

## HOW ACCURATELY DOES GEOCODING DETERMINE RESIDENTIAL LOCATIONS IN RURAL AND URBAN AREAS?

### ABSTRACTS

Epidemiology. 14(5) Supplement:S102-S103, September 2003.

*Ward, M H; Giglierano, J; Wolter, C; Miller, RS; Nuckols, JR; Hartge, P*

Geographic information systems are increasingly being used for the investigation of environmental exposures and health. Study participants' residences are often mapped after geocoding, in which the street address is matched to a digital street centerline database. Geocoding can result in positional errors in the residential location due to errors in the street centerline locations or due to incorrect or incomplete address ranges and street names. For rural addresses, the positional error may be greater than that for urban addresses, due to large distances between homes and the greater distance of homes from the roads. As a part of a population-based case-control study of non-Hodgkin lymphoma in Iowa, United States, we collected global positioning system (GPS) measurements near the front door of 634 participants' homes. We used the Garmin GPS12 Personal Navigator, a 12-channel small handheld receiver. The interviewers were instructed to find an area with an unobstructed view of most of the sky about 6 meters away from the house. GPS measurements were taken from December 1998 to February 2001. Seventy percent of the measurements were taken before selective availability (SA), the deliberate corruption of the satellite signals by the U.S. government, was ended on May 2, 2000. Because of the potential for positional errors of approximately 100 meters or more in these early data and the potential for interviewer error in recording the coordinates, we did extensive quality control checking of the GPS coordinates by comparing the participants' address information to various databases; we corrected the coordinates if necessary. We did this detailed checking for 235 residences located in the south central Iowa. We estimated that 14% of the GPS coordinates were more than 100 meters away from the actual residence location. The median distance was significantly greater before the removal of SA (31 meters; IQR = 0-87) compared with after (12 meters; IQR = 0-49). Based on the verified GPS coordinates, the 235 homes were classified as community residences (74%) if they were located within towns and rural (26%) if they were outside town boundaries. A commercial geocoding firm geocoded the addresses using a 12.2 meter (40 feet) offset from the street centerline. After correcting address misspellings, the final geocoding success rate was 94%. We calculated the distance between the geocoded address and the verified GPS coordinates. Overall, the median positional error was 57 meters (IQR 32-122). The error was significantly greater for the rural residences (median = 242 meters; IQR 88-828) compared with urban residences (median = 46; IQR 26-76) indicating the potential for greater misclassification of proximity-based exposures for rural study participants.

(C) 2003 Lippincott Williams & Wilkins, Inc.