SUPPLEMENTAL MATERIAL

Supplementary Appendix 1. **Geocoding Process**.

We used a multistep strategy for matching the geocode attributes of residences. Initially, all the residence information was entered into the automatic match process with ArcMap™ version 9.3.1. The sensitivity and specificity with which ArcMap™ suggests coordinates for entered addresses can be manually modified. Based on the automatic matching, there were three possibilities: exact match, tie; i.e. addresses for which more than one point on the map was proposed, and not-match; i.e. residence addresses for which there were no points identified in the ArcMap™ database of geocodes. We used a threshold score of ≥80 or higher (out of 100) to designate exact matches. To test the accuracy of such a cutoff, we also manually obtained the geocode coordinates for a randomly selected sample of 25 cases and received the exact coordinates as provided via the automatic matching system. Tied and unmatched addresses were entered into the automatic match process with ArcMap™ version 10.0, which includes updated information about street addresses. This process led to an incremental number of exact matches. We manually reviewed the addresses that remained tied or unmatched after two series of automatic matching. Exact match was assigned to some of these remaining addresses after manual review, with common problems being typographic mistakes during the initial data collection leading to unrecognizable addresses in the ArcMap™ Database, and suggestion of two identical match addresses in the proposed list of ArcMap™ geocode coordinates. If the exact match could not be assigned after manual address review, we looked up the proposed residence addresses in Google Maps™ (Google Inc., Santa Clara County, CA) and MapQuest™ (MapQuest, Inc.,

Denver, CO) to find the exact coordinates for the residence address. If exact coordinates were not obtainable through ArcMap™, Google Maps™, or MapQuest™, attempts were made to match the address to the smallest geographical region possible, i.e. the exact geographic coordinates, block groups, or census tracts. For rare cases wherein several candidate addresses remained, we chose the candidate address that lied at the top of the ArcMap™ candidate list. If the candidate address list in ArcMap™ was blank, then we chose a candidate address in Google Maps™ or MapQuest™ that had the closest street number, and noted down its coordinates (Figure 1). Accordingly, using the derived coordinates, we determined the census tract of residence for each patient.