# Evaluating socio-demographic risk factors of asthma control and effects of using a pilot asthma self-management mobile platform

Ruixuan Wang
*Institute of Health Informatics*
*University College London*
London, UK
ruixuan.wang.23@alumni.ucl.ac.uk
(ORCiD: 0009-0003-3960-6990)

Kevin C.H. Tsang
*Institute of Health Informatics*
*University College London*
London, UK
kevin.tsang@ucl.ac.uk
(ORCiD: 0000-0002-7530-5701)

***Abstract*—Asthma control is influenced by various socio-demographic and patient-centered factors. Through statistical analyses on Asthma Mobile Health Study (AMHS) data, 2,452 American patients, including Chi-Square, ANOVA, and mixed effects linear models, significant factors influencing asthma control were identified. The results showed improvements in asthma control during the first six months of using the self-management mobile platform. Notably, socio-demographic factors such as gender, smoking status, and insurance type were found to be significant predictors of asthma control. These findings underscore the potential of mobile health technologies and the importance of specific interventions to address disparities in asthma control, ultimately aiming to enhance patient outcomes and inform healthcare policies.**

***Keywords*—Asthma, Socio-demographic, Self-management, Longitudinal study, Mobile health**

## I. Introduction

Asthma is one of the most common respiratory conditions that affects individuals of all ages, and despite the availability of effective treatments, many patients have poor control and a terrible quality of life [1]. Socio-demographic factors, such as poverty and education, exacerbate asthma conditions by limiting access to care, adherence to treatment plans, and avoidance of environmental triggers [2]. However, traditional methods of asthma control often fail to address real-time needs of patients or capture the dynamic nature of the disease [3]. Mobile health (mHealth) has emerged as a transformative approach to enhance the understanding and management of chronic diseases like asthma. Through Asthma Health App (AHA), the Asthma Mobile Health Study (AMHS) was successfully conducted [7]. Although the data were collected in 2015, AMHS remains one of the largest longitudinal datasets capturing daily and monthly asthma symptoms through patient self-reports via the AHA. Also, AMHS provides high-frequency, real-world insights into asthma progression through remote monitoring.

Additionally, treatment adherence and quality-of-life factors might mediate or moderate the impact of socio-demographic factors on asthma control [4]. Higher adherence to treatment plans has been associated with improved disease control and a reduction in mortality [5]. Quality-of-life factors, such as depression, have been linked to poor asthma control due to its negative impact on self-care behaviors and adherence to treatment [6].

The primary aim of this study is to evaluate whether all asthma patients benefit equally from using the AHA and to investigate whether socio-demographic factors influence the effectiveness of the app in improving asthma control among participants of the AMHS [7]. It further hypothesizes that treatment adherence mediates this relationship, while quality-of-life factors predict changes in asthma control.

To achieve the aim above, the study will focus on the following specific objectives. First, it intends to statistically analyze and quantify the correlation between socio-demographic factors, and asthma control within AMHS participants, identifying specific intervention points based on the strength and direction of these relationships. Additionally, it explores the impacts of treatment adherence on asthma control outcomes by identifying key mediating or moderating variables and offering clear recommendations for optimizing participants' management. Furthermore, it determines how quality-of-life factors, such as mobility and mental well-being, predict changes in asthma control, identifying the most influential predictors.

## II. Methodology

### A. Data Source

Data for this study were collected through the AHA mobile application, released on March 9, 2015 on the Apple App Store. Participants were required to be 18 years or older, have asthma, not be pregnant, and be literate in English. Participants voluntarily downloaded the AHA and self-reported asthma symptoms and demographic information. Socio-demographic data were gathered using participant questionnaires on enrollment. The initial cohort included 5,875 U.S. participants who submitted survey data [7]. Although capturing diverse geographic locations, AHA users were generally more male, wealthier, more educated, and had more severe asthma than the American asthma population. [7]

This analysis focused on 2,452 participants who completed all required daily, weekly, and 90-day questionnaires for months 1 through 6, along with all the aforementioned questionnaires, between March 9, 2015 and September 5, 2015. Only those who fully completed these questionnaires were included in this analysis; incomplete data or failure to meet this criterion led to exclusion. The study enabled continuous data collection and comprehensive analysis of asthma control over time, integrating medication usage, environmental triggers, asthma symptoms, and socio-demographic details. And daily and weekly data collection ensures detailed health and environmental factors' tracking [8].

### *B. Exposures and Outcomes*

The exposures include socio-demographic factors, treatment adherence and quality-of-life variables. Asthma control, the primary outcome, was assessed using metrics from the Global Initiative for Asthma (GINA) [9]. GINA's framework evaluates asthma control based on the frequency and intensity of self-reported symptoms over the past four weeks, including daytime symptoms, night waking, reliever medication use, and activity limitation. Asthma control is categorized into three levels: Controlled, Partly Controlled, and Uncontrolled, as shown in Table I [10].

TABLE I. GINA CRITERIA FOR ASTHMA CONTROL LEVELS BASED ON WEEKLY SYMPTOM FREQUENCY AND ACTIVITY LIMITATION

| Asthma Control Level | Criteria |
|---|---|
| Well Controlled | - Total score = 4:<br>- Daytime symptoms ≤ 2 times/week<br>- No night waking<br>- Quick relief use ≤ 2 times/week<br>- No activity limitation |
| Partly Controlled | - Total score ≥ 2 but < 4:<br>- Some symptoms present (e.g., 1-2 of the conditions are not met) |
| Uncontrolled | - Total score < 2:<br>- Multiple symptoms present (e.g., 3 or more of the conditions are not met) |

### *C. Data Preprocessing and Aggregation*

The race data was one-hot encoded to accommodate mixed race individuals. Key metrics like symptom frequency, were summarized and merged with corresponding daily data, creating a comprehensive dataset. The monthly asthma control scores, based on GINA, were calculated and aggregated using mode by each participant during the month (in the case of equal frequency of two levels, the less controlled level was taken). This process repeatedly formed a six-month longitudinal dataset.

### *D. Analysis of Socio-Demographic, Treatment Adherence, and Quality-of-Life Factors*

This study employed statistical methods to evaluate the relationship between milestone asthma control and a range of socio-demographic, treatment adherence, and quality-of-life features. For categorical features, Chi-Square tests were performed, along with the calculation of Cramér's V to measure the effect size [11]. Cramér's V, derived from the Chi-squared statistic and adjusted for contingency table dimensions via degrees of freedom (*df*), quantifies the strength of association between categorical variables (0 = no association; 1 = perfect association) [12][13]. For $df = 1$, conventional thresholds were small (0.10), medium (0.30), and large (0.50) effect. For $df \geq 2$, thresholds decreased progressively (e.g., $df = 2$: small = 0.07, medium = 0.21, large = 0.35; $df = 5$: small = 0.04, medium = 0.13, large = 0.22), reflecting stricter criteria for complex contingency tables [14]. Statistical significance was defined as $p < 0.05$, with $p < 0.01$ and $p < 0.001$ indicating heightened confidence in rejecting the null hypothesis. The Chi-squared statistic, exact *p*-values, and *df* were reported alongside Cramér's V to ensure comprehensive evaluation of association strength and statistical reliability [14]. The analysis covered socio-demographic factors such as sex, race, and health insurance status, as well as treatment adherence variables like asthma action plans, awareness of the need for an asthma action plan, and lung tests. Additionally, the impact of quality-of-life variables, recorded using EQ-5D on day 180, on asthma control was investigated by analyzing features such as mobility, self-care ability, and so on [8].

For numerical features, One-way ANOVA analysis focused on numerical features including Body Mass Index (BMI) and age [17]. Instead of treating weight and height as separate features, the BMI was calculated and used as a single composite variable to better represent the participants' body composition, avoiding multicollinearity and redundancy. The F-statistic was used to assess the variance between asthma control groups, with a higher F-value indicating a greater likelihood of significant differences among groups [14]. Statistical significance was set at $p < 0.05$.

### *E. Mixed Effects Linear Model*

This study used statistical techniques to analyze factors influencing asthma control after six months. The dependent variable, representing asthma control at 6-month milestone, was converted into a binary variable where 'Well controlled' was coded as 1, and both 'Partly controlled' and 'Uncontrolled' were coded as 0. We used the 6-month milestone data to assess changes in asthma control over time.

To address multicollinearity, variance inflation factor (VIF) was calculated for the predictor variables [16]. A VIF threshold of 10, widely recognized in statistical methodology as an indicator of significant multicollinearity, was applied [12]. Variables exceeding this threshold were removed to ensure model stability and interpretability of fixed effects.

A mixed effects linear model was used to account for both fixed and random effects in the data. These fixed effects represent the population-level impact of these variables on asthma control. The random effects accounted for variability across different individuals. This model was chosen due to its ability to handle grouped data and provide insights into both population-level and individual-level effects. The combination of these methodologies ensured a rigorous and comprehensive analysis of the factors affecting asthma control. The use of mixed effects linear model allowed for the incorporation of individual variability.

### *F. Model Performance Evaluation*

The performance of the mixed effects linear model was evaluated using several metrics, including accuracy, root mean squared error (RMSE) defined as (1), mean absolute error (MAE) defined as (2), and area under the receiver operating characteristic curve (AUC). Accuracy was defined as (3)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (1)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (2)$$

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \tag{3}$$

Where n is the number of samples, $\hat{y}$ is the predicted probability of achieving asthma control at the milestone. The Area Under the Receiver Operating Characteristic (ROC) Curve (AUC) is calculated as the integral of the curve, which plots the true positive rate against the false positive rate. An AUC value of 0.5 indicates no discrimination, while a value of 1.0 indicates perfect discrimination. ROC curve was plotted to visualize the model's performance in distinguishing between well-controlled and poorly controlled asthma cases.

## III. Results

### A. Descriptive Analysis of Asthma Control Trends Over Six Months

The analysis of asthma control across the six months revealed meaningful trends in participants' asthma control over time. Over the six-month period, asthma control among participants showed a gradual improvement in Fig. 1. In the first month, 28.69% of participants were well-controlled, 44.67% were partly controlled, and 26.64% were uncontrolled. By the sixth month, the proportion of well-controlled participants increased to 30.45%, with a 1.76%pt. increase, while the partly controlled group remained steady at around 44%, and the uncontrolled group decreased to 25.58%, reflecting the effectiveness of asthma control. Among all participants, 13 patients achieved a two-level improvement in asthma control (i.e. from uncontrolled to well-controlled), while 74 patients improved by one level (i.e. from uncontrolled to partly controlled or from partly controlled to well-controlled). And 2,334 patients maintained the same control levels throughout the study. However, 31 patients experienced a one-level deterioration in control status, with no cases of two-level worsening observed.

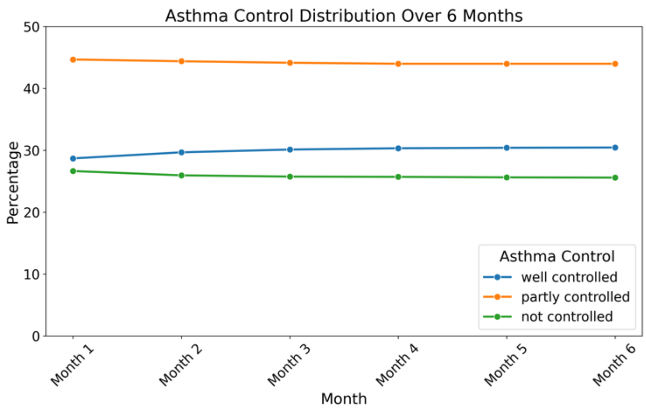


Fig. 1. Distribution of Asthma Control Levels Over Six Months

### B. Key Socio-Demographic Factors Influencing Asthma Control Achievement at Milestone

The Chi-Square tests revealed significant associations between asthma control at the 6-month milestone (completion of 6 months of AHA usage) and various demographic features, as shown in Table II.

Smoking status demonstrated the strongest effect (Cramér's V = 0.28, p < 0.001). Higher income (Cramér's V = 0.20, p < 0.001) and education levels (Cramér's V = 0.17, p < 0.001) were also moderately associated with asthma control, emphasizing the role of socioeconomic status.

Shown in Table II, health insurance coverage showed a moderate effect, while biological sex showed a weaker but significant effect. Among racial groups, American Indian or Alaskan Native participants exhibited a moderate association with asthma control, while Black/African American participants showed a weak association. White participants showed a very weak but statistically significant effect, while Asian participants demonstrated a weak association.

The ANOVA analysis [12] revealed highly significant relationships between numerical features and asthma milestone achievement. Age showed the strongest association (F = 57.53, p < 0.001), followed by BMI (F = 35.41, p < 0.001). These findings emphasize the need to address both age and BMI in asthma control.

TABLE II. Chi-Square Test Results and Effect Sizes for Various Demographic Features on Milestone Asthma Control

| Feature | Chi2 | P-value | Effect Size (Cramér's V) | Degrees of Freedom (df) |
|---|---|---|---|---|
| Biological Sex | 25.55 | < 0.001 | 0.08 | 2 |
| Income | 370.02 | < 0.001 | 0.20 | 12 |
| Education | 265.71 | < 0.001 | 0.17 | 12 |
| Smoking Status | 680.43 | < 0.001 | 0.28 | 4 |
| Health Insurance | 280.00 | < 0.001 | 0.18 | 6 |
| White | 7.93 | 0.02 | 0.04 | 2 |
| Asian | 33.76 | < 0.001 | 0.09 | 2 |
| Black/ African American | 44.56 | < 0.001 | 0.10 | 2 |
| American Indian or Alaskan Native | 75.99 | < 0.001 | 0.13 | 2 |
| Other race | 18.63 | < 0.001 | 0.06 | 2 |

### C. Key Treatment Adherence and Quality-of-Life Factors Influencing Asthma Control Achievement at Milestone

The findings revealed significant associations between asthma control and spirometry testing (Chi2 = 17.34, p < 0.01, Cramér's V = 0.10) and having with an established asthma action plan (AAP) (Chi2 = 7.87, p = 0.02, Cramér's V = 0.10). However, for those without an AAP, awareness of the need for an AAP showed no significant association with asthma control (Chi2 = 2.91, p = 0.23) and the weak effect size (Cramér's V = 0.07).

The Chi-Square tests revealed significant associations between asthma control and several quality-of-life factors, shown in Table III. The strongest association was with overall health perception, indicated by the highest Chi-Square value (Chi2 = 277.33, p < 0.001) and a substantial effect size (Cramér's V = 0.41).

Significant associations were also found for physical mobility (Chi2 = 49.58, p < 0.001, Cramér's V = 0.17) and self-care abilities (Chi2 = 38.56, $p < 0.001$, Cramér's V = 0.15). Usual activities (Chi2 = 20.55, $p < 0.01$, Cramér's V = 0.11) showed a smaller but still significant association.

While pain (Chi2 = 11.33, p = 0.08, Cramér's V = 0.08) and depression (Chi2 = 9.75, p = 0.14, Cramér's V = 0.08) did not reach conventional statistical significance, the low effect sizes suggest these variables may serve as partially independent indicators rather than being directly associated with milestone achievement.

These findings underscore the critical importance of health perception and physical capabilities in maintaining effective asthma control, emphasizing the need to consider quality-of-life factors in treatment plans.

TABLE III. CHI-SQUARE TEST RESULTS AND EFFECT SIZES FOR QUALITY-OF -LIFE FACTORS ON MILESTONE ASTHMA CONTROL

| Feature | Chi2 | P-value | Effect Size (Cramér's V) | Degrees of Freedom (df) |
|---|---|---|---|---|
| Overall Health Perception | 277.33 | < 0.001 | 0.41 | 96 |
| Physical Mobility | 49.58 | < 0.001 | 0.17 | 6 |
| Self-Care Abilities | 38.56 | < 0.001 | 0.15 | 4 |
| Usual Activities | 20.55 | < 0.01 | 0.11 | 6 |
| Pain | 11.33 | 0.08 | 0.08 | 6 |
| Depression | 9.75 | 0.14 | 0.08 | 6 |

## D. *Mixed Effects Linear Model*

The mixed effects linear model included a variety of socio-demographic factors such as biological sex, smoking status, and health insurance status, alongside psychological variables including depression, pain, self-care activities, usual activities, and lung function. Despite some variables showing non-significant associations with asthma control in individual Chi-Square tests and ANOVA test, they were included in the model to evaluate their collective impact.

To address multicollinearity among the predictor variables, VIF analysis was conducted. Variables with VIF values exceeding 10 were removed. The final set of features included in the model, as summarized in Table IV, which were then used in the subsequent mixed effects linear model. According to the categorical variables, they were converted into dummy variables to facilitate model interpretation in Table V.

TABLE IV. OVERVIEW OF SELECTED PREDICTOR VARIABLES FOR THE MIXED EFFECTS LINEAR MODEL AFTER VIF ANALYSIS

| Feature Group | Feature Name | VIF |
|---|---|---|
| Treatment Adherence | Asthma Plan Awareness | 1.95 |
| | Lung Test | 5.35 |
| Quality-of-Life | Depression | 5.50 |
| | Pain | 6.72 |
| | Usual Activities | 7.78 |
| Socio-demographic | Biological Sex | 3.33 |
| | Smoking Status | 3.32 |
| | Health Insurance | 7.09 |
| | Asian | 1.25 |
| | Black/African American | 1.29 |
| | American Indian or Alaskan Native | 1.16 |
| | Other | 1.07 |

The significant variables identified in the model include Asthma Plan Awareness, Gender, Race, Smoking Status, Insurance Type, Depression Levels, Pain Levels, Lung Function Measurement, and Usual Activities. These variables were selected based on their statistical significance ($p < 0.05$), indicating their potential impact on asthma control, their coefficients and p-values are listed in Table VI.

Besides that, the mixed effects linear model includes random effects, capturing the variability across individual participants. The random effects component, with a group variance of 0.032, indicates significant inter-individual variability in asthma control. This suggests that unobserved factors, such as genetic predisposition, may independently affect asthma control outcomes beyond the measured socio-demographic and behavioral variables.

TABLE V. VARIABLE ENCODING AND CORRESPONDING DUMMY VARIABLES FOR MIXED EFFECTS MODEL

| Feature Name | Dummy Variables |
|---|---|
| Asthma Plan Awareness | True, False |
| Lung Test | 1: Yes, 2: No, 3: Not sure |
| Depression | 1: Not anxious/depressed, 2: Slightly anxious/depressed, 3: Moderately anxious/depressed, 4: Severely anxious/depressed, 5: Extremely anxious/depressed |
| Pain | 1: No pain, 2: Slight pain, 3: Moderate pain, 4: Severe pain |
| Usual Activities | 1: No problems, 2: Slight problems, 3: Moderate problems, 4: Severe problems |
| Biological Sex | 1: Male, 0: Female |
| Smoking Status | 1: Never, 2: Current, 3: Former |
| Health Insurance | 1: Private 2: Public insurance, 3: No health insurance, 4: Choose not to answer |
| Asian | 1: True, 0: False |
| Black/African American | 1: True, 0: False |
| American Indian or Alaskan Native | 1: True, 0: False |
| Other | 1: True, 0: False |

TABLE VI. SUMMARY OF SIGNIFICANT FEATURES, COEFFICIENTS, AND P-VALUES IN THE MIXED EFFECTS LINEAR MODEL

| Feature | Coefficient | P-value |
|---|---|---|
| Asthma Plan Awareness | -0.062 | <0.001 |
| Gender (Female vs. Male) | -0.028 | <0.001 |
| Asian | 0.016 | <0.001 |
| American Indian/Alaskan Native | 0.058 | <0.001 |
| Black/African American | -0.035 | <0.001 |
| Smoking Status (Current Smoker) | 0.036 | < 0.001 |
| Smoking Status (Former Smoker) | 0.023 | < 0.001 |
| Insurance Type (Public) | -0.031 | < 0.001 |
| Insurance Type (No Insurance) | 0.011 | < 0.001 |
| Depression (Slight) | -0.017 | < 0.001 |
| Depression (Moderate) | -0.025 | < 0.001 |
| Depression (Severe) | -0.009 | < 0.001 |
| Pain (Slight) | -0.009 | < 0.001 |
| Pain (Moderate) | -0.005 | 0.002 |
| Pain (Severe) | -0.001 | 0.53 |
| Lung Function (No Test) | -0.017 | < 0.001 |
| Lung Function (Unsure) | -0.035 | < 0.001 |
| Usual Activities (Slight Limitations) | 0.008 | < 0.001 |
| Usual Activities (Moderate Limitations) | < -0.001 | 0.89 |
| Usual Activities (Severe Limitations) | 0.001 | 0.54 |

The performance of the mixed effects linear model was evaluated using established metrics. The model achieved an accuracy of 96.33%, with an RMSE of 0.18 and an MAE of 0.09, indicating a good overall fit to the data. The AUC was 0.91, reflecting strong discriminatory power in distinguishing between well-controlled (class 1: 693 observations) and poorly controlled asthma cases (class 0: 14,795 observations), see Fig. 2. However, because of the severe class imbalance, Precision-Recall (PR) curves were prioritized over ROC-AUC to better assess the model's ability to identify the minority class (well-controlled asthma) [18]. By focusing solely on the positive class, PR curves avoid the overoptimistic evaluations often driven by the majority class. As shown in Fig. 3, the PR curve illustrates the trade-off between precision and recall, achieving an average precision of 0.64, a 14.3-fold increase in precision.

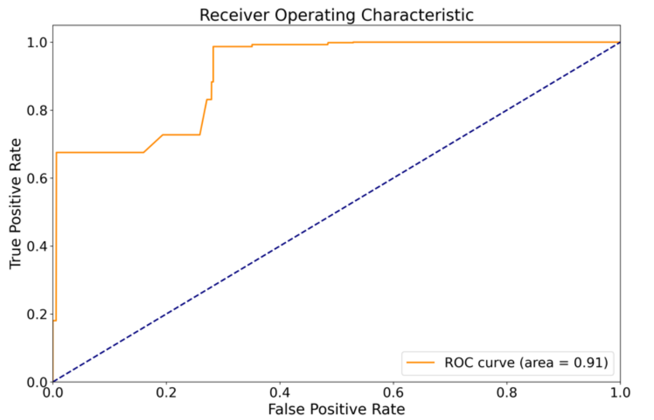


Fig. 2. ROC Curve for Mixed Effects Linear Model of Well- vs Poor-Controlled Asthma. ROC-AUC = 0.91

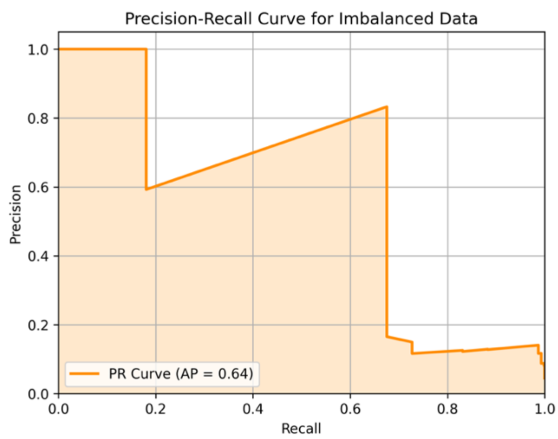


Fig. 3. Precision-Recall Curve for Asthma Control Prediction. PR-AUC = 0.64, class balance = 0.045

## IV. DISCUSSION

This study observed improvements in asthma control over the first six months of using the asthma self-management platform, with an increase in the proportion of participants achieving well-controlled asthma 28.69% at baseline to 30.45% after six months. Although the overall changes in proportions were modest, the improvements highlight the platform's potential to facilitate meaningful clinical progress for certain individuals. However, these findings also suggest that while the asthma self-management platform has demonstrated a positive effect for some patients, broader and more consistent improvements across the entire population may require further refinement of the intervention.

The mixed effects linear model also revealed significant inter-individual variability in asthma control, emphasizing that factors such as asthma severity and adherence behaviors contribute to the variability in outcomes. Therefore, it highlights the importance of personalized asthma control plans, as individual differences can significantly affect treatment outcomes. To better understand the factors driving these improvements, socio-demographic, adherence behaviors, and quality-of-life determinants were explored in detail.

### *A. Socio-demographic factors and asthma control after 6 months of AHA use*

For gender, the Chi-square test (Chi2 = 25.55, Cramér's V = 0.08) suggested a weak association with asthma control, while the model (coefficient = -0.028, $p < 0.001$) indicated that being female is significantly associated with better asthma control compared to being male, which may be due to women experiencing better quality of care than men, particularly in screening and treatment-related aspects [19].

For smoking status, the Chi-square test demonstrated a moderate association with asthma control (Chi2 = 680.43, Cramér's V = 0.28). And the mixed effects model showed that current smokers (coefficient = 0.036, $p < 0.001$) and former smokers (coefficient = 0.023, $p < 0.001$) had better asthma control compared to non-smokers, which is counterintuitive given the known negative impacts of smoking on Chronic Obstructive Pulmonary Disease [17]. It may reflect indication bias, where smokers may be aware of their heightened risks, adhere more strictly to asthma management plans. Furthermore, individuals with milder asthma may feel safer to smoke, underestimating the risks. Therefore, it suggests emphasizing the risks associated with smoking and to closely monitor adherence to smoking cessation plans.

For health insurance, the Chi-square test (Chi2 = 280.00, Cramér's V = 0.18) revealed a moderate association. Compared to private insurance, public insurance was associated with worse asthma control (coefficient = -0.031, $p < 0.001$, while no health insurance was linked to better control (coefficient = 0.011, $p < 0.001$). Public health insurance generally increases access to healthcare; however, its impact on financial protection and health outcomes varies, with some studies showing no significant effect or even negative effects on health outcomes [16]. Those without insurance may have milder asthma and only seek care during severe situations, leading to underreporting of consistent management. Therefore, enhancing the quality of public insurance and understanding alternative control strategies used by the uninsured could improve asthma outcomes.

Regarding race, the Chi-square test revealed weak to moderate associations. Asians (coefficient = 0.016, $p < 0.001$) and American Indians or Alaskan Natives (coefficient = 0.058, $p < 0.001$) demonstrated better asthma control than Black or African American participants, who showed worse control (coefficient = -0.035, $p < 0.001$). Existing literature shows that Black or African American individuals often experience poorer health than whites [20]. Conversely, better asthma control among American Indian or Alaskan Native and Asian participants may be linked to specific cultural practices, though this warrants further study.

### *B. Adherence to asthma action plans and clinical practices*

Although Chi-square analysis revealed no significant association on asthma control, for those without an existing AAP, AAP Awareness showed a coefficient of -0.062 ($p < 0.001$), indicating that increased awareness of AAP (as prompted by the app) was significantly associated with worse asthma control. This unexpected finding might stem from the complexity of asthma action plans. Patients with more severe asthma are often the ones who have developed an AAP, while patients with less severe asthma have not been made aware of the plan. Moreover, patients with moderate severity but poorer management might be aware of the AAP but have yet to fully develop or adhere to it. This suggests a need for further research to refine these plans and ensure they are both effective and manageable for patients.

Although the Chi-square test showed weak associations between asthma control and spirometry testing (Chi2 = 17.34, Cramér's V = 0.10), spirometry tests are crucial for effective asthma control. Lung function measurement status showed that participants who had not had their lung function tested (coefficient = -0.017, $p < 0.001$) or were unsure (coefficient = -0.035, $p < 0.001$) had worse asthma control. Lung function tests, such as spirometry and provocation tests, help diagnose and monitor asthma by measuring breathing and airway inflammation [21].

### *C. Mental health and quality of life*

Depression levels were negatively associated with asthma control. Participants with slight depression (coefficient = -0.017, $p < 0.001$), moderate depression (coefficient = -0.025, $p < 0.001$), and severe depression (coefficient = -0.009, $p < 0.001$) all showed poorer asthma control outcomes. Similarly, pain negatively affected asthma control, with slight pain (coefficient = -0.009, $p < 0.001$) and moderate pain (coefficient = -0.005, $p = 0.002$) linked to worse control, though severe pain had no significant impact (coefficient = -0.001, $p = 0.53$).

Although the Chi-square test revealed weak associations between pain (Chi2 = 11.33, Cramér's V = 0.08) and depression (Chi2 = 9.75, Cramér's V = 0.08) with asthma control, the model suggested that higher levels of both were consistently linked to poorer asthma control, as they may impair self-care and treatment adherence. Collaborative care intervention has shown success in improving self-care support, patient confidence, and clinical outcomes. This intervention increased participants' ability to adhere, which means similar approaches could benefit asthma control by enhancing self-efficacy and overall well-being [22].

For usual activities, the Chi-square test revealed a moderate association (Chi2 = 20.55, Cramér's V = 0.11). Slight limitations were associated with better asthma control (coefficient = 0.008, $p < 0.001$), while moderate or severe limitations showed no significant effect. Interventions with slight limitations in activities could be effective as slight problems with usual activities have a significant association with asthma control. Such patients could benefit from external support, like personalized AAP. However, those who have moderate to severe problems may have more variability in self-reported symptoms, potentially due to fluctuating physical quality. Therefore, extra care, including both physical and mental, could be given to these patients.

### *D. Limitations and Future Work*

The study has several limitations. The reliance on self-reported data may introduce reporting bias. Also, the generalizability of the findings may also be limited to the AMHS dataset's population. The use of an iPhone app skews the participant pool towards higher education, higher income levels and younger participants, alongside the underrepresentation of Black individuals [8].

There are also limitations in the analysis. The potential for multicollinearity among predictor variables, despite the use of VIF to mitigate this issue, Least Absolute Shrinkage and Selection Operator (LASSO) can effectively address multicollinearity by reducing VIF values below 10 and demonstrating stronger interpretative ability compared to using VIF alone [23]. Future studies might explore alternative modelling techniques.

Lastly, the reliance on a single modelling approach might limit the insights gained. Employing a combination of different statistical and machine learning models and feature importance could offer a more comprehensive understanding of the factors influencing asthma control [24].

## V. Conclusion

This study provides valuable insights into the trends and determinants of asthma control over six months of using a self-management platform. An increase in the proportion of well-controlled participants from 28.69% to 30.45% underscores the platform's potential to support meaningful clinical progress for specific individuals. Significant socio-demographic factors, such as smoking status, and race, alongside quality-of-life metrics and adherence behaviors, were identified as key predictors of asthma control. The mixed effects linear model demonstrated strong performance in identifying associations, achieving high accuracy and discriminatory power. However, the study also revealed limitations in achieving broader and consistent improvements across all participants, suggesting that further refinement of the platform and its interventions are needed. These results underscore the importance of integrating socio-demographic, behavioral, and clinical factors into comprehensive asthma management strategies to enhance the efficacy of self-management tools and improve overall patient outcomes.

## Acknowledgment

We would like to thank the patients who participated in this study through the Asthma Mobile Health Study (AMHS), conducted by the Icahn School of Medicine at Mount Sinai. Their valuable contributions and willingness to share their data were essential to this research.

## References

[1] Y. Liang and J. C. W. Mak, "Inhaled therapies for asthma and chronic obstructive pulmonary disease," Current Pharmaceutical Design, vol. 27, no. 12, pp. 1469–1481, Nov. 2020.

[2] W. Xi, S. Banerjee, M. Olfson, G. S. Alexopoulos, Y. Xiao, and J. Pathak, "Effects of social deprivation on risk factors for suicidal

ideation and suicide attempts in commercially insured US youth and adults," Scientific Reports, vol. 13, no. 1, 2023.

[3] S. Acharya and R. Sarra, "Proactive Real-Time Solution for Asthma Management," in 2014 IEEE International Conference on Bioinformatics and Bioengineering, 2014, pp. 270–276.

[4] B. G. Toelle, G. B. Marks, and S. M, "Psychological and Medical Characteristics Associated with Non-Adherence to Prescribed Daily Inhaled Corticosteroid," Dunn, Journal of Personalized Medicine, vol. 10, no. 3, September 2020.

[5] M. George and B. Bender, "New insights to improve treatment adherence in asthma and COPD," Patient Preference and Adherence, vol. Volume 13, pp. 1325–1334, July 2019.

[6] S. I. Leonard, E. R. Turi, J. S. Powell, J. Usseglio, K. K. MacDonell, and J.-M. Bruzzese, "Associations of asthma self-management and mental health in adolescents: A scoping review," Respiratory Medicine, vol. 200, p.106897, 2022.

[7] Y.-F. Y. Chan et al., "The Asthma Mobile Health Study, a large-scale clinical observational study using ResearchKit," Nature Biotechnology, vol. 35, no. 4, pp. 354–362, Apr. 2017.

[8] Y.-F. Y. Chan *et al.*, "The asthma mobile health study, smartphone data collected using ResearchKit," Scientific Data, vol. 5, no. 1, p. 180096, May 2018.

[9] Global Initiative for Asthma, "Reports - Global Initiative for Asthma - GINA," 2024. [Online]. Available at: https://ginasthma.org/reports/

[10] K. Pearson, "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling," The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, vol. 50, pp. 157–175, 1900.

[11] H. Cremér and Princeton University, Mathematical methods of statistics. Princeton, NJ: Princeton University, 1999.

[12] M. H. Kutner, Applied linear statistical model. Mcgraw-Hill Irwin, 2005.

[13] C. J. Ferguson, "An effect size primer: A guide for clinicians and researchers," Professional Psychology: Research and Practice, vol. 40, no. 5, pp. 532–538, 2009.

[14] J. Cohen, Statistical Power Analysis for the Behavioral Sciences, 2nd ed. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1988.

[15] E. Vittinghoff, Regression methods in biostatistics: Linear, logistic, survival, and repeated measures models. New York: Springer Publishing Co., 2005.

[16] C. E. Bird et al., "How Do Gender Differences in Quality of Care Vary Across Medicare Advantage Plans?," Journal of General Internal Medicine, vol. 33, no. 10, pp. 1752–1759, Aug. 2018.

[17] R. Laniado-Laborín, "Smoking and chronic obstructive pulmonary disease (COPD). Parallel epidemics of the 21st century," International Journal of Environmental Research and Public Health, vol. 6, no. 1, pp. 209–224, Jan. 2009.

[18] J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curves," Proceedings of the 23rd International Conference on Machine Learning (ICML), 2006, pp. 233–240.

[19] D. Erlangga, M. Suhrcke, S. Ali, and K. Bloor, "The impact of public health insurance on health care utilisation, financial protection and health status in low- and middle-income countries: A systematic review," PLOS ONE, vol. 14, no. 8, August 2019.

[20] D. R. Williams and M. Sternthal, "Understanding Racial-ethnic Disparities in Health: Sociological Contributions," Journal of Health and Social Behavior, vol. 51, no. 1_suppl, pp. S15–S27, March 2010.

[21] M. Gallucci, P. Carbonara, A. M. G. Pacilli, E. Di Palmo, G. Ricci, and S. Nava, "Use of symptoms scores, spirometry, and other pulmonary function testing for asthma monitoring," Frontiers in Pediatrics, vol. 7, Mar. 2019.

[22] E. J. Ludman *et al.*, "Improving Confidence for Self Care in Patients with Depression and Chronic Illnesses," Behavioral medicine (Washington, D.C.), vol. 39, no. 1, pp. 1–6, 2013.

[23] Y. Yuan, X. Wang, M. Shi, and P. Wang, "Performance comparison of RGB and multispectral vegetation indices based on machine learning for estimating Hopea hainanensis SPAD values under different shade conditions," Frontiers in Plant Science, vol. 13, Jul. 2022.

[24] K. C. H. Tsang, H. Pinnock, A. M. Wilson, and S. A. Shah, "Application of Machine Learning to Support Self-Management of Asthma with mHealth" in 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2020, pp. 5673–5677.