COMPARISON OF THREE DIFFERENT GEOCODING TECHNIQUES IN FRANCE

Sofia Temam, *Inserm, U1018, CESP & Univ Paris 11, UMRS 1018, Villejuif, France*
Christophe Declercq, *Institut de veille sanitaire, St Maurice, France*
Morgane Stempfelet, *Institut de veille sanitaire, St Maurice, France*
Francine Kauffmann, *Inserm, U1018, CESP & Univ Paris 11, UMRS 1018, Villejuif, France*
Françoise Clavel-Chapelon, INSERM, U1018, CESP & Univ Paris 11, UMRS 1018, Villejuif, France
Valérie Siroux, *Inserm, U823, Centre de recherche Albert Bonniot & Univ Joseph Fourier, Grenoble, France*
Raphaëlle Varraso, *Inserm, U1018, CESP & Univ Paris 11, UMRS 1018, Villejuif, France*
Nino Künzli, *Swiss Tropical and Public Health Institute & Basel University, Basel, Switzerland*
Benedicte Jacquemin,* *Inserm, U1018, CESP & Univ Paris 11, UMRS 1018, Villejuif, France*
*presenting author

Background and aims: The development of more and more precise techniques to assess environmental exposures, particularly using geographic information systems (GIS), has raised the necessity to geocode the addresses of participants in cohort studies. Our objective is to assess the distance of the geocodes obtained for the same address using different techniques of geocoding in French urban areas.
Methods: Within the context E3N, a French Cohort, we geocoded 1809 addresses in 4 urban areas, with 3 techniques:
A. automatically by a commercial firm that uses NavTEQ® streets network to geocode by spatial interpolation along street axis. Location of the coordinates corresponds to 15m to the left or the right from the street centre (depending on the street number if even or odd)
B. automatically using a free Internet service developed using Google Maps® Internet mapping application programming interface (API) where the returned location of the coordinates corresponds (in theory) to the entrance of the building/parcel on the street
C. manually using the free on-line French cadastral maps and locating the coordinates in the center of the building, these were considered as the "true" ones.
Results: We only kept the 1656 addresses for which the 3 techniques found the complete address without doubt. The median difference in m between A and C was 21 (IQR 12-41), and between B and C was 23 (14-45). Comparing A and C, the distance was between 50 and 75m for 9% of the addresses and above 75m for 11%, comparing B and C the percentages were 10 and 12 respectively.
Conclusion: We did not observe large differences between the different techniques however around 20% of the addresses were geocoded at more than 50m from the "true" location. It will be interesting to assess if this alters the exposure estimates and the health effects.
Funding: ESCAPE(EU-FP7)