
Figures and figure supplements

Approximating missing epidemiological data for cervical cancer through Footprinting: A case study in India

Irene Man *et al.*

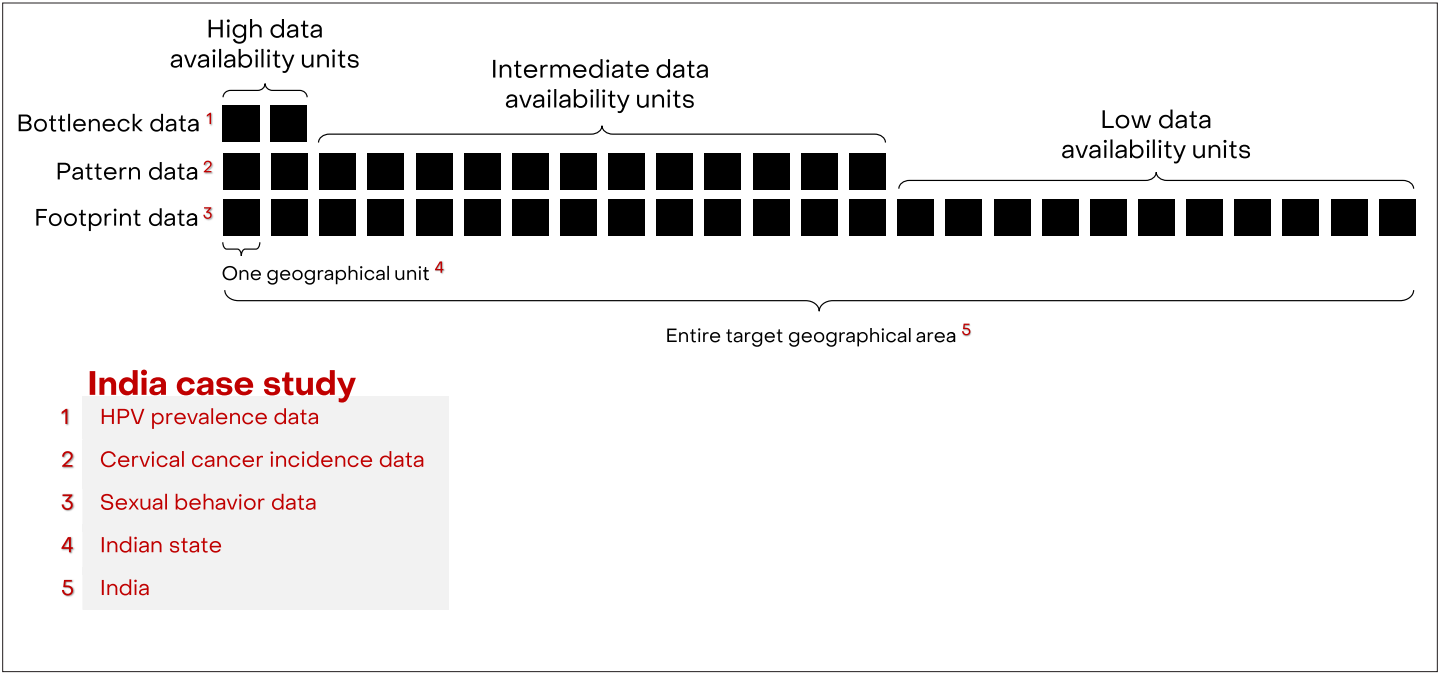
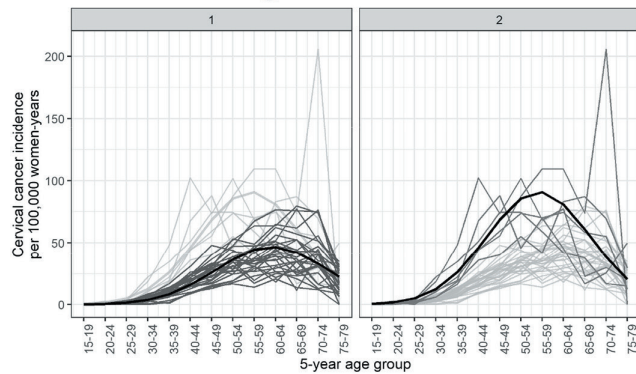
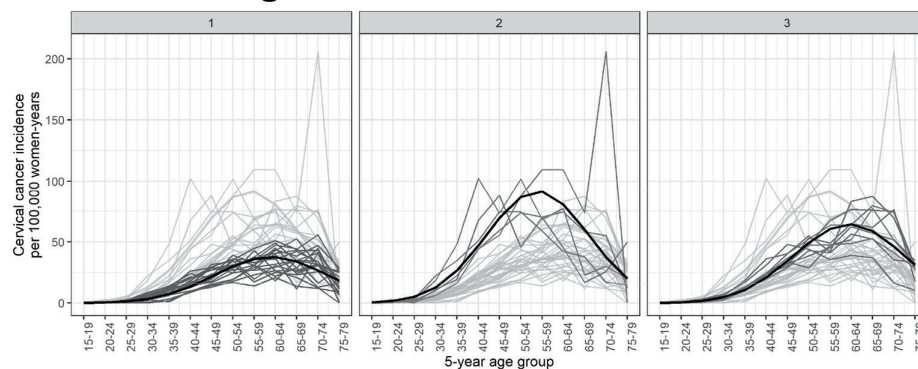


Figure 1. Hierarchical structure of availability of cervical cancer epidemiological data.

A. 2-clustering



B. 3-clustering



C. 4-clustering

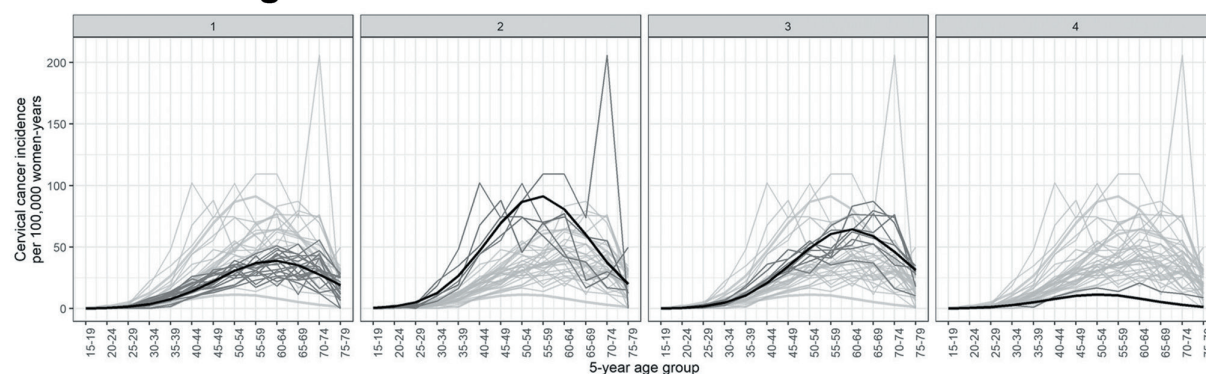


Figure 2. Identified clusters of registry-specific cervical cancer incidence. Clusterings under (A) 2, (B) 3, and (C) 4 prefixed clusters. Each panel within a row corresponds to a cluster within a k -clustering, with the cluster label given on top of the panel. The cervical cancer incidence data were extracted from volume XI of Cancer Incidence in Five Continents (CI5) (Bray et al., 2017) and the 2012–2016 report by the Indian National Centre for Disease Informatics and Research (NCDIR) (Report of National Cancer Registry Programme, 2020). Black: cluster mean of cervical cancer incidence; dark grey: registry incidence assigned to the cluster; light grey: registry incidence assigned to other clusters.

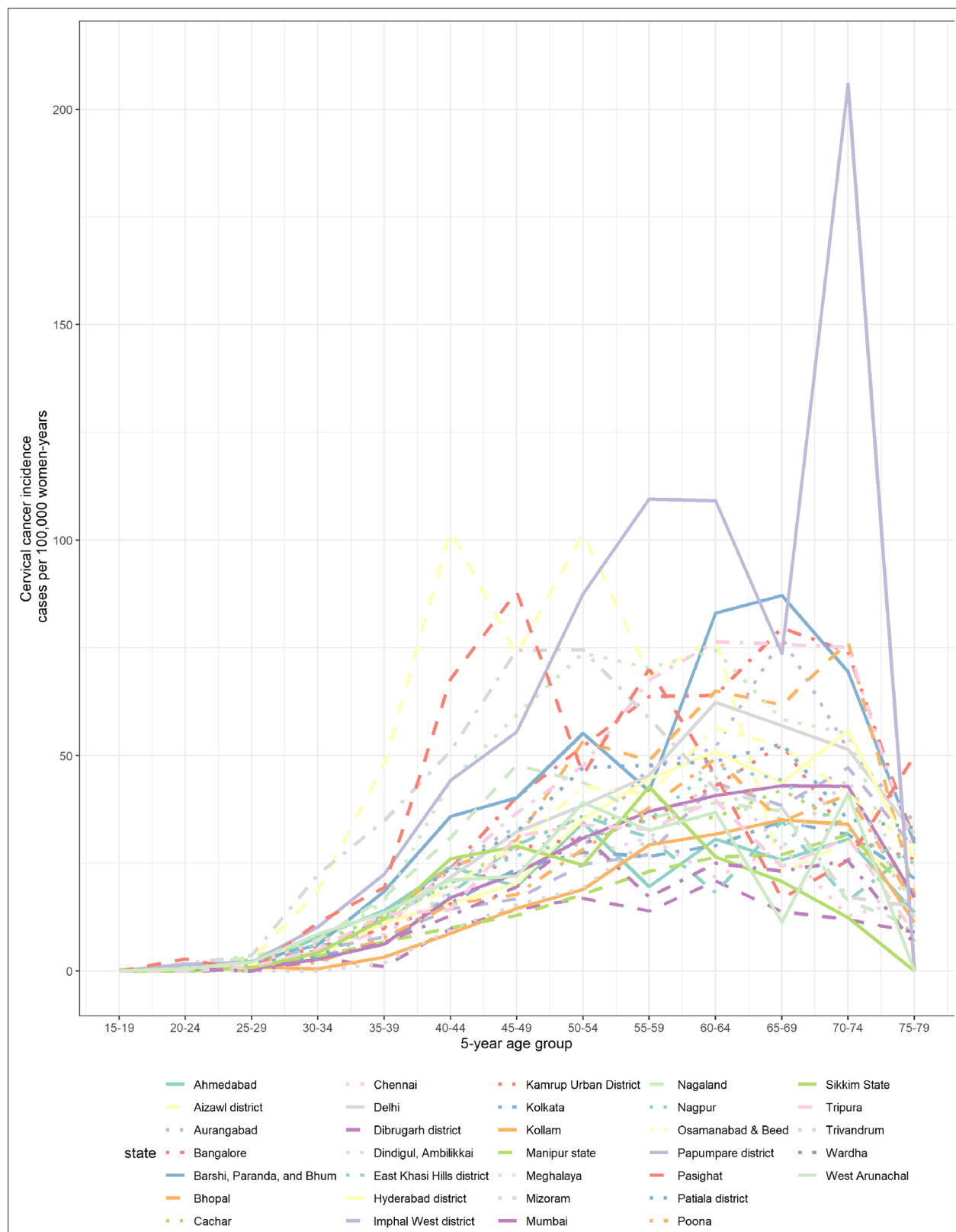


Figure 2—figure supplement 1. Registry-specific cervical cancer incidence data from Cancer Incidence in Five Continents (CI5) and National Centre for Disease Informatics and Research (NCDIR). See **Figure 1—source data 1** for whether registries belong to CI5 or NCDIR.

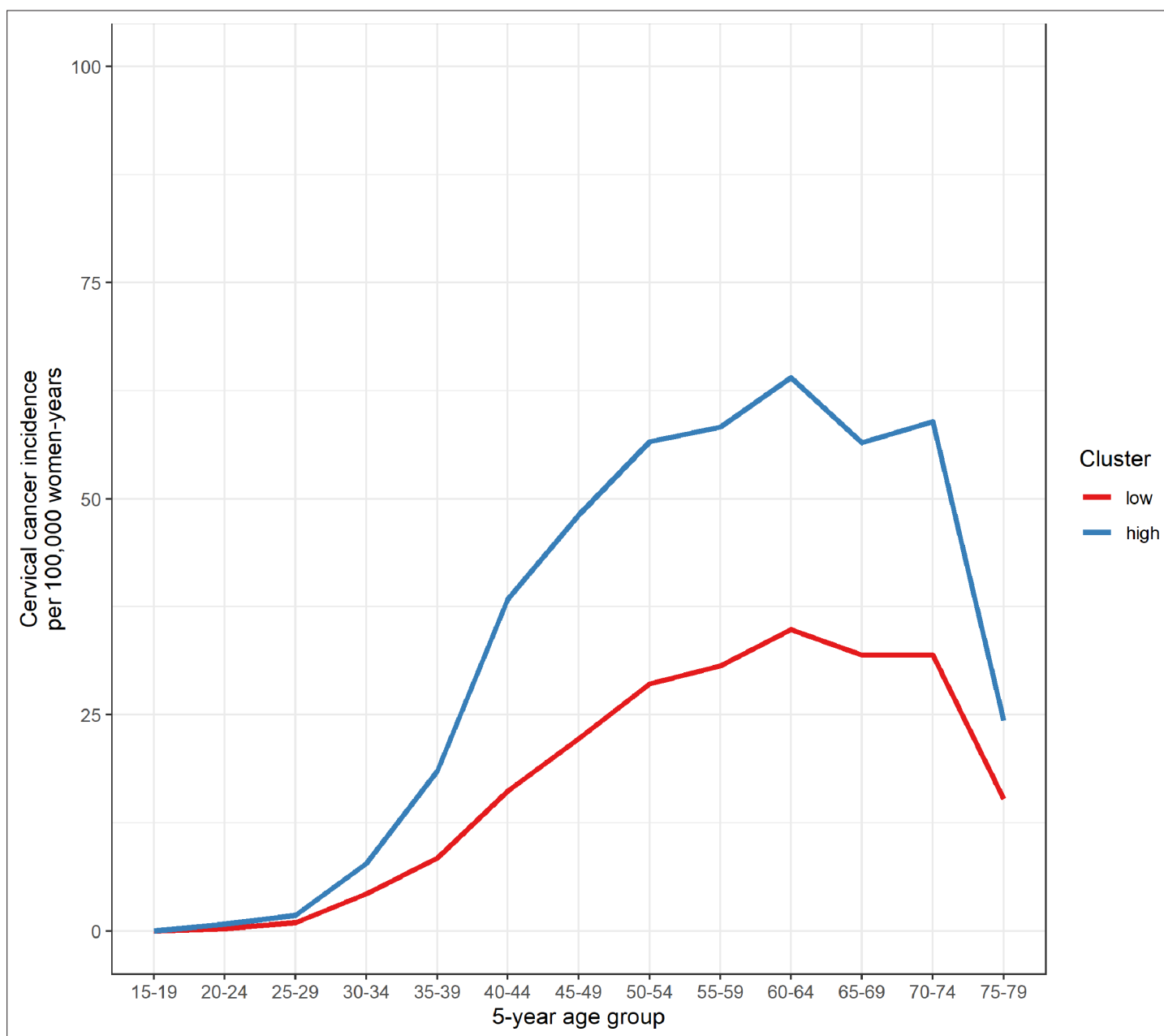


Figure 2—figure supplement 2. Mean age-specific cervical cancer incidence by cluster.

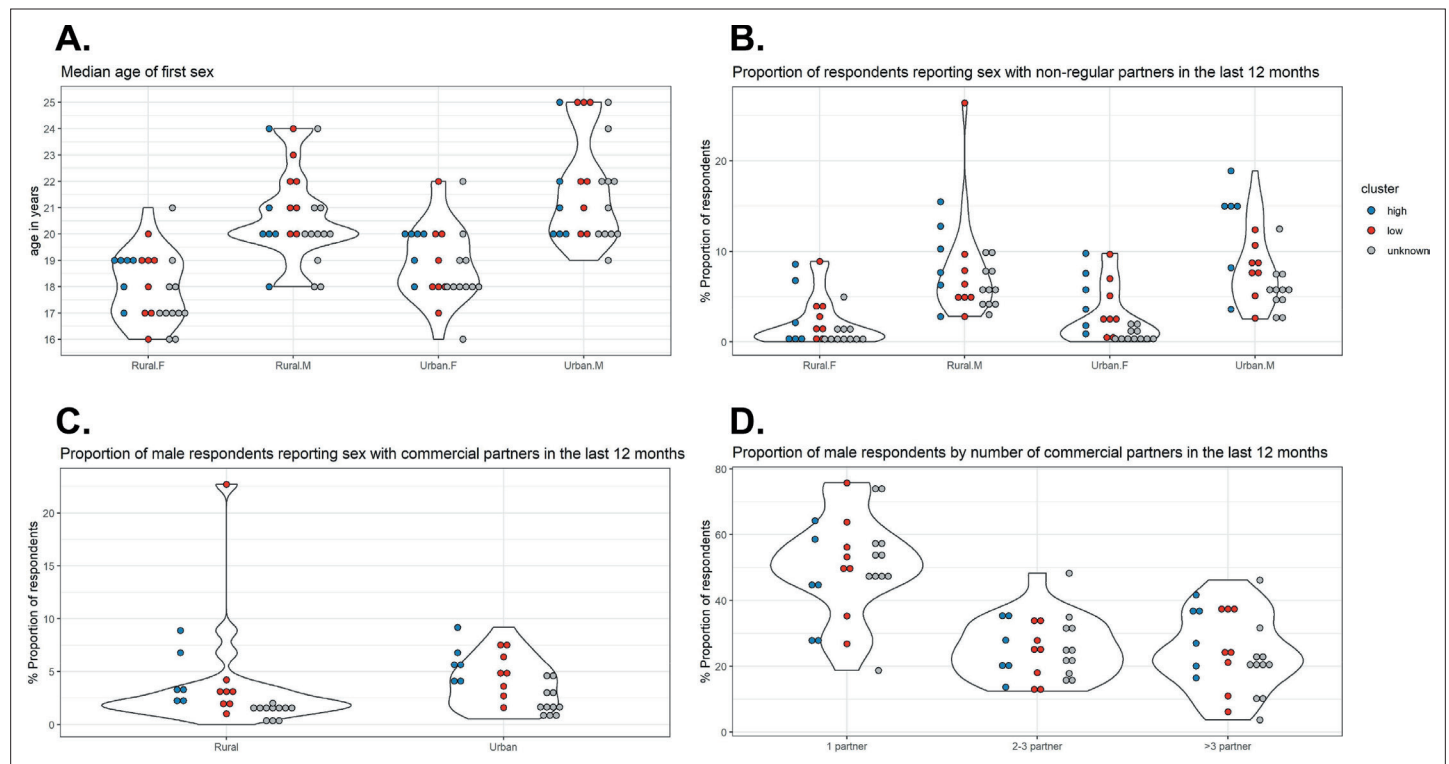


Figure 3. Sexual behaviour data from National AIDS Control Organization (NACO) by Indian state. Indian state-specific data on (A) median age of first sex, (B) proportion of respondents reporting sex with non-regular partners in the last 12 months, (C) proportion of male respondents reporting sex with commercial partners in the last 12 months, and (D) proportion of male respondents by number of commercial partners in the last 12 months. Each violin plot and the associated cloud of circles correspond to a sexual behaviour variable. Each circle corresponds to the data of a state (or group of states). The data were extracted from the 2006 National Behaviour Surveillance Survey of the National AIDS Control Organization of India (*National Behavioural Surveillance Survey: General Population, 2006*). Blue and red: Indian states identified in the high and low cervical cancer incidence clusters. Grey: states without cervical cancer incidence data and therefore unknown cluster.

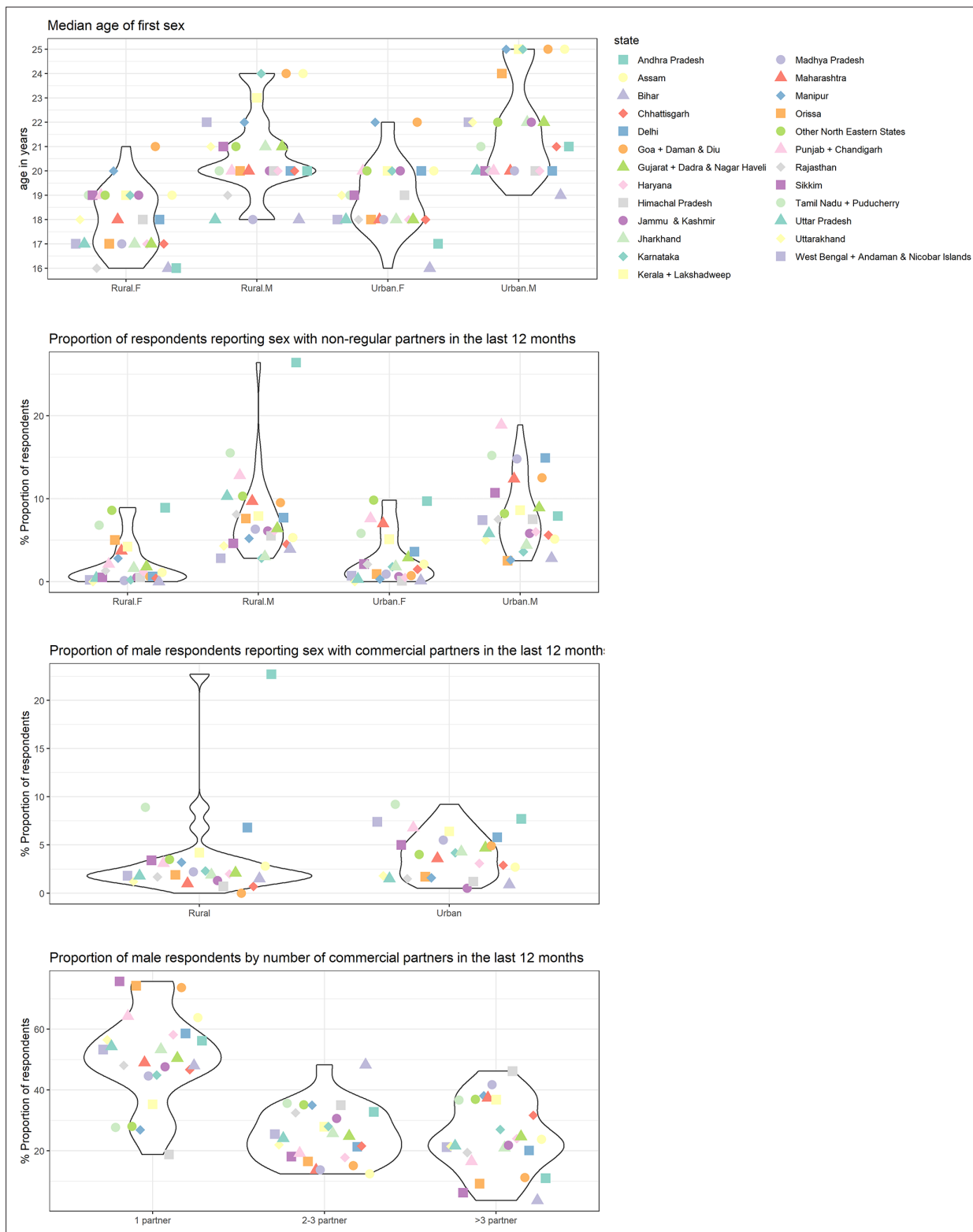


Figure 3—figure supplement 1. Indian state-specific sexual behaviour data from National AIDS Control Organization (NACO). States or groups of states as reported in the 2006 National Behaviour Surveillance Survey of the National AIDS Control Organization of India. Other North Eastern states include Arunachal Pradesh, Nagaland, Meghalaya, Mizoram, and Tripura.