380 **SUPPLEMENTARY DOCUMENT: STUDY PROTOCOL**

381 **TITLE:** Racial and ethnic disparities in healthcare utilization for eczema

382

383 **AIM:** To investigate the association between race/ethnicity and healthcare utilization (i.e.,

384 ambulatory visits, emergency visits, prescriptions filled) for eczema.

385

386 **HYPOTHESIS:** Healthcare utilization for eczema differs by race/ethnicity.

387

388 **SIGNIFICANCE:**

389 Differences in the prevalence of common dermatologic disorders, including eczema, have been

390 identified according to race/ethnicity. While the association between race/ethnicity and

391 healthcare utilization is amassing, studies examining multiple healthcare utilization outcomes

392 with adequate granularity in the context of dermatologic conditions are limited. Identification of

393 disparities in healthcare utilization can help identify subpopulations that could benefit from

394 specific allocation of education, social, and financial resources with respect to common

395 dermatologic conditions.

396

397 **INNOVATION:** Few studies have examined healthcare utilization for eczema according to

398 race/ethnicity. This study is intended to add to the literature in this understudied research area.

399

400 **STUDY DESIGN:** The Medical Expenditure Panel Surveys are a series of surveys with

401 longitudinal design consisting of two years of follow-up. The studies are large-scale United

402 States national surveys conducted from 1996-2013 of families and individuals, their medical

403 providers, and employers. Individuals included in these studies are a nationally representative

404 subsample of households from the previous year's National Health Interview Survey (National

405 Center for Health Statistics). The survey uses a complex design, which includes a stratified

406 multistage probability design with overlapping panels to collect detailed information for each

407 participant's demographic and socioeconomic characteristics, health status and conditions,

408 medical services usage, charges and source of payments, insurance coverage, income, and

409 employment. We will perform a cohort using data pooled from the serial longitudinal cohorts.

410

411 **SOURCE POPULATION:** Data will be obtained from the Medical Expenditure Panel Survey,

412 which is the most complete source of data on the cost and use of health care and health insurance

413 coverage in the United States. These data are publicly available through the Agency for

414 Healthcare Research and Quality.

415

416 Exposed groups will be defined as follows:

417 - Individuals with ICD9 code 691-692 (i.e., atopic dermatitis/contact dermatitis/other

418 eczema) AND

419 - Individuals who are of racial/ethnic minority status (i.e., Hispanic, non-Hispanic black,

420 non-Hispanic Asian/Pacific Islander, non-Hispanic Native American or Alaskan/Aleutian

421 Native, and other non-Hispanic non-white race/ethnicity) AND

422 - Individuals between the ages of 0 and 17

423

424 <u>Unexposed group will be defined by all of the following:</u>

425   -   Individuals with ICD9 code 691-692 (i.e., atopic dermatitis/contact dermatitis/other

426   eczema) AND

427   -   Individuals who are non-Hispanic white AND

428   -   Individuals between the ages of 0 and 17

429

430 Patients with incomplete follow-up information and thus missing information on population-

431 based longitudinal weights will be excluded from analyses. Patients without eczema will be

432 included as a separate subpopulation when conducting survey analyses to ensure accurate

433 calculation of variance estimation.

434

435 **EXPOSURE VARIABLE:** The exposure of interest will be minority race/ethnicity compared

436 with non-Hispanic white. Minority race/ethnicities include:

437   a. Hispanic white

438   b. Non-Hispanic black

439   c. Non-Hispanic Asian/Pacific Islander

440   d. Non-Hispanic Native American or Alaskan/Aleutian Native

441   e. Other non-Hispanic non-white race/ethnicity

442

443 **OUTCOME VARIABLE:** The primary outcomes will be healthcare utilization as defined by

444   the following within the two-year follow-up period

445   a. Any ambulatory visits associated with eczema

446   b. Dermatology ambulatory visits associated with eczema

447   c. Emergency visits associated with eczema

448   d. Prescription fills associated with eczema

449

450   **COVARIATES/POTENTIAL CONFOUNDERS:** We will examine possible confounding and

451   effect modification by factors known to be associated with healthcare utilization. The variables

452   that we will consider exploring in our multivariable statistical models include but are not limited

453   to the following: age, sex, region of residence, measures of socioeconomic status, household

454   income, employment status, health insurance status, atopic comorbidities. Care will be taken to

455   distinguish confounders from mediators, and only confounders will be adjusted for.

456

457   **STATISTICAL ANALYSIS:**

458   We will first perform descriptive statistics to summarize demographic information (e.g., age, sex,

459   region of residence, measures of socioeconomic status, household income, employment status,

460   health insurance status, atopic comorbidities) according to race/ethnicity. The global F-test from

461   simple linear regression will be used for continuous variables; the Rao-Scott design-based chi-

462   squared test will be used for categorical variables. We will use logistic regression to determine

463   the unadjusted and adjusted associations between race/ethnicity and dichotomous measures of

464   healthcare utilization for each of the dermatologic conditions of interest. The measure of

465   association will be odds ratios. We will use negative binomial or Poisson regression to determine

466   the unadjusted and adjusted associations between race/ethnicity and continuous measures of

467   healthcare utilization. The measure of association will be incidence rate ratios. In building a

468   multivariable statistical model, we will use all of the aforementioned potential confounding

469   variables in the model and then use a backward selection approach to reach a parsimonious

470   model and will remove any covariates from the model that are not statistically significant and

471   also do not change the point estimate of the exposure variable by more than 10%. All statistical

472   analyses will take into account the complex multi-stage sampling design to ensure appropriate

473   variance estimates as well as the population-based weights to ensure that the results are

474   generalizable to the United States population distribution.

475

476   **POTENTIAL BIAS:** Biases that may be introduced in our study include information bias (e.g.,

477   recall and reporting biases), which may result from misclassification of the exposure and/or

478   outcome. Furthermore, despite adjusting for covariates, residual confounding may exist and

479   introduce bias due to presence of unmeasured confounders.

480

481   **GENERALIZABILITY:** The Medical Expenditure Panel Survey data being used is highly

482   representative of the United States civilian non-institutionalized population, and is the most

483   complete source of data on the cost and use of health care and health insurance coverage in the

484   United States. Thus, our results will be generalizable to this large population.