and responsiveness to loss of asthma control, and higher scores indicate better control, with a score ≤ 19 indicating that asthma is inadequately controlled.\textsuperscript{28,30,31}

Secondary and Exploratory Outcome Measures

Monthly Cumulative Dose of ICS. At each follow-up visit, the blinded coordinator reviewed the doses used, as indicated by the dose counter, and calculated the cumulative ICS dose. In our focus group study, many parents and older children with asthma were concerned about the adverse effects of the medications and listed “less medication” as an important goal for their asthma care. In previous studies\textsuperscript{3,4}, the SBA group used less ICS compared with a guideline-based, daily ICS strategy (ie, PBA).

Rate of Asthma Exacerbation Over 12 Months. The blinded coordinator asked about and recorded participants’ asthma exacerbations every 3 months. We defined asthma exacerbation as (1) ED/urgent care/primary care visit for “asthma attack” for which the participant was treated with oral steroids (by self-report or by report from PCP); or (2) hospitalization for “asthma attack”. We also assessed the proportion of participants with ≥ 1 asthma exacerbation.

Pediatric Quality-of-Life Scores. We used the Children’s Health Survey for Asthma (CHSA) for caregivers\textsuperscript{32} for 6- to 11-year-olds and CHSA—Child Version\textsuperscript{33} for 12- to 17-year-olds, both of which we measured at baseline and at 12 months. In addition, we measured Patient-Reported Outcomes Measurement Information System (PROMIS) parent proxy item v1.0—asthma impact\textsuperscript{34} in all parents/guardians and PROMIS pediatric item\textsuperscript{35} v1.0—asthma impact in 8- to 17-year-olds at baseline and at each follow-up visit. The CHSA was developed by the
American Academy of Pediatrics to measure pediatric health-related quality of life in asthmatic children and families. It measures asthma-specific domains including physical health, activity, and emotional health and is listed as a supplemental outcome measurement for quality of life in asthma in NIH clinical trials. Both the proxy version and the child version have been validated. The score range is 0 to 100, with higher scores indicating better quality of life. No cut-off score for poor quality of life has been determined. PROMIS is an emerging tool developed by NIH that aims to standardize quality-of-life measurements across studies, and the asthma version measures asthma-specific quality of life during the past week. Higher scores indicate worse quality of life. For most PROMIS instruments, a score of 50 is the average for the US general population, with a standard deviation of 10 because calibration testing was performed on a large sample of the general population. The final score is represented by the T score, a standardized score with a mean of 50 and a standard deviation of 10.

**Change in Pre-bronchodilator Percent-Predicted FEV₁ From Baseline at 12 Months.** Lung function was measured by trained study staff using a portable spirometer (KoKo spirometer) at baseline and 12 months at PCPs’ offices. Change in percent-predicted FEV₁ pre-bronchodilator at 12 months from baseline was assessed.

**Missed School Days.** PCPs ascertained the number of missed school days due to asthma at each follow-up visit, and the study coordinator recorded the number every 3 months.

**Change in ACT.** The change in ACT (used in 12- to 17-year-olds) or cACT (used in 6- to 11-year-olds) score from baseline to 6 months was recorded.

**Satisfaction and Adherence Survey Score.** The study team developed the satisfaction survey, as there was no validated questionnaire that would accurately measure patients’ and/or parents’ satisfaction with care and their views about the positive and negative aspects of their asthma treatment assignment. The adherence question in this survey measured how much the caregiver/participant thought he or she was able to adhere to the management plan provided by the PCP. We used the measurement at 12 months to compare the overall
satisfaction between study arms. The questionnaire is included in Appendix 2 and the main questions are shown in Table 4.
Table 4. Self-reported Adherence and Sense of Self-management\textsuperscript{a}

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Treatment Arm</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PBA</td>
<td>SBA</td>
</tr>
<tr>
<td></td>
<td>Child (n = 35)</td>
<td>Child (n = 40)</td>
</tr>
<tr>
<td>Who do you feel took charge in adjusting your asthma medication during the study?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mostly me and/or caregiver</td>
<td>14 (44%)</td>
<td>28 (72%)</td>
</tr>
<tr>
<td>Doctor and me equally</td>
<td>10 (31%)</td>
<td>8 (21%)</td>
</tr>
<tr>
<td>Not mostly me and/or caregiver</td>
<td>8 (25%)</td>
<td>3 (8%)</td>
</tr>
<tr>
<td>How much do you think you were able to follow the instruction (do what you were supposed to do with your inhalers)?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>You followed mostly every day (75%-100% of the time)</td>
<td>24 (69%)</td>
<td>29 (73%)</td>
</tr>
<tr>
<td>You did not follow mostly every day</td>
<td>11 (31%)</td>
<td>11 (28%)</td>
</tr>
<tr>
<td>Who do you feel took charge of your child’s asthma management during the study?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mostly child and/or caregiver</td>
<td>33 (40%)</td>
<td>55 (63%)</td>
</tr>
<tr>
<td>Doctor and child equally</td>
<td>32 (39%)</td>
<td>27 (31%)</td>
</tr>
<tr>
<td>Not mostly child and/or caregiver</td>
<td>18 (22%)</td>
<td>5 (6%)</td>
</tr>
<tr>
<td>How much do you think you (or your child) were able to follow the instructions (do what you/your child were supposed to do with the inhalers)?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Your child followed mostly every day (75%-100% of the time)</td>
<td>60 (72%)</td>
<td>76 (86%)</td>
</tr>
<tr>
<td>Your child did not follow mostly every day</td>
<td>23 (28%)</td>
<td>12 (14%)</td>
</tr>
</tbody>
</table>

Abbreviations: PBA, provider-based adjustment; SBA, symptom-based adjustment.
\textsuperscript{a} Reported by child: top half; reported by parent: bottom half. P values in \textbf{bold} are significant at \( p < 0.05 \).
Exploratory Outcomes (time to first exacerbation, proportion of those with ACT decline, SBA participants’ preference, and narrative at the end of the study).

We performed an exploratory analysis of time to first exacerbation and the proportion of participants who had at least a 3-point decline in ACT or cACT scores at 12 months from baseline in both arms. We had not previously planned to perform these analyses.

To explore SBA participants’ preference, we asked the parents and participants (> 12 years old) in the SBA arm whether they preferred SBA, which they had followed during the study, or the management method that they had experienced prior to the study (as instructed by the PCP). This question was asked only of participants in the SBA arm, as they had experienced both SBA and PBA (the latter prior to the study). Also, at the end of the study, we asked patients to write narratives regarding their perspective on their treatment assignment.

Time Frame of the Study

The PCP conducted treatment adjustments in both arms. After randomization, the PCP was notified of the participant’s treatment assignment and then scheduled the initial follow-up visit in 4 to 6 weeks to review the treatment assignment. During the remainder of the study, the PCP determined the frequency and intensity of the follow-up visits per routine practice and addressed various management issues, including poor asthma control, missed visits, nonadherence to study assignment, and post-exacerbation care. Regularly scheduled provider visits as well as study follow-up visits occurred at the PCPs’ offices. Blinded coordinators measured outcomes, including asthma control, episodes of asthma exacerbations, and quality of life, every 3 months. We chose 3-month follow-ups because we felt this interval was adequate for distributing the study drugs to participants and for measuring outcomes.

Data Collection and Sources

To maximize follow-up rate, the study team aimed to engage the participants/parents by providing a study timeline sheet at each study visit, making reminder calls for upcoming visits, and sending newsletters and postcards with study updates. The study team also kept in close contact with PCPs and provided the enrollment status of each participant in the study.
every 2 to 3 months. We collected the reason for participant withdrawal when possible, but many who withdrew were lost to follow-up, and we were unable to regain contact.

Analytical and Statistical Approaches

Primary Outcome. Sample Size Calculation. We calculated the power and sample size using the primary outcome, change in ACT (or cACT) score from randomization to 12 months. Based on data from a previous study\(^{30}\) and unpublished data from our own RCT,\(^{37}\) the standard deviation of the ACT at a single time ranged from 3.5 to 5.0 and the correlation between baseline and follow-up measurements was approximately 0.5 for the cACT (6- to 11-year-olds) and ACT (12- to 17-year-olds). Therefore, we estimated the standard deviation for the change in ACT score to be 5.0. We also defined the upper and lower limits of the between-study arms difference in ACT score for equivalence to be \(\pm 2.5\) (for equivalence, the upper and lower bound of the 95% CI of the difference must fall within this range), considering that the minimally important clinical difference had been determined to be 3 for the ACT.\(^{30}\) We concluded that 200 participants (100 in each arm) would provide 90% power to establish equivalence with a 5% type I error rate with a dropout rate of 15%.

Primary Outcome Analysis. Recruitment ended with 103 participants in each study arm. We performed the primary outcome analysis as an intention-to-treat (ITT) analysis of the changes in ACT/cACT scores from randomization to 12 months. ACT/cACT measurements were taken at baseline and at 3, 6, 9, and 12 months. To allow for simultaneous analysis of the changes in ACT and cACT scores to provide study findings from the entire age range included in our study, we combined ACT and cACT scores in the analysis and created an indicator variable for which test was taken. We used a mixed-effects, repeated measures analysis of variance (ANOVA) model to evaluate the longitudinal data with study site as a random effect. The analysis of primary interest involved the group-by-time interaction and the statistical contrast that compared the changes in ACT/cACT scores from baseline to 12 months in the control arm with the corresponding change in the intervention arm. Covariates adjusted for in this model included gender, season of enrollment, and an indicator variable for which asthma control test was taken (ACT or cACT). We considered several covariance structures for the longitudinal
model and selected compound symmetry as the best fitting according to information criteria such as Akaike and Schwarz-Bayesian. The distribution of combined ACT and cACT scores was strongly left skewed (ie, to lower scores); therefore, we transformed the score as follows to provide an appropriate model fit. We subtracted each participant’s score from 30 and used the natural log of this value to achieve an approximate normal distribution (see Appendix 3 for graphs of the untransformed and transformed outcome). Because we used 2 separate scales of measurements (ACT and cACT), we provide the mean and 95% CI using 2 sample t tests for the change of ACT and cACT scores separately, as well as the corresponding P value from the statistical contrast. We assessed equivalence by comparing the predetermined upper limit of equivalence (change in ACT score of + 2.5) with the upper bound of the CI for the observed difference.

Since the distribution of gender between treatment arms was statistically significantly different by chance, we included it as a covariate in our model. A season of enrollment (ie, spring or fall) by treatment group interaction term was not statistically significant and neither was a season of enrollment term by itself; however, we felt it was clinically important to adjust for it even though it had minimal impact on the model.

**Secondary Outcomes.** Secondary outcomes include 1 or more occurrences of asthma exacerbation, quality-of-life scores (CHSA and PROMIS proxy and child surveys), cumulative monthly dose of ICS, missed school days, change in asthma control at 6 months, and change in pulmonary function (FEV₁) at 12 months. We captured PROMIS measure proxy and child scores every 3 months and analyzed them by repeated measures Tobit model. PROMIS scores were analyzed similar to the primary outcome, with a natural log transformation of the survey score in order to avoid violating the statistical modeling assumptions. We evaluated continuous variables with linear mixed models, where change at 12 months from baseline was the outcome variable. We adjusted the estimates of the secondary outcomes for gender, season of enrollment, and a random effect of practice site. We evaluated dichotomous outcomes such as an occurrence of an asthma exacerbation by logistic regression models and adjusted estimates for gender and season of enrollment. Practice site did not improve model fit for the
dichotomous outcomes and was not adjusted for in the logistic regression analyses. We calculated cumulative monthly dose of ICS based on the actuations of the ICS metered dose inhaler as measured by the dose counter and recorded at each study visit divided by the number of follow-up months. If an inhaler was not returned, we assumed no doses were taken. We estimated the expected monthly ICS dose in the PBA arm using the initial dose that patients were instructed to take in each age group. We compared average missed school days per year and average monthly dose of ICS using 2-sample independent \( t \) tests. We assessed the quality of each spirometry and excluded from the analysis those that did not meet American Thoracic Society criteria.\(^{38} \) We performed an exploratory analysis of time to first exacerbation using survival analysis and defining the development of an exacerbation as events. We also compared the proportion of those who had a 3-point or greater decline in ACT or cACT score at 12 months from baseline in both arms.

We performed all analyses using SAS version 9.4. An independent data safety monitoring board monitored the study.

**Missing Data.** We conducted outcome analyses including all available data from all randomized participants following the ITT principle. We treated missing data as missing at random and did not impute missing data. We measured the primary outcome (ACT/cACT scores) every 3 months and used a mixed-effects, repeated measures ANOVA model to include all the available data. Instead of imputing missing data, we used a mixed-model approach based on residual maximum likelihood to model our data. In contrast to least squares estimation, this maximum-likelihood approach adjusts parameter estimates to account for subjects with missing follow-up data. Appendix 4 contains a detailed description and examples of using this maximum likelihood approach to account for missing data, and Smith LJ, Kalhan R, Wise RA, et al.\(^{39} \) shows an example of the application of this approach in a clinical-trial setting.

**Plans to Examine the Heterogeneity of Treatment Effect (HTE).** We explored HTE by examining the interaction of the indicator variable—the intervention— with other prespecified variables. We explored the dependence of the intervention effect on baseline
characteristics such as gender, age at diagnosis, current age, parental marital status, parents’ income, mother’s employment status, number of people in household, and asthma exacerbation in past 12 months at baseline. To do this, we simply created an interaction term by multiplying the intervention variable by each listed covariate and tested the significance of the regression coefficient of the interaction term. We performed this analysis on the entire study population.

We originally planned to perform a subgroup analysis of participants with an adherence rate < 60%, as determined by the diary entry at 6 months and 12 months. However, only 60% of the participants returned the 6-month diary and 77% returned the 12-month diary; even among those who returned the diary, the quality of the data was very low, with only part of the days/sections filled out. Therefore, the study team determined that we could not perform this subgroup analysis.

Changes to the Original Study Protocol

Below are significant protocol modifications made during the study.

1. **Additional contact with asthma coaches after 6-month visits to enhance adherence to study arms:** During the study, we found that some participants were not adherent to their study arm assignment. Therefore, we sought advice from the advisory board. The board members felt that we should give additional feedback to the participants to remind them of their study assignment, but in a way that did not deviate too much from the original study plan. The board members also agreed on the importance of having participants follow their assignment correctly to ensure valid results. They suggested that the asthma coaches would be the best people to give feedback to participants, as their feedback could be part of the asthma education, which was already built into the study. Based on this input, we added an additional phone contact by the coaches after the 6-month visit. During this phone call, the coaches reviewed the instruction sheet and the study assignment given to the participants at the randomization visit. This protocol modification might have enhanced participants’ adherence to the study arm, increasing the internal validity of the study.

2. **Modification of inclusion criteria regarding asthma control at baseline:** We modified the study protocol to include those with AEQ scores of 0 to 2 in 2 out of 3 questions with
total score $\leq 5$ (previously 0-1 in all scores) to allow participants who were slightly more symptomatic to enroll. The randomization criteria remained the same (AEQ score of 0-1 in all questions), so we used the same randomization criteria for all participants, including those who were already in the study when we modified the screening protocol. The goal of this modification was to expand the study population to include well-controlled as well as partially controlled asthma patients. This protocol modification aimed to expand the study population so we could increase the external validity.

Results

Demographics of Study Participants

Between March 4, 2015, and October 20, 2016, we called 1825 potential participants. Following the initial phone screening, 311 participants were seen at the first screening visit. We randomized 206 participants to each adjustment strategy after a successful run-in period; 87 participants in the PBA arm and 92 participants in the SBA arm completed the 12-month visit. The flow of study participants is shown in Figure 3. Twenty-seven participants withdrew, were withdrawn after randomization, or were lost to follow-up. The number of such participants did not differ significantly between arms ($p = 0.30$). Five participants were withdrawn per study protocol, as they had too many asthma exacerbations (2 in 6 months or 3 in 12 months). The demographics of study participants are shown in Table 5. The average age of the participants was 10.2 years old. There was a significant gender difference between arms, with more girls in the SBA arm. Baseline PROMIS proxy score was higher in the SBA arm. Otherwise, the distribution of the participant characteristics was similar between arms. Overall, 42% of participants reported having had an asthma exacerbation in the previous year.
Figure 3. Flow of study participants

Abbreviations: FEV1, forced expiratory volume in the first second of expiration; ICU, intensive care unit; PBA, provider-based adjustment; PCP, primary care provider; Rz, randomization; SBA, symptom-based adjustment.
Table 5. Demographics and Characteristics of Study Participants at Entry

<table>
<thead>
<tr>
<th>Demographics</th>
<th>PBA (N = 103)</th>
<th>SBA (N = 103)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td>10.1 (3.2)</td>
<td>10.3 (3.4)</td>
<td>0.66</td>
</tr>
<tr>
<td>Age at diagnosis</td>
<td>3.0 (2.8)</td>
<td>2.7 (2.8)</td>
<td>0.42</td>
</tr>
<tr>
<td>% Female</td>
<td>39 (38)</td>
<td>56 (54)</td>
<td>0.02</td>
</tr>
<tr>
<td>% Parent married</td>
<td>33 (32)</td>
<td>23 (22)</td>
<td>0.12</td>
</tr>
<tr>
<td>Median household income &lt; $35 000 (%)</td>
<td>55 (53)</td>
<td>56 (61)</td>
<td>0.69</td>
</tr>
<tr>
<td>Mother employed (%)</td>
<td>80 (78)</td>
<td>74 (72)</td>
<td>0.33</td>
</tr>
<tr>
<td>Medicaid insured</td>
<td>58 (56)</td>
<td>66 (64)</td>
<td>0.63</td>
</tr>
<tr>
<td><strong>Asthma characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exacerbation in the past 12 months</td>
<td>45 (44)</td>
<td>42 (41)</td>
<td>0.67</td>
</tr>
<tr>
<td>ED visit for asthma in past 12 months</td>
<td>30 (29)</td>
<td>28 (27)</td>
<td>0.76</td>
</tr>
<tr>
<td>Number of asthma exacerbations in 12 months</td>
<td>0.7 (1.0)</td>
<td>0.8 (1.2)</td>
<td>0.65</td>
</tr>
<tr>
<td>Controller medications (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICS alone</td>
<td>89 (86)</td>
<td>88 (85)</td>
<td>0.84</td>
</tr>
<tr>
<td>Low-dose ICS/LABA</td>
<td>8 (8)</td>
<td>11 (11)</td>
<td>0.46</td>
</tr>
<tr>
<td>Daily antileukotriene</td>
<td>32 (31)</td>
<td>32 (31)</td>
<td>0.99</td>
</tr>
<tr>
<td><strong>Asthma score</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACT (n = 71; range 5-25)</td>
<td>22.0 (2.3)</td>
<td>21.1 (3.4)</td>
<td>0.18</td>
</tr>
<tr>
<td>cACT (n = 135; range 0-27)</td>
<td>21.6 (3.1)</td>
<td>21.8 (2.6)</td>
<td>0.72</td>
</tr>
<tr>
<td><strong>Quality-of-life scores, median (IQR)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHSA—Parent Version (Range = 0-100), a n = 135</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical, health</td>
<td>90 (78-95)</td>
<td>83 (75-95)</td>
<td>0.64</td>
</tr>
<tr>
<td>Activity, child</td>
<td>95 (80-100)</td>
<td>100 (85-100)</td>
<td>0.5</td>
</tr>
<tr>
<td>Activity, family</td>
<td>100 (92-100)</td>
<td>100 (92-100)</td>
<td>0.73</td>
</tr>
<tr>
<td>Emotion, child</td>
<td>100 (78-100)</td>
<td>95 (75-100)</td>
<td>0.81</td>
</tr>
<tr>
<td>Emotion, family</td>
<td>91 (80-99)</td>
<td>91 (81-100)</td>
<td>0.73</td>
</tr>
<tr>
<td>Demographics</td>
<td>PBA (N = 103)</td>
<td>SBA (N = 103)</td>
<td>P Value</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>---------------</td>
<td>---------------</td>
<td>---------</td>
</tr>
<tr>
<td>CHSA—Child Version (Range = 0-100), a n = 72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical, health, median (IQR)</td>
<td>93 (79-100)</td>
<td>89 (75-96)</td>
<td>0.1</td>
</tr>
<tr>
<td>Activities</td>
<td>95 (85-100)</td>
<td>100 (78-100)</td>
<td>0.32</td>
</tr>
<tr>
<td>Emotion</td>
<td>94 (92-100)</td>
<td>92 (87-98)</td>
<td>0.18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PROMIS Asthma Impact (T score, mean score = 50, SD 10)b</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROMIS proxy (n = 205)</td>
</tr>
<tr>
<td>PROMIS child (n = 145)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lung function</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEV₁ (L)</td>
</tr>
<tr>
<td>FEV₁, % predicted</td>
</tr>
<tr>
<td>FEV₁/FVC (%)</td>
</tr>
</tbody>
</table>

Abbreviations: ACT, Asthma Control Test (for 12 years old and above, scale = 5-25); cACT, childhood Asthma Control Test (for 6- to 11-year-olds, scale = 0-27; lower score indicates worse control in ACT and cACT); CHSA, Children's Health Survey for Asthma; ED, emergency department; FEV₁, forced expiratory volume in the first second of expiration; FVC, forced vital capacity; ICS, inhaled corticosteroid; IQR, interquartile range; LABA, long-acting beta-agonist; PBA, provider-based adjustment; PROMIS, Patient-Reported Outcomes Measurement Information System; SBA, symptom-based adjustment; SD, standard deviation.

Note: Data are mean (SD) or number (%) unless stated otherwise.

a CHSA: higher score indicates better quality of life.
b PROMIS: lower score indicates better quality of life.

**Primary Outcome**

*Change in ACT/cACT Scores at 12 Months.* Figure 4 shows the timeline of the ACT/cACT scores. At baseline, the PBA arm’s ACT score was 22.0 and cACT score was 21.6, while the SBA arm’s ACT score was 21.1 and cACT score was 21.8. At 12 months, the PBA arm’s ACT score was 23.4 and cACT score was 23.0, and the SBA arm’s ACT score was 21.8 and cACT score was 22.4, all above the established cut-off for adequately controlled asthma (a score ≤ 19 indicates that asthma is inadequately controlled).28,31,40
Figure 4. ACT and cACT scores at each visit

Panel A. ACT scores at each visit

Panel B. CACT scores at each visit

Abbreviations: ACT, Asthma Control Test; cACT, childhood Asthma Control Test; PBA, provider-based adjustment; SBA, symptom-based adjustment.

Note: Mean score and 95% CI at each visit are shown. ACT and cACT were measured every 3 months after randomization. The numbers in the 2 bottom rows indicate number of participants who attended each visit. The gray zone (score below 19) indicates inadequate asthma control.
For the PBA arm, the average change in ACT/cACT scores at 12 months from baseline was ACT 1.55 (n = 31) and cACT 1.43 (n = 56). For the SBA arm, the average change in ACT/cACT scores at 12 months from baseline was ACT 0.67 (n = 33) and cACT 0.69 (n = 59). The mean between-group difference in the change from baseline to 12 months was small for both measures, less than a 1-point score difference (ACT: –0.88 [95% CI –2.19 to 0.42]; cACT: –0.73 [95% CI –2.09 to 0.62]). In a modeled analysis of transformed combined ACT/cACT controlling for test, season, and gender with random effect of clinic, the difference between arms was not significant (p = 0.10; Table 6). The 95% CI for the true mean difference was between –2.19 and 0.42 for ACT and –2.09 and 0.62 for cACT, which is within the predefined clinical equivalence threshold of ± 2.5. Adjusting for gender and season did not have a major effect on the estimate. We found no significant difference between arms in the proportion of participants whose ACT or cACT score decreased by at least 3 points at 12 months (PBA: 12.6% vs SBA: 14.1%; p = 0.77).

Subgroup Analysis. We evaluated the interaction between the various prespecified baseline variables, including gender, age, age at diagnosis, parental marital status, median income < $35 000, mother’s employment status, number of people in household, and baseline history of asthma exacerbation in past 12 months, and the treatment indicator. None of these variables had a significant effect on the result of the primary outcome (gender: p = 0.126; age: p = 0.327; age at diagnosis: p = 0.249; parental marital status: p = 0.128; median income < $35 000: p = 0.056; mother’s employment status: p = 335; number of people in household [3-4 vs 5 or more]: p = 0.726; baseline history of asthma exacerbation in past 12 months: p = 0.270). We were unable to perform subgroup analyses based on the predefined adherence to study intervention due to lack of reliable adherence data (details provided in Study Limitations section).
Table 6. ACT/cACT Change From Baseline to 12 Months

<table>
<thead>
<tr>
<th></th>
<th>Difference (95% CI), Baseline to 12 Months&lt;sup&gt;a&lt;/sup&gt;</th>
<th>P Value&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SBA</td>
<td>PBA</td>
</tr>
<tr>
<td>ACT</td>
<td>0.67 (−0.33 to 1.66)</td>
<td>1.55 (0.68-2.42)</td>
</tr>
<tr>
<td>cACT</td>
<td>0.69 (−0.26 to 1.65)</td>
<td>1.43 (0.45-2.41)</td>
</tr>
</tbody>
</table>

Abbreviations: ACT, Asthma Control Test (for 12- to 17-year-olds); cACT, childhood Asthma Control Test (for 6- to 11-year-olds); PBA, provider-based adjustment; SBA, symptom-based adjustment.

<sup>a</sup> Differences shown are raw values. Model estimates are not obtainable due to the log transformation.

<sup>b</sup> P value is from model of transformed combined ACT/cACT controlling for test, season, and gender with random effect of clinic.

Secondary Outcomes

**Asthma Exacerbation.** The proportion of participants who had an asthma exacerbation was 23% in the PBA arm and 19% in the SBA arm (<i>p</i> = 0.62; see Table 7). The proportion of participants who had an asthma exacerbation requiring an ED visit was 11% in the PBA arm and 4% in the SBA arm, which was not significant (<i>p</i> = 0.072). The analysis of time to first exacerbation also did not show any statistically significant difference between arms (<i>p</i> = 0.49)<sup>39</sup>.

Table 7. Proportion of Exacerbation in Each Arm

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Treatment Arm</th>
<th>P Value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PBA</td>
<td>SBA</td>
</tr>
<tr>
<td><strong>Exacerbations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of events</td>
<td>29</td>
<td>22</td>
</tr>
<tr>
<td>Number of ≥ 1 events (%)</td>
<td>24 (23)</td>
<td>20 (19)</td>
</tr>
<tr>
<td><strong>Exacerbations resulting in ED visit</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of events</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>Number of ≥ 1 events (%)</td>
<td>11 (11)</td>
<td>4 (4)</td>
</tr>
</tbody>
</table>

Abbreviations: ED, emergency department; PBA, provider-based adjustment; SBA, symptom-based adjustment.

<sup>a</sup> P value is from model controlling for season and gender.
**Quality-of-Life Scores.** We measured quality of life using the PROMIS instrument and the CHSA. We found no differences between arms in PROMIS scores at 12 months in either the parent or child item. Neither the CHSA—Parent Version (used in 6- to 11-year-olds) nor the CHSA—Child Version (used in 12- to 17-year-olds) showed significant differences between arms in all domains (physical, activities, and emotional).

**Lung Function.** Participants’ percentage-predicted FEV₁ was well above the reference range (> 80% is considered normal) at baseline (see Table 5). One of our inclusion criteria was a pre-bronchodilator FEV₁ ≥ 80% of predicted value at entry. We observed a small but statistically significant decline in percentage-predicted FEV₁ at 12 months compared with baseline in both arms (PBA: –3.44%; SBA: –5.54%), but the mean change from baseline was not different between arms ($p = 0.14$; **Table 8**). Change in percentage-predicted FEV₁/FVC (forced vital capacity) was also similar between arms. The proportion of participants whose percentage-predicted FEV₁ was < 80% was also similar (PBA: 16%; SBA: 16%; $p > 0.99$).
Table 8. Secondary Outcomes: Change From Baseline to 12 Months

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Mean Difference, Baseline to 12 Months (95% CI)a</th>
<th>N PBA</th>
<th>N SBA</th>
<th>P Valueb</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CHSA—Parent</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical health</td>
<td>-0.09 (–3.41 to 3.22)</td>
<td>54</td>
<td>58</td>
<td>0.23</td>
</tr>
<tr>
<td>Activities, child</td>
<td>3.14 (–1.61 to 7.89)</td>
<td>55</td>
<td>58</td>
<td>0.37</td>
</tr>
<tr>
<td>Emotional, child</td>
<td>1.55 (–6.16 to 9.25)</td>
<td>55</td>
<td>58</td>
<td>0.57</td>
</tr>
<tr>
<td><strong>CHSA—Child</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical health</td>
<td>2.67 (–1.20 to 6.54)</td>
<td>31</td>
<td>33</td>
<td>0.93</td>
</tr>
<tr>
<td>Activities</td>
<td>4.35 (2.25-6.46)</td>
<td>31</td>
<td>33</td>
<td>0.51</td>
</tr>
<tr>
<td>Emotional</td>
<td>1.48 (–1.92 to 4.87)</td>
<td>31</td>
<td>33</td>
<td>0.95</td>
</tr>
<tr>
<td><strong>PROMIS scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROMIS proxy</td>
<td>-0.35 (–2.55 to 1.85)</td>
<td>85</td>
<td>90</td>
<td>0.86</td>
</tr>
<tr>
<td>PROMIS child</td>
<td>-4.97 (–7.34 to –2.60)</td>
<td>62</td>
<td>67</td>
<td>0.09</td>
</tr>
<tr>
<td><strong>Lung function</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEV1 % predicted</td>
<td>-3.44 (–5.53 to –1.34)</td>
<td>79</td>
<td>82</td>
<td>0.14</td>
</tr>
<tr>
<td>FEV1/FVC (%)</td>
<td>-0.05 (–1.30 to 1.19)</td>
<td>78</td>
<td>82</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Abbreviations: CHSA, Children’s Health Survey for Asthma; FEV1, forced expiratory volume in the first second of expiration; FVC, forced vital capacity; PBA, provider-based adjustment; PROMIS, Patient-Reported Outcomes Measurement Information System; SBA, symptom-based adjustment.

a Differences shown are raw values.
b P value is from model controlling for season and gender with random effect of clinic.

**Missed School Days.** We saw no significant difference in self-reported missed school days in either the PBA arm (1.05 days per year [95% CI, 0.56-1.55]) or the SBA arm (1.13 days per year [95% CI, 0.61-1.65]; p = 0.84).

**Changes in Asthma Control at 6 Months.** The changes in ACT/cACT scores at 6 months showed no statistically significant difference (p = 0.71).

**Participant Satisfaction and Sense of Self-management at 12 Months.**
Satisfaction was assessed at 12 months by both parent and child (> 12 years old). The
proportion of those who reported they were satisfied with the treatment strategy was similar between arms (child very satisfied: PBA 80% vs SBA 80%, \( p = 0.99 \); caregiver very satisfied: PBA 92% vs SBA 97%, \( p = 0.20 \)). More children and/or their caregivers in the SBA arm felt that they rather than their PCP were managing their asthma (see Table 4; child: PBA 44% vs SBA 72%, \( p = 0.044 \); parent: PBA 40% vs SBA 63%, \( p = 0.001 \)). Fewer parents and participants in the PBA arm correctly identified the strategy they were assigned to (parent: PBA 69% vs SBA 93%, \( p < 0.001 \); child: PBA 78% vs SBA 92%, \( p = 0.09 \)).

**Adherence to Study Arm.** More parents in the PBA arm reported that they were not following the PBA instructions (use of daily ICS) than parents in the SBA arm (28% vs 14%; \( p = 0.023 \)). We attempted to obtain estimates of adherence by asking the participants to complete diaries of symptoms and medication use at 6 and 12 months for 4 weeks. However, the completion rate for the diaries was very low and the quality of the diary data was not sufficient to conduct reliable analysis on adherence. Therefore, we did not use the information from the diaries in the analysis.

**Average Monthly Dose of ICS Used.** Randomized participants received 2235 beclomethasone inhalers, 1778 of which we recovered (80%) to assess ICS use by the dose counter. The recovery rate of inhalers did not differ between the 2 arms (PBA 81% vs SBA 79%, \( p = 0.24 \)). On average, PBA arm participants used 1961.3 mcg (95% CI, 1681.5-2241.1 mcg) of beclomethasone (approximately 60% of the expected dose of 3289.4 mcg) and SBA arm participants used 526.2 mcg (95% CI, 412.8-639.5 mcg) per month (see Figure 5). SBA arm participants used significantly less ICS than expected (\( p < 0.0001 \)). This pragmatic study did not measure adherence.
Adverse Events

Adverse events were similar between arms. Details are provided in Appendix 5 (Adverse Events). Five participants (3 in the PBA arm and 2 in the SBA arm) were withdrawn from the study per protocol as they had too many exacerbations (2 in 6 months or 3 in 12 months). One participant in the PBA arm had thrush, a complication of ICS. Four hospitalizations were recorded: 3 in the PBA arm (1 asthma exacerbation requiring hospitalization, 1 pneumonia requiring hospitalization, and 1 hospitalization due to chronic constipation) and 1 in the SBA arm (an asthma exacerbation requiring hospitalization).

Exploratory Analysis: Narrative and SBA Participants’ Preference at the End of the Study

End of the Study Narratives. We asked all participants and parents in both arms to provide written comments regarding the treatment strategy they received. Many parents stated that they appreciated the education the study provided to help control their children’s asthma. One parent wrote, “It helped to educate us more about asthma and its treatment.” Most parents in the PBA arm mentioned they liked the treatment arm and that they had no
issues. One parent wrote, “I like the program. It helped to educate us more about asthma and its treatment.” Another parent wrote, “It was a plan that we were pretty much used to. No issues at hand.” One parent admitted that her child did not take daily ICS: “Yes, the study was great; however, he did not take his Qvar [beclomethasone] daily.” Many parents in the SBA arm also mentioned that they liked the strategy and had no issues; however, several comments were unique to the SBA arm. One parent wrote, “I like it because it taught me a lot about what to look for when she [my child] was having symptoms.” The following are a few more comments from parents: “I liked it; easy and another way to control his asthma.” “We were assigned to as-needed group. I liked this group because I had control on when to give meds.” One child in the SBA arm wrote, “I liked it because I didn’t have to constantly take it every day.”

To explore SBA participants’ preference, we asked participants (older than 12 years) and parents the following question: “After you experienced both ways of treating your asthma, which do you prefer: symptom-based adjustment or what you were told to do by your doctor before the study?” Fifty-two parents and 25 children answered the question—34 parents (65%) said they preferred SBA, 9 (17%) were neutral, and 9 (17%) preferred the prior method; 19 children (76%) said they preferred SBA and 6 (24%) said they were neutral.

Provider Exit Interview. We completed the provider interview with 10 providers after the study’s completion. During the interview, we discussed (1) how SBA did and did not work for their patients, (2) barriers to implementation, (3) suggestions on dissemination of study results, and (4) their perspective on how things went during the study. The providers unanimously stated that they benefited from participating in ASIST. Many providers said they heard positive feedback from many of their patients who participated in ASIST. For some providers, logistics for the study visits were confusing, especially at the beginning, but they stated that overall the study worked out well. Many providers said they had kept patients on SBA even after the trial was over if the patient did well with it during the study. Providers identified the following possible barriers to implementing SBA for all patients in their practice: (1) not enough evidence yet, (2) not in the current NAEPP guideline, (3) lack of
resources to provide education for alternative management strategy, (4) uncertainty in patient selection, and (5) lack of patient understanding and adherence.
DISCUSSION

Context for Study Results

In a community primary care–based RCT, we demonstrated that asthma control using symptom-based, intermittent use of ICS with a rescue bronchodilator (ie, SBA) among African American children with mild asthma was similar to using provider-based asthma management (ie, PBA). Other secondary outcomes, including the proportion of those with asthma exacerbation, quality of life (measured by the CHSA), and lung function measurements (percentage-predicted FEV₁ and FEV₁/FVC), were also similar between the 2 arms. This result was achieved with fewer than 30% of the exposure to ICS used in the SBA arm participants. The mean difference was much smaller than the reported minimally important clinical difference for the ACT and cACT measures of asthma control.³⁰

Asthma guidelines recommend daily use of ICS for persistent asthma.¹,⁹ However, adherence to daily asthma therapy is often challenging and remains low despite decades of efforts to improve adherence,¹⁶ especially in vulnerable minority populations. Because asthma symptoms are often episodic, patients often discontinue taking their daily ICS shortly after the exacerbation,¹⁹ and many rely on ED visits to receive asthma care.¹⁵ In addition to poor adherence to daily controller therapy, concerns about adverse effects (eg, impaired growth,⁴¹ long-term steroid-induced complications⁴²) exist, and these concerns negatively affect adherence both in adults⁴³ and children.²³ Also, we and other authors have shown that medication cost for asthma therapy is one of the major concerns of urban minority families.²³,²⁶ Reflecting this concern and the beliefs of patients and parents, PCPs often elect not to recommend daily ICS to half of their patients with persistent asthma,⁴⁴ especially those with mild persistent asthma.¹³ These ongoing obstacles to guideline-based asthma management have led to consideration of alternative, nondaily strategies that have been evaluated in clinical trials in the past decade.

Among alternative nondaily asthma management strategies, intermittent, as-needed use of ICS is an emerging, patient-centered strategy in which the dose of ICS is adjusted based on a patient’s rescue inhaler needs. ICS is delivered with a short-acting beta-agonist (SABA;
albuterol) or short-onset beta-agonist (budesonide/formoterol) with the onset of symptoms, and patients take the initiative to manage their dose of ICS. Intermittent, as-needed use of ICS with SABA has proven to be effective in controlling asthma and preventing asthma exacerbation in adults and children with mild persistent asthma. More recently, in a large study in patients with mild persistent asthma, a similar strategy—intermittent, as-needed use of ICS with short-onset beta-agonist (budesonide/formoterol)—was shown to deliver better asthma control than SABA only and resulted in a similar rate of asthma exacerbation compared with maintenance daily dose of budesonide/formoterol.

Our study builds on and extends the conclusions of previous studies and is the first to our knowledge to evaluate the effectiveness of the as-needed use of ICS with short-acting bronchodilator compared with a guideline-based strategy in a real-world setting. Our study went one step further than the previous studies, in which highly selected asthma participants managed both daily and as-needed use of inhalers in a blinded fashion. The open-label design of our study allowed the provider and patient to use their asthma controllers as they do in typical practice, and the providers implemented both strategies (PBA and SBA) following routine practice. The participating providers of AA children addressed management issues during the 12-month study for both treatment strategies, including worsening of asthma control, nonadherence, no-shows to clinic visits, and treatment and follow-up of asthma exacerbations. In fact, in ASIST, the estimated adherence to the guideline-based daily ICS strategy (ie, PBA) was approximately 60%, despite counseling to help patients adhere to the daily schedule in a study setting. We designed our pragmatic trial to evaluate an alternative asthma management strategy (ie, SBA) in a primary care setting at the location where patients’ usual medical care takes place. We believe that the ASIST study was able to answer a question that explanatory RCTs could not address and played an important role in filling the gap between the controlled trials and real-world setting when implementing this alternative strategy.

Our findings suggest there is little difference in real-world effectiveness between as-needed use of ICS with short-acting bronchodilator (SBA) and the use of guideline-based daily ICS (PBA) in mild asthma and indicates that intermittent, as-needed use of ICS with short-acting
bronchodilator can be an alternative strategy to daily ICS use in AA children. It is important to point out several factors that PCPs should consider when discussing the choice of strategy with their patients. The first is adherence to guideline-based, daily ICS strategy. In our community-based study, the overall estimated adherence to the PBA arm was approximately 60%, which was consistent with other asthma studies. However, when treating individual patients who are more adherent and willing to take daily ICS (current standard of care), the PBA strategy may be considered preferable to achieve asthma control. Another factor to consider is patient/parent preferences and beliefs about ICS. Negative beliefs about using ICS contribute to nonadherence and poor asthma control, and concern about the adverse effects of ICS are especially prevalent in minority populations. Therefore, an intermittent, as-needed strategy may be more acceptable for those with strong negative beliefs about ICS. In our focus group study conducted in caregivers with African American asthmatic children, caregivers and their providers thought that an intermittent, as-needed ICS strategy was in line with their goal of using less medicine. Future studies that include patient/parent preference in the choice of management strategy are needed. In our study, we randomly assigned the treatment strategies, and it was not within the scope of ASIST to include patient preference when assigning participants to treatment arms. Finally, the strategy tested in our study is as-needed, symptom-based use of ICS with short-acting bronchodilator and is different from intermittent use of ICS at higher doses or for a fixed period around the time of asthma exacerbation. Studies of intermittent use of ICS at higher doses during the early symptoms of asthma exacerbation have not been shown to reduce the risk of subsequent asthma exacerbations.

Secondary outcomes including asthma exacerbation and quality of life were also similar between arms, supporting that there was little difference in patient-reported outcomes. A small but significant decrease from baseline lung function occurred at 12 months in both study arms, although the between-group difference was not significant. This decrease may be attributed in part to measuring baseline lung function right after the run-in period, during which time the participants were instructed to take daily ICS with close monitoring. Lung function was measured only once after the treatment assignments were made, so we are
unable to conclude whether this decrease represents a true long-term decline. A longer study with serial lung function measurements would be needed to answer that question.

**Generalizability of the Findings**

We conducted our study in African American children with mild asthma. Racial differences in the response to as-needed, symptom-based use of ICS with short-acting bronchodilator have not been studied, and there is no scientific rationale to believe that symptom-based use of ICS works for only a particular race. Moreover, most previous studies have included both white and AA populations. Because ASIST included only children with asthma, it was not within the scope of the study to assess applicability of SBA to adults with asthma in the primary care setting. However, most prior studies of SBA have been conducted in adults and have showed that SBA is as effective as PBA. Symptom-based use of ICS may be easier to implement in adults, as they can directly act on their own symptoms. In the ASIST study, all participants received education from a peer coach about appropriate use of albuterol via 2 to 4 phone calls, with an additional call for reinforcement at 6 months. This type of peer education over several conversations is not a part of usual care and may limit the generalizability of our findings. In ASIST, we provided all study inhalers (beclomethasone and albuterol) to the participants to avoid access to medication as a confounder. This could affect generalizability of the findings in a real-world setting; however, it is unlikely that this procedure had a major effect on our findings. If we had not provided the study medication, adherence to the treatment strategies would have likely declined in both arms, decreasing the validity of the study findings.

**Implementation of Study Results**

Although our study went a step further than previous RCTs and showed the effectiveness of as-needed, symptom-based use of ICS with short-acting bronchodilator (ie, SBA), several important barriers remain to implementing this strategy widely into general practice. One significant barrier is lack of inclusion of intermittent, as-needed therapy in the NAEPP guideline.¹ For full acceptance from the PCP community, incorporation of SBA into the standard of care is especially important. It is anticipated that NIH will publish an update of the
NAEPP guideline in late 2020. In our focus group study, caregivers and providers emphasized that patient education about symptom recognition as well as close communication between physician and patient would facilitate SBA implementation. At our exit interview with the participating providers, they identified several barriers to implementing SBA into their practice, which include (1) not enough evidence, (2) not in the current NAEPP guideline, (3) lack of resources to provide education for alternative management strategy, (4) uncertainty in patient selection, and (5) lack of patient understanding and adherence. Developing educational materials for PCPs so that they can implement SBA will be the next step. Finally, electronic remote monitoring of inhaler use could be considered as more options for this technology become available. This technology could greatly benefit PCPs in assessing their patients’ adherence to SBA.

Subpopulation Considerations

Baseline characteristics that may affect asthma control (gender, age, age of asthma onset, parental marital status, median income < $35 000, mother’s employment status, number of people in household, and history of asthma exacerbation in past 12 months) did not have significant interaction with the results of our primary outcome. We were unable to measure estimates of participant adherence to the study intervention (details stated below), so the effect of adherence on the outcome is unknown.

Study Limitations

First, we did not include a placebo group in our study design, which made comparing the effectiveness of our study arms vs placebo impossible. We chose not to include a placebo because we believed it would be unethical to withhold care and because withdrawal of asthma maintenance therapy has been shown to have negative outcomes in children with mild persistent asthma. Our study population included children who were receiving daily asthma controller medications (ICS/LTRA) prescribed by the participating providers. Although not all the enrolled children were adherent to the prescribed daily medications and might not have required full daily dose to control their asthma at enrollment, we honored the treatment
decision of the providers in this pragmatic trial in which the PCP implemented the study arms following routine practice.

Second, our open-label design could have affected the internal validity of the trial, but the providers needed to know the randomly assigned treatment arm in order to manage the study participants under their care, and the participants/parents needed to follow the instructions for SBA in a real-world setting. To avoid measurement bias, study staff who were blinded to the study arm assignment independently obtained all outcome measurements and used validated questionnaires.

We were unable to accurately measure adherence to the study arms to evaluate the effect of adherence on the study outcomes. We considered multiple options for adherence monitoring in the study design phase but determined that it was not feasible to accurately evaluate adherence to the study intervention in this pragmatic trial for the following reasons: (1) The definition of perfect adherence to PBA and SBA is fundamentally different, which makes a valid comparison difficult; (2) it is not possible to fully capture the presence of symptoms (especially in the SBA arm) since it can be measured only by self-report; and (3) requiring a daily diary of medication-taking for 12 months is not feasible in a primary care setting. Therefore, we attempted to conduct an exploratory analysis using a short-term written diary for 4 weeks at 6 months and 12 months as a surrogate marker. Unfortunately, compliance with and quality of the diary entries were insufficient to conduct an exploratory analysis that would support reliable conclusions.

Finally, our study included mainly children with mild asthma, so the effectiveness of SBA in those with moderate asthma with daily symptoms or FEV$_1 <$ than 80% was not tested. However, most of the asthma population has mild symptoms$^{25}$; in fact, our screening data confirm the high prevalence of mild asthma in the practices of PCPs. We observed that only a few patients were excluded because their asthma was too severe to participate in our study when we conducted our initial phone screening for eligibility. Therefore, we believe the findings of our study are highly relevant to PCPs.
Future Research

Future research should include 2 main areas not included in our study. The first would be a study expanding the as-needed, symptom-based use of the ICS strategy (ie, SBA) to a broader population. A larger pragmatic study including other races, adults, and other geographical regions would be necessary to expand this alternative strategy to adjust to multiple different populations. The second area would be studies to address barriers to implementing this strategy. Notably, our participating PCPs identified the need for provider-friendly education tools to teach patients how to recognize symptoms and how to conduct SBA. In ASIST, the education was delivered by our asthma coaches and research staff, but researchers need to develop and test a tool that is not resource-intense for providers. Engagement of stakeholders, including providers and patient partners, is necessary to develop and implement such a user-friendly educational tool.
CONCLUSIONS

Our study found that asthma control and rates of asthma exacerbation with the use of intermittent, symptom-based use of inhaled corticosteroids with a rescue bronchodilator (ie, SBA) in the primary care setting among African American children with mild asthma was not inferior to that achieved with provider-based asthma management (ie, PBA) and with much lower ICS exposure. Patients and their caregivers who practiced SBA reported a greater sense of self-management in managing their asthma than those practicing PBA. This patient-centered strategy can be considered as an alternative for managing mild asthma in African American children, especially in those who have difficulty adhering to daily ICS therapy.
REFERENCES


ACKNOWLEDGMENTS

We are grateful to our study advisory board members (Lisa Henry, Laura Turner, Barbara White, Alison Nash, Shadrach Martin, Taylor Prince, and Shawni Jackson-Triggs) for providing many valuable contributions to ASIST. We also thank the participating primary care providers and their staff for implementing the study:

- St Louis Tots Thru Teens Pediatric: Isabel Rosenbloom, MD
- St. Louis Pediatric Practitioners: Alison Nash, MD; Diane Schmitt, NP; and Andrea Mouldon, NP
- Brownridge Pediatrics: Seth Brownridge, MD; and Terianne Lindsey, NP
- Johnson Pediatrics: Denise Johnson, MD
- Esse Health Florissant: John Galgani, MD; and Shirley Knight, MD
- Strashun Pediatrics: Robert Strashun, MD
- Rainbow Pediatrics: April Tyus, MD; and Anita Stiffelman, MD
- Edgemont Pediatrics: Justin Ogbevoen, MD
- Beeks Pediatrics: Earl Beeks, MD
- Joyce Johnson Pediatrics: Joyce Johnson, MD

We also thank Charis Johnson and Yulanda Tomin-Watson for providing phone education to the participants. Additionally, we thank the ASIST data safety monitoring board members (James Wedner, MD; Stephanie Fritz, MD; Linda Rogers, MD; and Yan Yan, PhD) for monitoring ASIST. Finally, we would like to the thank the ASIST participants and families for their participation and support throughout the study.
APPENDICES

Appendix 1: ASIST Asthma Evaluation Questionnaire (AEQ)
Appendix 2: Satisfaction Questionnaire Child and Parent Version
Appendix 3: Histogram of Transformed ACT/cACT Data
Appendix 4: Allison Missing Data Reference
Appendix 5: Adverse Events in ASIST
Appendix 6: Participant With Missing Data
Appendix 1: ASIST Asthma Evaluation Questionnaire (AEQ)

Please consider your (you refers to the child) last 2 weeks of asthma control in answering these questions. Check the number next to the response that best describes the your (child’s) asthma symptoms.

1. In the past two weeks, how often have you experienced asthma symptoms?
   0. Less than or equal to 2 days a week
   1. 3-4 days a week
   2. 5 or more days per week, but not continual
   3. Continual (multiple time every day)

2. In the past two weeks, how often have you used you rescue beta-agonist medicine (e.g. albuterol (Proventil, Ventolin, Proair)) aside from preventive use prior to exercise?
   0. Less than or equal to 2 days a week
   1. 3-4 days a week
   2. 5 or more days per week, but not continual
   3. Continual (multiple time every day)

3. In the past two weeks, how often have you awakened at night due to asthma symptoms?
   0. No awakening or awakened one night during the 2 weeks
   1. 1 night per week
   2. 2 or 3 nights per week
   3. 4 or more nights per week
Appendix 2: ASIST Satisfaction and Adherence Child Questionnaire (at 12 month)

1. What were the good things about your treatment plan you were assigned to? (check all apply)
   1. Easy to follow the instruction
   2. Better asthma control (better symptoms)
   3. Less medicine
   4. Better quality of life (be able to do things you want to do)
   5. More involvement on treatment adjustments by you
   6. Being preventive (avoid asthma attacks)
   7. Others (please state:  )

2. What were the bad things about your treatment plan you were assigned to? (check all apply)
   1. Difficult to take the medication as instructed (taking it every day, or taking them with albuterol)
   2. Worse asthma control (worse symptoms)
   3. More medicine
   4. Worse quality of life (not able to do things you want to do)
   5. Less involvement of on treatment adjustments by you
   6. Not being preventive (not avoiding asthma attack)
   7. Others (please state:  )

3. What was your over all satisfaction on your treatment assignment?
   1. Very satisfied (80-100%)
   2. Mostly satisfied (50-79%)
   3. Mostly not satisfied (20-40%)
   4. Not satisfied (below 20%) (please state reason:  )

4. Do you think you will continue to take your medicine as you were told during the study, even after the study is over?
   Yes
   No

5. Would you recommend your treatment plan to another child or friend with asthma?
   Yes
   No
6. Who do you feel took charge in adjusting your asthma medicine during the study? Was it you (and your caregiver, like your mom or dad), or your doctor, or both? Please select one best answer.

I feel:

1. I (and/or my caregiver) took charge all the time (100% me)
2. I (and/or my caregiver) adjusted most of the time (~75% me), but sometimes my Doctor (~25% doctor)
3. I (and/or my child) and my doctor both adjusted equally (50% each)
4. My doctor adjusted most of the time (~75% doctor), but I (and/or my child) also did sometimes (~25% me)
5. My doctor did all the time (100% doctor) and we just followed

7. During the study, you followed (or tried to follow) which direction during the study?
   1. Take Qvar every day as directed by your doctor
   2. Take albuterol and Qvar as needed based on symptoms
   3. Others (please state: [ ])

8. Regarding your given treatment plan above, how much do you think you were able to follow the instruction (do what your were supposed to do with their inhalers) ?
   1. Your followed mostly every day (or 75-100% of the time)
   2. You followed few times a week (~50% of the time)
   3. Your followed few times a month (~25% of the time)
   4. Did not do what you were told to do most of the time (less than 25%)

9. Can you give us any comments about the treatment plan you were assigned to.
   Any comments are welcomed.
   -Did you like it, not like it?
   -Any issues you had?

   [ ]
10. **How was participating in our study?**
   - *How was the study visits?*
   - *How were our staff?*
   - *Any thing you liked, or did not like?*

Any feedback is welcomed!
Thank you for completing the ASIST study. We would like to hear about your experience during the study.

1. First, we would like to know how treated your asthma before entering the study. 
BEFORE entering ASIST study, what did your doctor tell you to do for your asthma?
   1. Take asthma medicine (controller inhaler and/or pills) every day whether or not you have symptoms
   2. Take asthma medicine (controller inhaler and/or pills) only when you need them
   3. Others: Please state here

2. DURING the study, if you did symptom based adjustment (take albuterol 2 puffs plus Qvar 2 puffs when your had symptoms of asthma), please answer the following questions.

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>During the study, the way I managed my asthma (symptom base adjustment) was different from what I was told to do before entering the ASIST study</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Symptom based adjustment was similar to what I was told to do before entering the ASIST study</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

3. After you experienced both ways of treating your asthma, which do you prefer, symptom-based adjustment or what were told to do by your doctor before the study? Please circle one.

1. I liked the symptom-based much more
2. I liked the symptom-based slightly better
3. Neutral
4. I liked what I was told to do before slightly better than symptom base
5. I liked what I was told to do before much more

Please state the reason that you prefer symptom-base or the previous way.

____________________________________________________________________

____________________________________________________________________

_________________
Appendix 3: Histogram of Transformed ACT/cACT Data

Histogram of original ACT/cACT score
Box plot of original ACT/cACT score by treatment group
Histogram of transformed ACT/cACT score
Box plot of transformed ACT/cACT score by treatment group
Supplement:

Time to first exacerbation

\[
\begin{align*}
\text{Survival Curves of Both Groups} \\
\text{With Number of Subjects at Risk}
\end{align*}
\]

<table>
<thead>
<tr>
<th>Months since Randomization</th>
<th>PBA</th>
<th>SBA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>103</td>
<td>103</td>
</tr>
<tr>
<td>3</td>
<td>89</td>
<td>93</td>
</tr>
<tr>
<td>6</td>
<td>80</td>
<td>89</td>
</tr>
<tr>
<td>9</td>
<td>77</td>
<td>84</td>
</tr>
<tr>
<td>12</td>
<td>46</td>
<td>51</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: Time to follow-up for those with no exacerbation is from randomization until the latest date of ACT or cACT.

P-value: 0.49
APPENDIX 4:
Handling Missing Data by Maximum Likelihood
Paul D. Allison, Statistical Horizons, Haverford, PA, USA

ABSTRACT
Multiple imputation is rapidly becoming a popular method for handling missing data, especially with easy-to-use software like PROC MI. In this paper, however, I argue that maximum likelihood is usually better than multiple imputation for several important reasons. I then demonstrate how maximum likelihood for missing data can readily be implemented with the following SAS® procedures: MI, MIXED, GLIMMIX, CALIS and QLIM.

INTRODUCTION
Perhaps the most universal dilemma in statistics is what to do about missing data. Virtually every data set of at least moderate size has some missing data, usually enough to cause serious concern about what methods should be used. The good news is that the last twenty five years have seen a revolution in methods for handling missing data. The new methods have much better statistical properties than traditional methods, while at the same time relying on weaker assumptions.

The bad news is that these superior methods have not been widely adopted by practicing researchers. The most likely reason is ignorance. Many researchers have barely even heard of modern methods for handling missing data. And if they have heard of them, they have little idea how to go about implementing them. The other likely reason is difficulty. Modern methods can take considerably more time and effort, especially with regard to start-up costs. Nevertheless, with the development of better software, these methods are getting easier to use every year.

There are two major approaches to missing data that have good statistical properties: maximum likelihood (ML) and multiple imputation (MI). Multiple imputation is currently a good deal more popular than maximum likelihood. But in this paper, I argue that maximum likelihood is generally preferable to multiple imputation, at least in those situations where appropriate software is available. And many SAS users are not fully aware of the available procedures for using maximum likelihood to handle missing data.

In the next section, we'll examine some assumptions that are commonly used to justify methods for handling missing data. In the subsequent section, we'll review the basic principles of maximum likelihood and multiple imputation. After I present my arguments for the superiority of maximum likelihood, we'll see how to use several different SAS procedures to get maximum likelihood estimates when data are missing.

ASSUMPTIONS
To make any headway at all in handling missing data, we have to make some assumptions about how missingness on any particular variable is related to other variables. A common but very strong assumption is that the data are missing completely at random (MCAR). Suppose that only one variable Y has missing data, and that another set of variables, represented by the vector X, is always observed. The data are missing completely at random (MCAR) if the probability that Y is missing does not depend on X or on Y itself (Rubin 1976). To represent this formally, let R be a “response” indicator having a value of 1 if Y is missing and 0 if Y is observed. MCAR means that

\[ \Pr(R = 1 \mid X, Y) = \Pr(R = 1) \]

If Y is a measure of delinquency and X is years of schooling, MCAR would mean that the probability that data are missing on delinquency is unrelated to either delinquency or schooling. Many traditional missing data techniques are valid only if the MCAR assumption holds.

A considerably weaker (but still strong) assumption is that data are missing at random (MAR). Again, this is most easily defined in the case where only a single variable Y has missing data, and another set of variables X has no missing data. We say that data on Y are missing at random if the probability that Y is missing does not depend on Y, once we control for X. Formally, we have

\[ \Pr(R = 1 \mid X, Y) = \Pr(R = 1 \mid X) \]

where, again, R is the response indicator. Thus, MAR allows for missingness on Y to depend on other variables that are observed. It just cannot depend on Y itself (after adjusting for the observed variables).

Continuing our example, if Y is a measure of delinquency and X is years of schooling, the MAR assumption would be satisfied if the probability that delinquency is missing depends on years of schooling, but within each level of schooling, the probability of missing delinquency does not depend on delinquency.
In essence, MAR allows missingness to depend on things that are observed, but not on things that are not observed. Clearly, if the data are missing completely at random, they are also missing at random.

It is straightforward to test whether the data are missing completely at random. For example, one could compare men and women to test whether they differ in the proportion of cases with missing data on income. Any such difference would be a violation of MCAR. However, it is impossible to test whether the data are missing at random, but not completely at random. For obvious reasons, one cannot tell whether delinquent children are more likely than nondelinquent children to have missing data on delinquency.

What if the data are not missing at random (NMAR)? What if, indeed, delinquent children are less likely to report their level of delinquency, even after controlling for other observed variables? If the data are truly NMAR, then the missing data mechanism must be modeled as part of the estimation process in order to produce unbiased parameter estimates. That means that, if there is missing data on Y, one must specify how the probability that Y is missing depends on Y and on other variables. This is not straightforward because there are an infinite number of different models that one could specify. Nothing in the data will indicate which of these models is correct. And, unfortunately, results could be highly sensitive to the choice of model. A good deal of research has been devoted to the problem of data that are not missing at random, and some progress has been made. Unfortunately, the available methods are rather complex, even for very simple situations.

For these reasons, most commercial software for handling missing data, either by maximum likelihood or multiple imputation, is based on the assumption that the data are missing at random. But near the end of this paper, we’ll look at a SAS procedure that can do ML estimation for one important case of data that are not missing at random.

**MULTIPLE IMPUTATION**

Although this paper is primarily about maximum likelihood, we first need to review multiple imputation in order to understand its limitations. The three basic steps to multiple imputation are:

1. Introduce random variation into the process of imputing missing values, and generate several data sets, each with slightly different imputed values.
2. Perform an analysis on each of the data sets.
3. Combine the results into a single set of parameter estimates, standard errors, and test statistics.

If the assumptions are met, and if these three steps are done correctly, multiple imputation produces estimates that have nearly optimal statistical properties. They are consistent (and, hence, approximately unbiased in large samples), asymptotically efficient (almost), and asymptotically normal.

The first step in multiple imputation is by far the most complicated, and there are many different ways to do it. One popular method uses linear regression imputation. Suppose a data set has three variables, X, Y, and Z. Suppose X and Y are fully observed, but Z has missing data for 20% of the cases. To impute the missing values for Z, a regression of Z on X and Y for the cases with no missing data yields the imputation equation

\[ \hat{Z} = b_0 + b_1X + b_2Y \]

Conventional imputation would simply plug in values of X and Y for the cases with missing data and calculate predicted values of Z. But those imputed values have too small a variance, which will typically lead to bias in many other parameter estimates. To correct this problem, we instead use the imputation equation

\[ \hat{Z} = b_0 + b_1X + b_2Y + sE \]

where E is a random draw from a standard normal distribution (with a mean of 0 and a standard deviation of 1) and s is the estimated standard deviation of the error term in the regression (the root mean squared error). Adding this random draw raises the variance of the imputed values to approximately what it should be and, hence, avoids the biases that usually occur with conventional imputation.

If parameter bias were the only issue, imputation of a single data set with random draws would be sufficient. Standard error estimates would still be too low, however, because conventional software cannot take account of the fact that some data are imputed. Moreover, the resulting parameter estimates would not be fully efficient (in the statistical sense), because the added random variation introduces additional sampling variability.

The solution is to produce several data sets, each with different imputed values based on different random draws of E. The desired model is estimated on each data set, and the parameter estimates are simply averaged across the multiple runs. This yields much more stable parameter estimates that approach full efficiency.
With multiple data sets we can also solve the standard error problem by calculating the variance of each parameter estimate across the several data sets. This "between" variance is an estimate of the additional sampling variability produced by the imputation process. The "within" variance is the mean of the squared standard errors from the separate analyses of the several data sets. The standard error adjusted for imputation is the square root of the sum of the within and between variances (applying a small correction factor to the latter). The formula (Rubin 1987) is:

\[
\sqrt{\frac{1}{M} \sum_{k=1}^{M} s_k^2 + \left(1 + \frac{1}{M}\right) \left(\frac{1}{M-1}\right) \sum_{k=1}^{M} (a_k - \bar{a})^2}
\]

In this formula, \( M \) is the number of data sets, \( s_k \) is the standard error in the \( k \)th data set, \( a_k \) is the parameter estimate in the \( k \)th data set, and \( \bar{a} \) is the mean of the parameter estimates. The factor \((1+1/M)\) corrects for the fact that the number of data sets is finite.

How many data sets are needed? With moderate amounts of missing data, five are usually enough to produce parameter estimates that are more than 90% efficient. More data sets may be needed for good estimates of standard errors and associated statistics, however, especially when the fraction of missing data is large.

THE BAYESIAN APPROACH TO MULTIPLE IMPUTATION

The method just described for multiple imputation is pretty good, but it still produces standard errors that are a bit too low, because it does not account for the fact that the parameters in the imputation equation \((b_0, b_1, b_2, \text{ and } s)\) are only estimates with their own sampling variability. This can be rectified by using different imputation parameters to create each data set. The imputation parameters are random draws from an appropriate distribution. To use Bayesian terminology, these values must be random draws from the posterior distribution of the imputation parameters.

Of course, Bayesian inference requires a prior distribution reflecting prior beliefs about the parameters. In practice, however, multiple imputation almost always uses non-informative priors that have little or no content. One common choice is the Jeffreys prior, which implies that the posterior distribution for \( \beta \) is the same as the prior distribution, which is unusual. The problem arises when data are missing on one or more of the potential predictors, \( Z \) and \( X \), used in imputing \( Z \). Then no regression that we can actually estimate utilizes all of the available information about the relationships among the variables. Iterative methods of imputation are necessary to solve this problem.

There are two major iterative methods for doing multiple imputation for general missing data patterns: the Markov chain Monte Carlo (MCMC) method and the fully conditional specification (FCS) method. MCMC is widely used for Bayesian inference (Schafer 1997) and is the most popular iterative algorithm for multiple imputation. For linear regression imputation, the MCMC iterations proceed roughly as follows. We begin with some reasonable starting values for the means, variances, and covariances among a given set of variables. For example, these could be obtained by listwise or pairwise deletion. We divide the sample into subsamples, each having the same missing data pattern (i.e., the same set of variables present and missing). For each missing data pattern, we use the starting values to construct linear regressions for imputing the missing data, using all the observed variables in that pattern as predictors. We then impute the missing values, making random draws from the simulated error distribution as described above, which results in a single "completed" data set. Using this data set with missing data imputed, we recalculate the means, variances and covariances, and then make a random draw from the posterior distribution of these parameters. Finally, we use these drawn parameter values to update the linear regression equations needed for imputation.

This process is usually repeated many times. For example, PROC MI runs 200 iterations of the algorithm before selecting the first completed data set, and then allows 100 iterations between each successive data set. So producing the default number of five data sets requires 600 iterations (each of which generates a data set). Why so many iterations? The first 200 ("burn-in") iterations are designed to ensure that the algorithm has converged to the correct posterior distribution. Then, allowing 100 iterations between successive data sets gives us confidence that the imputed values in the different data sets are statistically independent. In my opinion, these numbers are far larger than necessary for the vast majority of applications.
If all assumptions are satisfied, the MCMC method produces parameter estimates that are consistent, asymptotically normal, and almost fully efficient. Full efficiency would require an infinite number of data sets, but a relatively small number gets you very close. The key assumptions are, first, that the data are missing at random. Second, linear regression imputation implicitly assumes that all the variables with missing data have a multivariate normal distribution.

The FCS Algorithm

The main drawback of the MCMC algorithm, as implemented in PROC MI, is the assumption of a multivariate normal distribution. While this works reasonably OK even for binary variables (Allison 2006), it can certainly lead to implausible imputed values that will not work at all for certain kinds of analysis. An alternative algorithm, recently introduced into PROC MI in Release 9.3, is variously known as the fully conditional specification (FCS), sequential generalized regression (Raghunathan et al. 2001), or multiple imputation by chained equations (MICE) (Brand 1999, Van Buuren and Oudshoorn 2000). This method is attractive because of its ability to impute both quantitative and categorical variables appropriately. It allows one to specify a regression equation for imputing each variable with missing data—usually linear regression for quantitative variables, and logistic regression (binary, ordinal, or unordered multinomial) for categorical variables. Under logistic imputation, imputed values for categorical variables will also be categorical. Some software can also impute count variables by Poisson regression.

Imputation proceeds sequentially, usually starting from the variable with the least missing data and progressing to the variable with the most missing data. At each step, random draws are made from both the posterior distribution of the parameters and the posterior distribution of the missing values. Imputed values at one step are used as predictors in the imputation equations at subsequent steps (something that never happens in MCMC algorithms). Once all missing values have been imputed, several iterations of the process are repeated before selecting a completed data set.

Although attractive, FCS has two major disadvantages compared with the linear MCMC method. First, it is much slower, computationally. To compensate, the default number of iterations between data sets in PROC MI is much smaller (10) for FCS than for MCMC (100). But there is no real justification for this difference. Second, FCS itself has no theoretical justification. By contrast, if all assumptions are met, MCMC is guaranteed to converge to the correct posterior distribution of the missing values. FCS carries no such guarantee, although simulation results by Van Buuren et al. (2006) are very encouraging.

AUXILIARY VARIABLES

For both multiple imputation and maximum likelihood, it is often desirable to incorporate auxiliary variables into the imputation or modeling process. Auxiliary variables are those that are not intended to be in the final model. Ideally, such variables are at least moderately correlated with the variables in the model that have missing data. By including auxiliary variables into the imputation model, we can reduce the uncertainty and variability in the imputed values. This can substantially reduce the standard errors of the estimates in our final model.

Auxiliary variables can also reduce bias by getting us to a closer approximation of the MAR assumption. Here’s how. Let W be a measure of annual income and let X be a vector of observed variables that will go into the final model, along with W. Suppose that 30% of the cases are missing income, and suppose that we have reason to suspect that persons with high income are more likely to be missing income. Letting R be a response indicator for W, we can express this suspicion as

\[ \Pr(R = 1 \mid X, W) = f(X, W) \]

That is, the probability that W is missing depends on both X and W, which would be a violation of the MAR assumption. But now suppose that we have another vector of observed variables Z that, together, are highly correlated with W. These might include such things as education, IQ, sex, occupational prestige, and so on. The hope is that, if we condition on these variables, the dependence of the probability of missingness on W may disappear, so that we have

\[ \Pr(R = 1 \mid X, W, Z) = f(X, Z) \]

In practice, it's unlikely the dependence will completely disappear. But we may be able to reduce it substantially.

In sum, in order to reduce both bias and standard errors, it’s considered good practice to include auxiliary variables in multiple imputation models. The same advice applies to the maximum likelihood methods that we now consider, although the reasons may be less obvious.

MAXIMUM LIKELIHOOD

Now we’re ready to consider maximum likelihood (ML), which is a close competitor to multiple imputation. Under identical assumptions, both methods produce estimates that are consistent, asymptotically efficient and
asymptotically normal.

With or without missing data, the first step in ML estimation is to construct the likelihood function. Suppose that we have \( n \) independent observations \( (i = 1, \ldots, n) \) on \( k \) variables \( (y_{i1}, y_{i2}, \ldots, y_{ik}) \) and no missing data. The likelihood function is

\[
L = \prod_{i=1}^{n} f_i(y_{i1}, y_{i2}, \ldots, y_{ik} ; \theta)
\]

where \( f_i(.) \) is the joint probability (or probability density) function for observation \( i \), and \( \theta \) is a set of parameters to be estimated. To get the ML estimates, we find the values of \( \theta \) that make \( L \) as large as possible. Many methods can accomplish this, any one of which should produce the right result.

Now suppose that for a particular observation \( i \), the first two variables, \( y_1 \) and \( y_2 \), have missing data that satisfy the MAR assumption. (More precisely, the missing data mechanism is assumed to be ignorable). The joint probability for that observation is just the probability of observing the remaining variables, \( y_3 \) through \( y_k \). If \( y_1 \) and \( y_2 \) are discrete, this is the joint probability above summed over all possible values of the two variables with missing data:

\[
F_i^*(y_1, \ldots, y_k ; \theta) = \sum_{y_1, y_2} f_i(y_{i1}, \ldots, y_{ik} ; \theta)
\]

If the missing variables are continuous, we use integrals in place of summations:

\[
F_i^*(y_1, \ldots, y_k ; \theta) = \int_{y_1, y_2} f_i(y_{i1}, \ldots, y_{ik} ; \theta) dy_1 dy_2
\]

Essentially, then, for each observation’s contribution to the likelihood function, we sum or integrate over the variables that have missing data, obtaining the marginal probability of observing those variables that have actually been observed.

As usual, the overall likelihood is just the product of the likelihoods for all the observations. For example, if there are \( m \) observations with complete data and \( n-m \) observations with data missing on \( y_1 \) and \( y_2 \), the likelihood function for the full data set becomes

\[
L = \prod_{i=1}^{m} f_i(y_{i1}, y_{i2}, \ldots, y_{ik} ; \theta) \prod_{i=m+1}^{n} F_i^*(y_{i3}, \ldots, y_{ik} ; \theta)
\]

where the observations are ordered such that the first \( m \) have no missing data and the last \( n-m \) have missing data. This likelihood can then be maximized to get ML estimates of \( \theta \). In the remainder of the paper, we will explore several different ways to do this.

**WHY I PREFER MAXIMUM LIKELIHOOD OVER MULTIPLE IMPUTATION**

Although ML and MI have very similar statistical properties, there are several reasons why I prefer ML. Here they are, listed from least important to most important:

1. **ML is more efficient than MI.**

   Earlier I said that both ML and MI are asymptotically efficient, implying that they have minimum sampling variance. For MI, however, that statement must be qualified—MI is almost efficient. To get fully efficiency, you would have to produce and analyze an infinite number of data sets. Obviously, that’s not possible. But we can get very close to full efficiency with a relatively small number of data sets. As Rubin (1987) showed, for moderate amounts of missing data, you can get over 90% efficiency with just five data sets.

2. **For a given set of data, ML always produces the same result.** On the other hand, MI gives a different result every time you use it.

   Because MI involves random draws, there is an inherent indeterminacy in the results. Every time you apply it to a given set of data, you will get different parameter estimates, standard errors, and test statistics. This raises the possibility that different investigators, applying the same methods to the same data, could reach different conclusions. By contrast, ML always produces the same results for the same set of data.
The indeterminacy of MI estimates can be very disconcerting to those who are just starting to use it. In reality, it’s no more problematic than probability sampling. Every time we take a different random sample from the same population, we get different results. However, with random samples, we almost never get to see the results from a different random sample. With MI, it’s easy to repeat the process, and it can be alarming to see how different the results are.

What makes this problem tolerable is that we can reduce the random variation to as little as we like, just by increasing the number of data sets produced in the imputation phase. Unlike random sampling, the only cost is more computing time. So, for example, instead of using five imputed data sets, use 50. How do you know how many is enough? A couple rules of thumb have been proposed, but I won’t go into them here.

3. The implementation of MI requires many different decisions, each of which involves uncertainty. ML involves far fewer decisions.

To implement multiple imputation, you must decide:
   a. Whether to use the MCMC method or the FCS method.
   b. If you choose FCS, what models or methods to use for each variable with missing data.
   c. How many data sets to produce, and whether the number you’ve chosen is sufficient.
   d. How many iterations between data sets.
   e. What prior distributions to use.
   f. How to incorporate interactions and non-linearities.
   g. Which of three methods to use for multivariate testing.

And this list is by no means exhaustive. To be sure, most software packages for MI have defaults for things like prior distributions and numbers of iterations. But these choices are not trivial, and the informed user should think carefully about whether the defaults are appropriate.

ML is much more straightforward. For many software packages, you just specify your model of interest, tell the software to handle the missing data by maximum likelihood, and you’re done. It’s just a much “cleaner” technology than multiple imputation.

4. With MI, there is always a potential conflict between the imputation model and the analysis model. There is no potential conflict in ML because everything is done under one model.

To do MI, you must first choose an imputation model. This involves choosing the variables to include in the model, specifying the relationships among those variables, and specifying the distributions of those variables. As noted earlier, the default in PROC MI is to assume a multivariate normal model, which implies that every variable is linearly related to every other variable, and each variable is normally distributed.

Once the missing data have all been imputed, you can then use the resulting data sets to estimate whatever model you please with a different SAS procedure. For the results to be correct, however, the analysis model must, in some sense, be compatible with the imputation model. Here are two common sources of incompatibility that can cause serious bias in the estimates:

- The analysis model contains variables that were not included in the imputation model. This is never a good idea, but it can easily happen if the analyst and the imputer are different people.
- The analysis model contains interactions and non-linearities, but the imputation model is strictly linear. This is a very common problem, and it can lead one to conclude that the interactions or non-linearities are not present when they really are.

As indicated by these two examples, most such problems arise when the imputation model is more restrictive than the analysis model.

One implication is that it may be very difficult to generate an imputed data set that can be used for all the models that one may want to estimate. In fact, it may be necessary to generate different imputed data sets for each different model.

If you’re careful and thoughtful, you can avoid these kinds of problems. But it’s so easy not to be careful that I regard this as a major drawback of multiple imputation.

With ML by contrast, there’s no possibility of incompatibility between the imputation model and the analysis model. That’s because everything is done under a single model. Every variable in the analysis model will be taken into account in dealing with the missing data. If the model has nonlinearities and interactions, those will automatically be
incorporated into the method for handling the missing data.

Even though there’s no potential conflict with ML, it’s important to be aware that ML, like MI, requires a model for the relationships among the variables that have missing data. For the methods available in SAS, that’s usually a multivariate normal model, just like the default imputation method in PROC MI.

Given all the advantages of ML, why would anyone choose to do MI? The big attraction of MI is that, once you’ve generated the imputed data sets, you can use any other SAS procedure to estimate any kind of model you want. This provides enormous flexibility (although it also opens up a lot of potential for incompatibility). And you can use familiar SAS procedures rather than having to learn new ones.

ML, by contrast, typically requires specialized software. As we shall see, there are SAS procedures that can handle missing data by ML for a wide range of linear models, and a few nonlinear ones. But you won’t find anything in SAS, for example, that will do ML for logistic regression with missing data on the predictors. I’m optimistic, however, that it’s only a matter of time before that will be possible.

We’re now ready to take a look at how to use SAS procedures to do maximum likelihood for missing data under several different scenarios. In each of these scenarios, the goal is estimate some kind of regression model.

SCENARIO 1: REGRESSION MODEL WITH MISSING DATA ON THE DEPENDENT VARIABLE ONLY.

In this scenario, our goal is to estimate some kind of regression model. It doesn’t matter what kind—it could be linear regression, logistic regression, or negative binomial regression. Data are missing on the dependent variable for, say, 20% of the cases. The predictor variables have no missing data, and there is no usable auxiliary variable. That is, there is no other variable available that has a moderate to high correlation with the dependent variable, and yet is not intended to be a predictor in the regression model.

If we are willing to assume that the data are missing at random, this scenario is easy. Maximum likelihood reduces to listwise deletion (complete case analysis). That is, we simply delete cases that are missing on the dependent variable and estimate the regression with the remaining cases.

Under this scenario, there is no advantage at all in doing multiple imputation of the dependent variable. It can only make things worse by introducing additional random variation.

What if there’s a good auxiliary variable available? Well, then we can do better than listwise deletion. If the regression model is linear, one can use the methods discussed later based on PROC CALIS.

What if the data are not missing at random? That is, what if the probability that data are missing depends on the value of the dependent variable itself? Near the end of this paper, we’ll see a maximum likelihood method designed for that situation using PROC QLIM. However, I’ll also suggest some reasons to be very cautious about the use of this method.

Finally, what if some of the cases with observed values of the dependent variable have missing data on one or more predictors? For those cases, one may do better using either MI or special ML methods to handle the missing predictors. But you’re still better off just deleting the cases with missing data on the dependent variable.

SCENARIO 2. REPEATED MEASURES REGRESSION WITH MISSING DATA ON THE DEPENDENT VARIABLE ONLY

Scenario 1 was instructive but didn’t offer us any new methods. Now we come to something more interesting.

In Scenario 2, we want to estimate some kind of regression model, and we have repeated measurements on the dependent variable. Some individuals have missing data on some values of the dependent variable, and we assume that those data are missing at random. A common cause for missing data in repeated-measures situations is dropout. For example, someone may respond at time 1 but not at any later times. Someone else may respond at times 1 and 2 but not at any later times, and so on. However, we also allow for missing data that don’t follow this pattern. Thus, someone could respond at times 1 and 3, but not at time 2. There are no missing data on the predictors and no auxiliary variables.

To make this scenario more concrete, consider the following example. The data set has 595 people, each of whom was surveyed annually for 7 Years (Cornwell and Rupert 1988). The SAS data set has a total of 4165 records, one for each person in each of the seven years. Here are the variables that we shall work with:

\[
\begin{align*}
\text{LWAGE} &= \log \text{ of hourly wage (dependent variable)} \\
\text{FEM} &= 1 \text{ if female, 0 if male} \\
T &= \text{ year of the survey, 1 to 7}
\end{align*}
\]
The original data set had no missing data. For that data set, I estimated a least-squares linear regression model with LWAGE as the dependent variable. As predictors, I included FEM, T and their interaction. I used PROC SURVEYREG so that the standard errors would be adjusted for within-person correlation:

``` SAS
proc surveyreg data=my.wages;
    model lwage = fem t fem*t;
    cluster id;
run;
```

Results are shown in Output 1. There is no evidence for an interaction between FEM and T. The main effect of FEM is highly significant, however, as is the main effect of T. Women make about 37% less than men (obtained by calculating 100(exp(-.4556)-1)). For each additional year, there’s about a 10% increase in wages.

| Parameter | Estimate | Standard Error | t Value | Pr > |t| |
|-----------|----------|----------------|---------|------|------|
| Intercept | 6.339174 | 0.01569031     | 404.07  | <.0001|
| FEM       | -0.4555891 | 0.05227259    | -8.72  | <.0001|
| T         | 0.0974641 | 0.00188545     | 51.69  | <.0001|
| FEM*T     | -0.0047193 | 0.00553191   | -0.85  | 0.3940|

**Output 1. Least Squares Regression With No Missing Data.**

Next, I made some of the data missing. I constructed a drop-out mechanism such that in each year, the probability of dropping out in the following year (never to return) was a function of the wage recorded in the current year. Specifically, the probability that someone dropped out in year t+1 given that he or she responded in year t was determined by

\[ p_{t+1} = \frac{1}{1 + \exp(-8.5 - lwage_t)} \]

which says that those with higher wages were more likely to drop out. This missing data mechanism satisfies the missing at random assumption.

Here are the number of cases still present in each of the 7 years:

<table>
<thead>
<tr>
<th>Year</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>595</td>
</tr>
<tr>
<td>2</td>
<td>526</td>
</tr>
<tr>
<td>3</td>
<td>468</td>
</tr>
<tr>
<td>4</td>
<td>399</td>
</tr>
<tr>
<td>5</td>
<td>338</td>
</tr>
<tr>
<td>6</td>
<td>284</td>
</tr>
<tr>
<td>7</td>
<td>224</td>
</tr>
</tbody>
</table>

At the end of the 7 years, more than half of the cases had dropped out. The total number of records without missing data was 2834, about 68 percent of the original, balanced data set.

I then re-estimated the linear model on the data set with missing data, again using PROC SURVEYREG. Results are shown in Output 2. Now we see evidence for interaction. Specifically, the rate of increase in wages is greater for women than for men. However, for both groups, the rate of increase is lower than it was in Output 1.
Here’s the explanation for this result. People with higher wages are more likely to drop out, but those people are disproportionately men. So the fraction of men in the sample gets smaller in each year (from 89% in year 1 to 81% in year 7), and the remaining men tend to be those with relatively lower income. So, although both men and women have increasing wages, men’s wages appear to increase more slowly.

Apparently, the missing data pattern is leading us astray. How can we do better? It’s been known for some time that maximum likelihood estimation of mixed models is an effective method for dealing with missing data on repeated measures, but few researchers make appropriate use of that method (Molenberghs and Kenward 2007).

The simplest and best-known mixed model is the “random intercepts” model, given by

\[ y_{it} = \mu + \beta x_{it} + \alpha_i + \varepsilon_{it} \]

where \( y_{it} \) is the dependent variable measured for person \( i \) at time \( t \), \( x_{it} \) is a vector of predictors (some of which may vary with time), \( \alpha \) and \( \varepsilon \) are random disturbances with mean 0, constant variance, and are uncorrelated with \( x \) and with each other. What distinguishes \( \alpha \) from \( \varepsilon \) is that \( \alpha \) has only the \( i \) subscript, while \( \varepsilon \) has both \( i \) and \( t \). The variation in \( \alpha \) induces a correlation among the repeated measurements of \( y \).

By estimating this model with maximum likelihood (or residual maximum likelihood), we allow for data that are missing at random on the dependent variable. Let’s try it on our data set. The remarkable thing is that you don’t have to do anything special. Just specify the model in the usual way with PROC MIXED:

```plaintext
proc mixed data=my.wagemiss;
  model lwage = fem t fem*t / solution;
  random intercept / subject=id;
run;
```

Output 3. PROC MIXED With Missing Data.

Now our results are consistent with those in Output 1 which was based on the original sample with no missing data. The interaction is far from significant and the two main effect estimates are quite close to those in Output 1. Notice that 1331 records with missing data are deleted. Those who drop out contribute records for those years in which they respond, but not for the later years. Despite the fact that those later records are deleted, maximum likelihood estimation of mixed models is an effective method for dealing with missing data on repeated measures.
“borrows” information from the values of the dependent variable in the earlier years to project what would happen in the later years. At the same time, it fully accounts for the uncertainty of this projection in the calculation of the standard errors and test statistics.

One could produce essentially the same results by doing multiple imputation. But it’s much simpler to handle the missing data problem with PROC MIXED.

A potential disadvantage of the random intercepts model is that it implies “exchangeability” or “compound symmetry”, which means that the correlation between values of the dependent variable at any two points in time will be the same for any pair of time points (after adjusting for covariates). In actuality, most longitudinal data sets show a pattern of decreasing correlations as the time points get farther apart. Molenberghs and Kenward (2007) argue that, for missing data applications, it’s very important to get the covariance structure right. So they recommend models that impose no structure on the covariances among the repeated measures.

In PROC MIXED, an unstructured correlation matrix can be specifying by using the REPEATED statement rather than the RANDOM statement. When there are missing data, it is essential that the time variable be specified on the REPEATED statement and also on a CLASS statement. But the model we want to estimate treats time as a quantitative variable on the MODEL statement. So it is necessary to create a copy of the time variable so that it can be used as both a quantitative variable and a CLASS variable. Here’s the code:

```
data wagemiss;
  set my.wagemiss;
tcat=t;
proc mixed data=wagemiss;
  class tcat;
  model lwage = fem t fem*t / solution;
  repeated tcat / subject=id type=un;
run;
```

Results are quite similar to what we got in Output 3 with the simpler random intercepts model. The downside of the unstructured model is that it has 27 additional parameters, leading to somewhat larger standard errors for the coefficients. A good compromise might be a Toeplitz structure that allowed the correlations to decline with the time between measurements, but with far fewer parameters than the unstructured model.

```
  Effect     Estimate   Error  DF    t Value  Pr > |t|
Intercept   6.3229     0.01562 593  404.76    <.0001
FEM        -0.4500    0.04557 593    -9.88    <.0001
           0.09700    0.00244 593    39.72    <.0001
FEM*t       0.001247  0.006194 593     0.20     0.8405
```

Output 4. PROC MIXED With Unstructured Covariance Matrix.

LINEAR MIXED MODELS WITH AUXILIARY VARIABLES

It’s possible to include auxiliary variables in a repeated-measures, mixed model analysis with missing data, but it takes a bit of effort. Here’s an outline of what’s required for a single auxiliary variable:

1. Treat the auxiliary variable as an additional value of the dependent variable. This will require that there be an additional record for each individual, with the value of the auxiliary variable as the dependent variable.
2. Create an indicator variable D with a value of 1 if the record is for the auxiliary variable, otherwise 0.
3. Using the REPEATED statement, fit a model with an unstructured covariance matrix.
4. On the MODEL statement, include interactions between the indicator variable D and each of the variables on the right-hand side, as well as a main effect of D itself.
In the output, the “main effects” of the variables will be the parameters of interest. The interactions are included simply to allow the independent variables to have different effects on the auxiliary variable.

**MIXED MODELS WITH BINARY OUTCOMES**

If the repeated-measures dependent variable is binary, you can get the same benefits with PROC GLIMMIX. Here’s an example. (The data used here come from clinical trials conducted as part of the National Drug Abuse Treatment Clinical Trials Network sponsored by National Institute on Drug Abuse). The sample consisted of 154 opioid-addicted youths, half of whom were randomly assigned to a treatment consisting of the administration of buprenorphine-naloxone over a 12-week period. The other half received a standard short-term detox therapy. The primary outcome of interest is a binary variable coded 1 if the subject tested positive for opiates. These urine tests were intended to be performed in each of the 12 weeks following randomization. However, for various reasons, there was a great deal of missing data on these outcomes. Twenty persons had missing data for all 12 outcomes, reducing the effective sample size to 134. After eliminating these 20 cases, Table 1 shows the proportion of cases with data present in each of the 12 weeks.

<table>
<thead>
<tr>
<th>Week</th>
<th>Proportion Present</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.90</td>
</tr>
<tr>
<td>2</td>
<td>.74</td>
</tr>
<tr>
<td>3</td>
<td>.60</td>
</tr>
<tr>
<td>4</td>
<td>.78</td>
</tr>
<tr>
<td>5</td>
<td>.48</td>
</tr>
<tr>
<td>6</td>
<td>.45</td>
</tr>
<tr>
<td>7</td>
<td>.44</td>
</tr>
<tr>
<td>8</td>
<td>.69</td>
</tr>
<tr>
<td>9</td>
<td>.40</td>
</tr>
<tr>
<td>10</td>
<td>.37</td>
</tr>
<tr>
<td>11</td>
<td>.37</td>
</tr>
<tr>
<td>12</td>
<td>.67</td>
</tr>
</tbody>
</table>

*Table 1. Proportion of cases with drug test data present in each of 12 weeks.*

The proportions were substantially higher in weeks 1, 4, 8, and 12, presumably because the subjects were paid $75 for participation in those weeks, but only $5 for participation in the other weeks.

The objective is to estimate the effect of the treatment on the probability of a positive drug test, as well as evaluating how that effect may have changed over the 12-week period. We shall estimate a random effects (mixed) logit model using PROC GLIMMIX.

Like PROC MIXED, GLIMMIX expects data in the “long form”, that is, one record for each longitudinal measurement for each individual. Thus, a person who was tested on all 12 occasions would have 12 records in the data set. Each record has an ID variable that has a unique value for each person and serves to link the separate records for each person. There is also a WEEK variable that records the week of the measurement (i.e., 1 through 12). Any variables that do not change over time (e.g., the treatment indicator) are simply replicated across the multiple observations for each person. If a person has a missing outcome at a particular time point, no record is necessary for that time point. Thus, a person who only was tested at weeks 1, 4, 8, and 12, has only four records in the data set.

After some exploration, the best fitting model was a “random intercepts” logit model of the following form:
\[
\text{pit}^2 = \log\left(\frac{1}{1-p}^2\right) = \mu + \beta z_i + \gamma t_i + \alpha + \delta t_i + \phi t_i z_i + \alpha_i
\]

Here \(p_{it}\) is the probability that person \(i\) tests positively for opiates at week \(t\). \(z_i\) is coded 1 if the person \(i\) is in the treatment group, otherwise 0. The model includes an interaction between treatment and week, a quadratic effect of week, and an interaction between treatment and week squared. \(\alpha_i\) represents all the causes of \(y_i\) that vary across persons but not over time. It is assumed to be a random variable with the following properties:

- \(E(\alpha_i) = 0\).
- \(\text{Var}(\alpha_i) = \tau^2\).
- \(\alpha_i\) independent of \(x_i\) and \(z_i\).
- \(\alpha_i\) normally distributed.

Here is the SAS code for estimating this model:

```sas
proc glimmix data=my.nidalong method=quad(qpoints=5);
  class usubjid;
  where week ne 0;
  model opiates=treat week week*treat week*week week*week*treat/ d=b solution ;
  random intercept / subject=usubjid;
  output out=a pred(ilink noblup)=yhat2 ;
run;
```

The METHOD option specifies that the likelihood function will be evaluated numerically using Gaussian quadrature. I also specified 5 quadrature points because default number failed. In general, the more quadrature points the greater the accuracy. If the METHOD option is omitted, GLIMMIX does pseudo-likelihood estimation, which is known to be inaccurate for dichotomous outcomes.

The WHERE statement excludes the baseline drug test from the analysis. In the MODEL statement, D=B specifies a binomial distribution for the dependent variable with a logit link (the default). The SOLUTION option requests parameter estimates—otherwise only test statistics would be reported. The RANDOM statement specifies the inclusion of the \(\alpha_i\) term in the model. The OUTPUT statement produces a data set that will be used to construct a graph of the treatment and time effects.

Note that even though person-weeks with missing data on the outcome are excluded, this analysis should produce approximately unbiased estimates under the MAR assumption.

Estimates and test statistics for this model, seen in Output 5, show highly significant interactions of treatment with both the linear and quadratic components of time. These effects are hard to interpret from the numerical output, but we'll see a graph shortly.

| Effect                | Estimate | Standard Error | DF   | t Value | Pr > |t| |
|-----------------------|----------|----------------|------|---------|------|---|
| Intercept             | -0.5545  | 0.4928         | 132  | -1.13   | 0.2625 |
| treat                 | 1.1327   | 0.6766         | 786  | 1.67    | 0.0945 |
| week                  | 0.4198   | 0.1714         | 786  | 2.45    | 0.0145 |
| treat*week            | -1.2558  | 0.2336         | 786  | -5.38   | <.0001 |
| week*week             | -0.02804 | 0.01297        | 786  | -2.16   | 0.0309 |
| treat*week*week       | 0.09299  | 0.01776        | 786  | 5.24    | <.0001 |

Output 5. Estimates for Mixed Model with Quadratic Effect of Time and Interactions with Treatment.

I also estimated a comparable model using PROC SURVEYLOGISTIC with robust standard errors. Although SURVEYLOGISTIC does ML estimation, the estimated model does not allow for any over-time correlations. Therefore, it should not be expected to produce estimates that appropriately adjust for the missing data. As in
Output 4, the results showed strong interactions between the treatment and both the linear and quadratic components of WEEK. But the magnitudes of the coefficients were somewhat different, leading to noticeable differences in the graphs of the predicted values.

Figure 1 displays predicted probabilities of a positive drug test for treatment and control groups for the two estimation methods. The green and brown curves are based on the MIXED model; the blue and red curves (the two curves in the middle) are from the SURVEYLOGISTIC model. For both methods, the graph shows that the treatment differences are minimal in the early weeks, increase dramatically to about week 7, and then gradually diminish toward week 12. But we also see that the treatment effect is larger for the MIXED model for all weeks up to week 10.

As noted earlier, random intercept models have the undesirable feature of implying that the association between measurements that are close in time are no larger than measurements that are farther apart. With PROC MIXED, we were able to relax that assumption by using the REPEATED statement. PROC GLIMMIX does not have a REPEATED statement, but it does have an option on the RANDOM statement that allows one to specify an unstructured model. Unfortunately, this option does not work if you request maximum likelihood rather than the default pseudo-likelihood method. (Even if the option would work with ML, the resulting coefficients would be population-averaged rather than subject-specific.)

**SCENARIO 3. LINEAR MODELS WITH DATA MISSING ON PREDICTOR VARIABLES**

When it comes to missing data, the big limitation of PROC MIXED and PROC GLIMMIX is that they do nothing about missing data on predictor variables. In particular, cases with valid observations on the dependent variable but missing values on one or predictor variables contain potentially valuable information that will be completely lost. So we now turn other methods that can effectively utilize the available information for those cases. As before, we will assume that any missing data are missing at random. And we will focus exclusively on linear models.

For this situation, there are two options in SAS: the EM algorithm in PROC MI and the “full-information” maximum likelihood method in PROC CALIS.
EM ALGORITHM IN PROC MI

Although PROC MI was designed primarily to do multiple imputation, the first step in the default algorithm is to do maximum likelihood estimation of the means and the covariance matrix using the EM algorithm. These estimates are used as starting values for the MCMC multiple imputation algorithm. But they can also be useful in their own right.

Consider the following example. The data set NLSYMISS (available at www.StatisticalHorizons.com), has records for 581 children who were surveyed in 1990 as part of the National Longitudinal Survey of Youth. Here are the variables:

- ANTI: antisocial behavior, measured with a scale ranging from 0 to 6.
- SELF: self-esteem, measured with a scale ranging from 6 to 24.
- POV: poverty status of family, coded 1 for in poverty, otherwise 0.
- BLACK: 1 if child is black, otherwise 0.
- HISPANIC: 1 if child is Hispanic, otherwise 0.
- CHILDA GE: child’s age in 1990.
- DIVORCE: 1 if mother was divorced in 1990, otherwise 0.
- GENDER: 1 if female, 0 if male.
- MOMAGE: mother’s age at birth of child.
- MOMWORK: 1 if mother was employed in 1990, otherwise 0.

Our goal is to estimate the linear regression of ANTI on the other variables. Four of the variables have missing data: SELF, POV, BLACK, HISPANIC, and MOMWORK. If you use PROC REG to run the regression, you get the results shown in Output 6. Only one of the variables, POV, is statistically significant at the .05 level. But PROC REG deletes 356 of the 581 cases because of missing data on one or more variables.

| Variable   | DF | Parameter Estimate | Standard Error | t Value | Pr > |t| |
|------------|----|--------------------|----------------|---------|------|---|
| Intercept  | 1  | 2.86533            | 1.99117        | 1.44    | 0.1516 |
| self       | 1  | -0.04531           | 0.03135        | -1.45   | 0.1498 |
| pov        | 1  | 0.71946            | 0.23739        | 3.03    | 0.0027 |
| black      | 1  | 0.05069            | 0.24918        | 0.20    | 0.8390 |
| hispanic   | 1  | -0.35696           | 0.25537        | -1.40   | 0.1636 |
| childata   | 1  | 0.00197            | 0.17072        | 0.01    | 0.9908 |
| divorce    | 1  | 0.08703            | 0.24499        | 0.36    | 0.7228 |
| gender     | 1  | -0.33470           | 0.19844        | -1.69   | 0.0931 |
| momage     | 1  | -0.01198           | 0.04611        | -0.26   | 0.7953 |
| momwork    | 1  | 0.25440            | 0.21751        | 1.17    | 0.2435 |

Output 6. PROC REG Output for ANTI as Dependent Variable, Complete Case Analysis.

We can do better by using PROC MI to produce ML estimates of the means and covariance matrix for all the variables in the model. This is accomplished by:

```latex
proc mi data=nlsymiss nimpute=0;
    var anti self pov black hispanic childata divorce gender momage momwork;
    em outem=nlsyem;
run;
```
The NIMPUTE=0 option suppresses multiple imputation. The EM statement requests EM estimates and writes the means and covariance matrix into a SAS data set called NLSYEM;

EM stands for expectation-maximization (Dempster et al. 1977). EM is simply a convenient numerical algorithm for getting ML estimates in certain situations under specified assumptions. In this case, the assumptions are that the data are missing at random, and all the variables have a multivariate normal distribution. That implies that every variable is a linear function of all the other variables (or any subset of them) with homoskedastic errors. I’ll have more to say about this assumption later.

What can we do with the EM covariance matrix? Well, for many linear models including linear regression, the parameters are functions of the means, variances and covariances. So we can use the output data set from PROC MI as input to PROC REG:

```
proc reg data=nlsyem;
model anti= self pov black hispanic childage divorce gender momage momwork;
run;
```

Results in Output 6 are very different than those in Output 5. Most notably, the p-values are much lower than before, and all the coefficients but one are statistically significant at the .01 level or better.

| Variable   | DF | Parameter Estimate | Standard Error | t Value | Pr > |t|   |
|------------|----|--------------------|----------------|---------|------|-----|
| Intercept  | 1  | 2.59110            | 0.28964        | 8.95    | <.0001         |
| self       | 1  | -0.06718           | 0.00454        | -14.79  | <.0001         |
| pov        | 1  | 0.64626            | 0.03357        | 19.25   | <.0001         |
| black      | 1  | 0.08501            | 0.03395        | 2.50    | 0.0123         |
| hispanic   | 1  | -0.32439           | 0.03636        | -8.92   | <.0001         |
| childage   | 1  | -0.00386           | 0.02391        | -0.16   | 0.8716         |
| divorce    | 1  | -0.10598           | 0.03422        | -3.10   | 0.0020         |
| gender     | 1  | -0.56116           | 0.02783        | -20.16  | <.0001         |
| momage     | 1  | 0.02076            | 0.00668        | 3.11    | 0.0019         |
| momwork    | 1  | 0.21896            | 0.03074        | 7.12    | <.0001         |


The good news is that the parameter estimates in Output 6 are the true maximum likelihood estimates of the regression coefficients. So in that sense, we’ve accomplished our goal. The bad news is that the p-values are useless. We know that because of a warning message in the log window:

```
WARNING: The data set WORK.NLSYEM does not indicate how many observations were used to compute the COV matrix. The number of observations has been set to 10000. Statistics that depend on the number of observations (such as p-values) are not interpretable.
```

There’s a way to edit the output data set from PROC MI to specify a sample size. But what size? It would be wrong to choose 581, the original sample size, because that would presume that there are no missing data. Consequently, the standard errors and p-values would be too small. In fact, there is no single sample size that will give the correct standard errors for all the coefficients.

One solution to this problem is to estimate the standard errors and p-values by bootstrapping. Here are the steps for doing that:

1. From the original sample of size N, draw many samples of size N with replacement.
2. Produce the EM estimates of the means/covariance matrix for each sample.
3. Estimate the regression model from each covariance matrix.
4. Calculate the standard deviation of each regression coefficient across samples.

And here's the SAS code for accomplishing that with 1000 bootstrap samples:

```sas
proc surveyselect data=nlsymiss method=urs n=581 reps=1000 out=bootsamp outhits;
proc mi data=bootsamp nimpute=0 noprint;
    var anti self pov black hispanic childage divorce gender momage momwork;
    em outem=nlsyem;
    by replicate;
proc reg data=nlsyem outest=a noprint;
    model anti= self pov black hispanic childage divorce gender momage momwork;
    by replicate;
proc means data=a std;
    var self pov black hispanic childage divorce gender momage momwork;
run;
```

Table 2 shows the ML coefficients (from Output 6), the bootstrapped standard errors, the z-statistics (ratios of coefficients to standard errors) and the associated p-values. Comparing this with Output 6—the complete case analysis—we see some major differences. SELF and GENDER are now highly significant, and HISPANIC and MOMWORK are now marginally significant. So ML really made a difference for this example.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std_Err</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>self</td>
<td>-.06718</td>
<td>.022402</td>
<td>-2.99888</td>
</tr>
<tr>
<td>pov</td>
<td>.64626</td>
<td>.166212</td>
<td>3.888174</td>
</tr>
<tr>
<td>black</td>
<td>.08501</td>
<td>.168117</td>
<td>.5056583</td>
</tr>
<tr>
<td>hispanic</td>
<td>-.32439</td>
<td>.163132</td>
<td>-1.988516</td>
</tr>
<tr>
<td>childage</td>
<td>-.00386</td>
<td>.103357</td>
<td>-.0373463</td>
</tr>
<tr>
<td>divorce</td>
<td>-.10598</td>
<td>.150277</td>
<td>-.705231</td>
</tr>
<tr>
<td>gender</td>
<td>-.56116</td>
<td>.114911</td>
<td>-4.883444</td>
</tr>
<tr>
<td>momage</td>
<td>.02076</td>
<td>.028014</td>
<td>.7410475</td>
</tr>
<tr>
<td>momwork</td>
<td>.21896</td>
<td>.145777</td>
<td>1.502024</td>
</tr>
</tbody>
</table>

Table 2. ML Coefficients With Bootstrap Standard Errors.

This example had no auxiliary variables, but it would be easy to incorporate them into the analysis. Just add them to the VAR statement in PROC MI.

The primary weakness of this methodology is the assumption of multivariate normality for all the variables. That assumption can’t possibly be true for binary variables like POV, BLACK, HISPANIC and MOMWORK, all variables with missing data. However, that assumption is also the basis for the default MCMC method in PROC MI. So it’s no worse than multiple imputation in that regard. A considerable amount of simulation evidence suggests that violation of the multivariate normality assumption are not very problematic for multiple imputation (e.g., Schafer 1997, Allison 2006). Given the close similarity between multiple imputation and maximum likelihood under identical assumptions, I would expect ML to be robust to such violations also. However, toward the end of this paper, I will briefly discuss alternative ML methods with less restrictive assumptions.

**FIML IN PROC CALIS**

There’s an easier way to get the ML estimates in SAS, one that only requires a single procedure and does not require any additional computation to get the standard errors. Beginning with release 9.22, PROC CALIS can do what SAS calls “full information maximum likelihood” (FIML), which is just maximum likelihood estimation with appropriate incorporation of the cases with missing data. In other literature, the method is sometimes described as “direct” maximum likelihood (because it directly maximizes the likelihood for the specified model rather than doing the two
steps that we used with PROC MI) or “raw” maximum likelihood (because the model must be estimated using raw
data rather than a covariance matrix).

First a little background. PROC CALIS is designed to estimate linear “structural equation models” (SEM), a very large
class of linear models that involves multiple equations and often latent variables. It encompasses both the
confirmatory factor model of psychometrics and the simultaneous equation model of econometrics. The SEM method
was first introduced in the 1970s by Karl Jöreskog, who also developed the first stand-alone software package (called
LISREL) to implement it.

Like most SEM software, the default estimation method in PROC CALIS is maximum likelihood under the assumption
of multivariate normality. But the default is also to delete cases with missing data on any of the variables in the
specified model. By specifying METHOD=FIML on the PROC statement, we can retain those cases by integrating the
likelihood function over the variables with missing data. This results in a somewhat more complicated likelihood
function, one that takes more computation to maximize.

Let’s apply this method to the NLSY data that we just analyzed with PROC MI and PROC REG. Again, our goal is to
estimate a linear regression model, which is just a special case of the general structural equation model. PROC
CALIS is unusual in that it has eight different languages that can be used to specify equivalent models. Why so
many? Well, different disciplines have developed different terminologies and ways of formalizing these models, and
the SAS developers wanted to create a software environment where everyone could feel at home. Until recently, I
used only the EQS language. But I’ve come to realize that the PATH language is a lot easier, at least for relatively
simple models. And the model that we want to estimate is about as simple as you can get.

Here’s the CALIS code:

```
proc calis data=my.nlsymiss method=fiml;
    path anti <- self pov black hispanic childage divorce gender momage momwork;
run;
```

The PATH statement specifies a linear model with ANTI as the dependent variable and the others as predictors.
Results are shown in Output 8. Actually, this is only a small part of the standard output. SEM software tends to
produce a lot of output, and CALIS is no exception.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Error</th>
<th>t Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>anti self</td>
<td>-0.06718</td>
<td>0.02193</td>
<td>-3.06412</td>
</tr>
<tr>
<td>anti pov</td>
<td>0.64627</td>
<td>0.16366</td>
<td>3.94874</td>
</tr>
<tr>
<td>anti black</td>
<td>0.08500</td>
<td>0.16081</td>
<td>0.52857</td>
</tr>
<tr>
<td>anti hispanic</td>
<td>-0.32439</td>
<td>0.17011</td>
<td>-1.90694</td>
</tr>
<tr>
<td>anti childage</td>
<td>-0.00387</td>
<td>0.10055</td>
<td>-0.03844</td>
</tr>
<tr>
<td>anti divorce</td>
<td>-0.10599</td>
<td>0.14583</td>
<td>-0.72680</td>
</tr>
<tr>
<td>anti gender</td>
<td>-0.56116</td>
<td>0.11691</td>
<td>-4.79985</td>
</tr>
<tr>
<td>anti momage</td>
<td>0.02076</td>
<td>0.02817</td>
<td>0.73702</td>
</tr>
<tr>
<td>anti momwork</td>
<td>0.21895</td>
<td>0.14169</td>
<td>1.54528</td>
</tr>
</tbody>
</table>

Output 8. PROC CALIS Output With FIML Estimates of Linear Regression Model.

Comparing Output 8 with Table 2, we see that the coefficients are identical (at least to the first four decimal places),
which is exactly as it should be. Maximum likelihood is maximum likelihood, regardless of how you get there. The
standard errors shown in Output 8, were produced by the “observed information matrix” method that is standard for
most applications of maximum likelihood. While not identical to the bootstrapped standard errors in Table 2 (we
wouldn’t expect them to be), they’re pretty close. Consequently, the t and z statistics are about the same as well.

So we’ve gotten to the same place with PROC MI and PROC CALIS. But CALIS only required three lines of code,
while MI required 16 lines to get both the coefficients and the standard errors. Furthermore, MI with bootstrapped
standard errors used about 8 times as much computing time as PROC CALIS.

What assumptions were made here? The same as for PROC MI: missing at random and multivariate normality.
How can we incorporate auxiliary variables into the analysis? With PROC MI it was easy—just put the auxiliary variables on the VAR statement along with all the others. It's a little more complicated with CALIS. Suppose we decide to remove SELF as a predictor in the linear regression model and treat it as an auxiliary variable instead. To do that properly, we must allow SELF to be freely correlated with all the variables in the regression model. An easy way to do that is to specify a second regression model with SELF as the dependent variable. Here's how:

```
proc calis data=my.nlsymiss method=fiml;
  path anti <- pov black hispanic childage divorce gender momage momwork,
              self <- anti pov black hispanic childage divorce gender momage momwork;
run;
```

Notice that the two “paths” or equations are separated by a comma, not a semicolon. Now suppose that we have two auxiliary variables rather than just one. Specifically, let's remove MOMAGE from the main equation and treat it as an auxiliary variable:

```
proc calis data=my.nlsymiss method=fiml;
  path anti <- pov black hispanic childage divorce gender momwork,
              momage self <- anti pov black hispanic childage divorce gender momwork,
              momage <-> self;
run;
```

The second line in the PATH statement has both MOMAGE and SELF on the left-hand side. This specifies two linear equations, one with MOMAGE as the dependent variable and the other with SELF as the dependent variable. But that's not enough. We must also allow for a correlation between MOMAGE and SELF (or rather between their residuals). That's accomplished with the third line in the PATH statement.

So that's essentially how you use PROC CALIS to do ML estimation of a linear regression model when data are missing. But CALIS can do so much more than single equation linear regression. It can estimate confirmatory factor models, simultaneous equation models with feedback loops, and structural equation models with latent variables. For examples, see the paper by Yung and Zhang (2011) presented at last year's Global Forum. As shown in my book *Fixed Effects Regression Methods for Longitudinal Data Using SAS* (2005), CALIS can also be used to estimate fixed and random effects models for longitudinal data. For any of these models, CALIS can handle missing data by maximum likelihood—as always, under the assumption of multivariate normality and missing at random.

**SCENARIO 4. GENERALIZED LINEAR MODELS WITH DATA MISSING ON PREDICTORS**

As we just saw, PROC CALIS can handle general patterns of missing data for almost any linear model that you might want to estimate. But suppose the dependent variable is dichotomous, and you want to estimate a logistic regression. Or maybe your dependent variable is a count of something, and you think that a negative binomial regression model would be more appropriate. If data are missing on predictors, there's no SAS that can handle the missing data by maximum likelihood—as always, under the assumption of multivariate normality and missing at random.

I know of only one commercial package that can do ML with missing predictors for generalized linear models, and that's Mplus. Since this is a SAS forum, I won't go into any of the details. But I do want to stress that Mplus can handle general patterns of missing data for an amazing variety of regression models: binary logistic, ordered logistic, multinomial logistic, poisson, negative binomial, tobit, and Cox regression. What's more, with Mplus you can specify models that do not assume that the missing predictors have a multivariate normal distribution. Instead, you can model them as dependent variables using the same wide variety of regressions: logistic, poisson, etc. And you can do this for both cross-sectional and longitudinal data.

Hopefully, there will someday be a SAS procedure that can do this.

**SCENARIO 5. DATA MISSING ON THE DEPENDENT VARIABLE, BUT NOT AT RANDOM**

All the methods we have considered so far have been based on the assumption that the data are missing at random. That's a strong assumption, but one that is not easy to relax. If you want to go the not-missing-at-random route, you must specify a model for the probability that data are missing as a function of both observed and unobserved variables. Such models are often not identified or may be only weakly identified.

Nevertheless, there is one such model that is widely used and can be estimated with PROC QLIM (part of the ETS product). It's the "sample selection bias" model of James Heckman, who won the Nobel Prize in economics for this and other contributions.
Heckman’s (1976) model is not typically thought of as a missing data method, but that’s exactly what it is. The model is designed for situations in which the dependent variable in a linear regression model is missing for some cases but not for others. A common motivating example is a regression predicting women’s wages, where wage data are necessarily missing for women who are not in the labor force. It is natural to suppose that women are less likely to enter the labor force if their wages would be low upon entry. If that’s the case, the data are not missing at random.

Heckman formulated his model in terms of latent variables, but I will work with a more direct specification. For a sample of \( n \) cases (\( i=1,\ldots, n \)), let \( Y_i \) be a normally distributed variable with a variance \( \sigma^2 \) and a conditional mean given by

\[
E(Y_i|X_i) = \beta'X_i
\]

where \( X_i \) is a column vector of independent variables (including a value of 1 for the intercept) and \( \beta \) is a row vector of coefficients. The goal is to estimate \( \beta \). If all \( Y_i \) were observed, we could get ML estimates of \( \beta \) by ordinary least squares regression. But some \( Y_i \) are missing. Let \( R \) be an indicator variable having a value of 1 if \( Y \) is observed and 0 if \( Y \) is missing. The probability of missing data on \( Y \) is assumed to follow a probit model

\[
Pr(R_i = 0 | Y_i, X_i) = \Phi(\alpha_0 + \alpha_1Y_i + \alpha_2X_i)
\]

where \( \Phi(.) \) is the cumulative distribution function for a standard normal variable. If \( \alpha_1=0 \), the data are missing at random. Otherwise the data are not missing at random because the probability of missing \( Y \) depends on \( Y \) itself. If both \( \alpha_1=0 \) and \( \alpha_2=0 \), the data are missing completely at random.

This model is identified (even when there are no \( X_i \) or when \( X_i \) does not enter the probit equation) and can be estimated by maximum likelihood.

Here’s how to do it with QLIM. The data from Mroz (1987) contain information on logged women’s wages (LWAGE) and labor force participation (INLF). Predictors include household income not from wife (NWIFEINC), years of schooling (EDUC), years in the labor force (EXPER), AGE, number of children less than 6 years old (KIDSLT6) and number of children greater than or equal to 6 (KIDSGE6).

Here’s the program:

```plaintext
proc qlim data=my.mroz;
   model inlf = nwifeinc educ exper expersq age kidslt6 kidsge6 / discrete;
   model lwage = educ exper expersq / select(inlf=1);
run;
```

The first MODEL statement specifies the probit model for INLF. The second MODEL statement specifies the linear regression model for LWAGE. The SELECT option specifies that data are present on LWAGE only when INLF=1. Notice that the INLF equation contains predictors that are not included in the LWAGE equation. This is desirable for achieving robust identification of the model, but it is not essential. Results are shown in Output 8.

<table>
<thead>
<tr>
<th>Index</th>
<th>Value</th>
<th>Total Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>325</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>426</td>
</tr>
</tbody>
</table>

| Parameter Estimate | DF | Estimate | Standard Error | t Value | Pr > |t| |
|--------------------|----|----------|----------------|---------|------|---|
| LWAGE.Intercept    | 1  | -0.534470| 0.262809       | -2.03   | 0.0420 |
| LWAGE.EDUC         | 1  | 0.107985 | 0.014908       | 7.24    | <.0001|
| LWAGE.EXPER        | 1  | 0.041617 | 0.014974       | 2.78    | 0.0054|
Output 9. PROC QLIM Output for Heckman model.

In the output, we see estimates of the regression coefficients in each of the two equations. The estimates in the LWAGE equation correct for bias that would otherwise occur because the data are not missing at random. The parameter RHO is a function of the $\alpha_1$ parameter in the equation above. Specifically, if RHO is 0, then $\alpha_1$ is zero, implying that the probability that $Y$ is missing does not depend on $Y$ itself. In this example, we see no evidence that RHO is anything other than 0 ($p=.91$).

Although this seems like an attractive method, it has a fundamental flaw. The estimates produced by this method are extremely sensitive to the assumption that $Y$ has a normal distribution. If $Y$ actually has a skewed distribution, ML estimates obtained under Heckman’s model may be severely biased, perhaps even more than estimates obtained under an MAR model (Little and Rubin 1987). This is fairly typical of models for data that are not missing at random.

CONCLUSION

Maximum likelihood is a great way to handle missing data. It has optimal statistical properties (if assumptions are met), and it has several advantages over multiple imputation. The most important advantage is that there is no potential conflict between an imputation model and an analysis model. When estimating mixed models for repeated measurements, PROC MIXED and PROC GLIMMIX automatically handle missing data by maximum likelihood, as long as there are no missing data on predictor variables. When data are missing on both predictor and dependent variables, PROC CALIS can do maximum likelihood for a large class of linear models with minimal programming effort. All three of these procedures assume that data are missing at random. PROC QLIM can estimate models for data that are not missing at random on the dependent variable. But results may be very sensitive to distributional assumptions about the response variable.

REFERENCES


**APPENDIX 5: Adverse events in ASIST**

There were more adverse events reported in PBA group than SBA group, but the difference was not statistically significant (p=0.059) and most of the AE was not related to study participation.

**PBA (n=27)**
3 participants were withdrawn per protocol as they had too many exacerbations (two in 6 months, or three in 12 months).

**Adverse events**
- 3 Influenza infection
- 3 Strep pharyngitis
- 2 Arm fracture
- 2 Upper respiratory infection
- 2 Ear infection
- 1 ADHD
- 1 Stomach Flu
- 1 Shaking after albuterol
- 1 Rash
- 1 allergy reaction
- 1 Tonsillectomy outpatient
- 1 transient sleepiness
- 1 chest wall pain
- 1 Tic disorder
- 1 Finger injury
- 1 Conjunctivitis
- 1 Thrush

**Serious adverse events**
- 1 asthma exacerbation requiring hospitalization
- 1 pneumonia requiring hospitalization
- 1 hospitalization due to chronic constipation

**SBA (n=16)**
2 participants were withdrawn per protocol as they had too many exacerbations (two in 6 months, or three in 12 months).

**Adverse events**
- 2 Upper respiratory infection
- 2 Ear infection
- 2 Influenza infection
- 2 strep throat
- 1 bronchitis
- 1 increased appetite
- 1 finger fracture
- 1 viral syndrome
- 1 skin infection
- 1 tooth abscess
- 1 hairloss

**Serious adverse events**
One asthma exacerbation requiring hospitalization
### Descriptive statistics of demographic characteristics for individuals with and without missing follow-up data

**07MAR2019**

<table>
<thead>
<tr>
<th>Variable Label</th>
<th>Variable Value</th>
<th>Baseline data only (missing follow-up data)</th>
<th>Follow-up data available at 1 or more time points</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment group</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PBA</td>
<td>7/14 (50%)</td>
<td>96/192 (50%)</td>
<td>&gt; 0.99</td>
</tr>
<tr>
<td></td>
<td>SBA</td>
<td>7/14 (50%)</td>
<td>96/192 (50%)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td>11.47 +/- 3.23 [14]</td>
<td>10.77 +/- 3.34 [192]</td>
<td>0.44</td>
</tr>
<tr>
<td>Age at Diagnosis</td>
<td></td>
<td>2.45 +/- 2.31 [14]</td>
<td>2.84 +/- 2.88 [192]</td>
<td>0.62</td>
</tr>
<tr>
<td>Participant gender</td>
<td></td>
<td></td>
<td></td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>6/14 (43%)</td>
<td>89/192 (46%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>8/14 (57%)</td>
<td>103/192 (54%)</td>
<td></td>
</tr>
<tr>
<td>Season of Randomization</td>
<td></td>
<td></td>
<td></td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>Spring</td>
<td>1/14 (7.1%)</td>
<td>49/192 (26%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Summer</td>
<td>5/14 (36%)</td>
<td>76/192 (40%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fall</td>
<td>4/14 (29%)</td>
<td>37/192 (19%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Winter</td>
<td>4/14 (29%)</td>
<td>30/192 (16%)</td>
<td></td>
</tr>
<tr>
<td>In the past 12 months, has the participant received steroids (e.g., prednisone, IM, IV) for their asthma?</td>
<td>No</td>
<td>8/14 (57%)</td>
<td>111/192 (58%)</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>6/14 (43%)</td>
<td>81/192 (42%)</td>
<td></td>
</tr>
<tr>
<td>In the past 12 months, has the participant visited a hospital emergency room or urgent care because of breathing problems?</td>
<td>No</td>
<td>8/14 (57%)</td>
<td>140/192 (73%)</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>6/14 (43%)</td>
<td>52/192 (27%)</td>
<td></td>
</tr>
<tr>
<td>Inhaled corticosteroid alone</td>
<td></td>
<td></td>
<td></td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0/14 (0%)</td>
<td>29/192 (15%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>14/14 (100%)</td>
<td>163/192 (85%)</td>
<td></td>
</tr>
<tr>
<td>Low dose ICS/LABA</td>
<td></td>
<td></td>
<td></td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>12/14 (86%)</td>
<td>174/191 (91%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>2/14 (14%)</td>
<td>17/191 (8.9%)</td>
<td></td>
</tr>
</tbody>
</table>

Appendix 6: Participant With Missing Data
Descriptive statistics of demographic characteristics for individuals with any missing data and those without any missing data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group</th>
<th>Any missing data</th>
<th>No missing data</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment group</td>
<td>PBA</td>
<td>20/36 (56%)</td>
<td>83/170 (49%)</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>SBA</td>
<td>16/36 (44%)</td>
<td>87/170 (51%)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td>10.2 +/- 3.29 [36]</td>
<td>10.94 +/- 3.33 [170]</td>
<td>0.23</td>
</tr>
<tr>
<td>Age at Diagnosis</td>
<td></td>
<td>2.68 +/- 2.37 [36]</td>
<td>2.85 +/- 2.94 [170]</td>
<td>0.75</td>
</tr>
<tr>
<td>Participant gender</td>
<td>Female</td>
<td>19/36 (53%)</td>
<td>76/170 (45%)</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>17/36 (47%)</td>
<td>94/170 (55%)</td>
<td></td>
</tr>
<tr>
<td>Season of Randomization</td>
<td>Spring</td>
<td>7/36 (19%)</td>
<td>43/170 (25%)</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>Summer</td>
<td>13/36 (36%)</td>
<td>68/170 (40%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fall</td>
<td>10/36 (28%)</td>
<td>31/170 (18%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Winter</td>
<td>6/36 (17%)</td>
<td>28/170 (16%)</td>
<td></td>
</tr>
<tr>
<td>In the past 12 months, has the participant received steroids (e.g., prednisone, IM, IV) for their asthma?</td>
<td>No</td>
<td>19/36 (53%)</td>
<td>100/170 (59%)</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>17/36 (47%)</td>
<td>70/170 (41%)</td>
<td></td>
</tr>
<tr>
<td>In the past 12 months, has the participant visited a hospital emergency room or urgent care because of breathing problems?</td>
<td>No</td>
<td>25/36 (69%)</td>
<td>123/170 (72%)</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>11/36 (31%)</td>
<td>47/170 (28%)</td>
<td></td>
</tr>
<tr>
<td>Inhaled corticosteroid alone</td>
<td>No</td>
<td>3/36 (8.3%)</td>
<td>26/170 (15%)</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>33/36 (92%)</td>
<td>144/170 (85%)</td>
<td></td>
</tr>
<tr>
<td>Low dose ICS/LABA</td>
<td>No</td>
<td>31/36 (86%)</td>
<td>155/169 (92%)</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>5/36 (14%)</td>
<td>14/169 (8.3%)</td>
<td></td>
</tr>
</tbody>
</table>

P-values obtained using t-tests for continuous outcomes, and Chi-square tests or Fisher’s exact tests were used for categorical outcomes as appropriate.

For continuous outcomes, sample size is reported in brackets.