WEB MATERIAL

Cancer Stage at Diagnosis, Historical Redlining, and Current Neighborhood Characteristics: Breast, Cervical, Lung, and Colorectal Cancers, Massachusetts, 2001–2015

(Version: March 5, 2020)

Nancy Krieger, Emily Wright, Jarvis T. Chen, Pamela D. Waterman, Eric R. Huntley, and Mariana Arcaya

Correspondence to Dr. Nancy Krieger, Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA 0115 (e-mail: nkrieger@hsph.harvard.edu).

TABLE OF CONTENTS

Web Appendix. Data sources and methods for generating the Home Owners' Loan Corporation (HOLC) categories

Home Owners' Loan Corporation (HOLC) data: source

Source of HOLC maps: https://dsl.richmond.edu/panorama/redlining/#loc =4/36.71/-96.93&opacity =0.8

Date of HOLC maps for the Massachusetts: 1937 and 1938

Methods for determining percent of census tract (by land area) in type of HOLC area:

A) Method using SAS

1) The Massachusetts (MA) block level shapefile (2018 Release of the 2010 Block Level Shapefiles containing Land Use Data) was downloaded from the Census website (source: https://www.census.gov/cgi-bin/geo/shapefiles2013/main) and restricted to the 28 MA municipalities for which HOLC maps were available (see Table 1)

2) In ArcGIS 10.4.1 for Desktop, the MA boundary shapefile (source: Census Tiger Boundary Files, 2010 Massachusetts Census Tracts; see: https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.2010.html) was used to clip block polygons along the coastline or other bodies of water.

3) The Feature to Point tool was used to convert block polygons to centroids with the option to place the point inside the polygon selected.

4) These points were then joined with HOLC areas maps for the 28 MA municipalities with HOLC maps (source: University of Richmond Digital Scholarship Lab, https://dsl.richmond.edu/panorama/redlining/#loc =4/36.71/-96.93&opacity =0.8) using the ArcGIS spatial join feature, thus assigning HOLC grades to individual blocks. All maps were dated 1938, except for three from 1937 (Haverhill, Chicopee, Brockton) and one map had no date visible (Holyoke).

5) Census tract HOLC composition (using the 2010 Census Tract Boundary Files) was determined in SAS by adding up the total land area (Census variable ALAND10) of blocks designated as one of four HOLC grades:

A (green): "Best"
B (blue): "Still Desirable"
C (yellow): "Definitely Declining"
D (red): "Hazardous"

6) The total land area for blocks that were either unclassified by HOLC or that did not exist when the HOLC maps were created was also tallied in SAS and designated as 'Unknown'

B) Method using R

The R Script that was used is reproduced below and is available on Github (https://gist.github.com/ericmhuntley/febc10378f73dc4c28bc49a67800d6ec). Executing this script requires, at minimum, R and the packages tigris, tidyr, dplyr, and sf. The script was run using R 3.5.3.

1. Massachusetts (MA) census tract geographies were downloaded from the U.S. Census Bureau API using the R tigris package (source: https://CRAN.R-project.org/package =tigris), converted to a simple features object, and projected using the NAD 83 / Massachusetts Mainland coordinate reference system (EPSG:2249).

2. A bounding box describing the extent of the state of Massachusetts was calculated for the downloaded census tracts.

3. The boundaries of areas mapped and graded by the HOLC intersecting the bounding box generated above were downloaded from an API endpoint maintained by the University of Richmond Digital Scholarship Lab and hosted on the CARTO platform (data source: https://dsl-ur.carto.com/u/digitalscholarshiplab/tables/holc_polygons/public/ API endpoint: https://digitalscholarshiplab.carto.com:443/api/v2/).

4. The HOLC areas were subsetted to include only those areas intersecting Massachusetts census tracts.

5. County geometries were downloaded from the U.S. Census Bureau API using the R 'tigris' package. A list of counties whose area included HOLC areas was produced.

6. Water features for those counties whose area includes HOLC areas were downloaded from the U.S. Census Bureau API.

7. These water features were 'subtracted' from census tract areas in order to exclude non-land area from our calculation of the proportion of tracts included in HOLC areas.

8. The total land area of census tracts (i.e., area excluding water features) was calculated.

9. The geometric intersection of the HOLC areas and the census tract boundaries was computed.

10. The total area included in HOLC zones corresponding to each grade was calculated on a per census tract basis:

A: "Best"
B: "Still Desirable"
C: "Definitely Declining"
D: "Hazardous"

11. The calculated areas per-census-tract corresponding to each HOLC grade were merged with the census tracts downloaded from the US census. Tracts with no land area within HOLC areas received NULL values.

```
require('tigris')
require('tidyr')
require('dplyr')
require('sf')

tracts <- tracts("MA") %>%
  st_as_sf() %>%
  mutate(cens_land = as.numeric(ALAND) * 10.76391) %>%
  select(-c(NAME, NAMELSAD, MTFCC, FUNCSTAT, ALAND, AWATER, INTPTLAT, INTPTLON)) %>%
  rename_all(tolower) %>%
  st_transform(2249)

# Create bounding box from Census tracts to narrow query results from CARTO database.
bbox <- paste(st_bbox(tracts), collapse = ",")

# Download only those HOLC areas that are within the rectangle created by the extents of census geographies.
# Uses a PostGIS query to digitalscholarshiplab CARTO API.
holc.url <- "https://digitalscholarshiplab.carto.com:443/api/v2/sql?format =GeoJSON&q ="
holc.query <- paste("SELECT holc_grade, the_geom
                    FROM digitalscholarshiplab.holc_polygons AS holc
                    WHERE ST_Intersects(holc.the_geom, ST_Transform(ST_MakeEnvelope(", bbox, ",2249), 4326))", sep ='') %>%
  URLencode(.)

# Download and project to Massachusetts Mainland (NAD 1983), EPSG: 2249.
holc <- st_read(paste(holc.url, holc.query, sep = '')) %>%
  st_transform(2249) %>%
  .[tracts,]

# Determine which counties intersect with HOLC zones
# (To constrain number of census hydrology downloads)
counties <- counties("MA") %>%
  st_as_sf() %>%
  st_transform(2249) %>%
```

```
  .[holc,] %>%
  rename_all(tolower) %>%
  st_set_geometry(NULL) %>%
  pull(countyfp) %>%
  as.numeric()

# Download water for counties with HOLC districts.
for (c in counties) {
  w <- area_water("MA", as.numeric(c)) %>%
    st_as_sf()
  if (exists('water')) {
    water <- rbind(water, w)
  } else {
    water <- w
  }
}

# Select only those census tracts that intersect HOLC zones.
# (To constrain size and complexity of unioned water polygon)
tracts_holc <- tracts[holc,]
# Project water polygon, limit to only those tracts with HOLC zones.
water <- water %>%
  st_transform(2249) %>%
  .[tracts_holc,] %>%
  st_union()

# 'Erase' water polygons.
# (This is by far the most computationally taxing step.)
diff_tracts <- st_difference(tracts_holc, water)

# Calculate total land area of census geographies.
diff_tracts$calc_area <- as.numeric(st_area(diff_tracts$geometry))

# Intersect holc polygons with census tracts.
int <- st_intersection(holc, diff_tracts)

# Calculate area of intersected polgons.
int$holc_area <- st_area(int$geometry)

# Calculate total area in redlined areas by census geographies and remove geometries (can't merge two spatial dataframes).
holc_by_tract <- int %>%
  group_by(geoid, holc_grade) %>%
  summarise(
    holc_area = as.numeric(sum(holc_area)),
    calc_area = as.numeric(max(calc_area))
  ) %>%
  st_set_geometry(NULL) %>%
  spread(holc_grade, holc_area, fill = 0)

# Merge redlined areas with original census geographies.
```

```
tracts <- merge(tracts, holc_by_tract, by = "geoid", all.x = TRUE) %>%
  st_transform(4326)

# Export to GeoJSON.
st_write(tracts, dsn = '<path>/<filename>.<format>', delete_dsn = TRUE)
```

Distribution of HOLC grades by CT:

A grade: Among the 50 CTs with any land in an A HOLC area (10.5% of the total 474 CTs for the 28 municipalities),0 (0%) were 100% A, 5 (10%) were ≥50% and <100% A (mean: 72.8%), and the remaining 45 (90%) were 0.05% to 44.3% A (mean: 13.4%).
-- Hence, among those CT assigned an A grade (≥50% A; 0 + 5 = 5, equal to 10% of the total 50 CTs), 0% were homogenous and 100% were mixed.

B grade: Among the 171 CTs with any land in a B HOLC area (36.1% of the total 474 CTs for the 28 municipalities), 1 (0.6%) was 100% B, 37 (21.6%) were ≥50% and≥ <100% B (mean: 67.3%), and the remaining 133 (77.8%) were 0.01 % to 46.2% B (mean: 16.0%).
-- Hence, among those CT assigned a B grade (≥50% B; 1 + 37 = 38, equal to 22.2% of the total 171 CTs), 2.6% were homogenous and 97.4% were mixed.

C grade: Among the 360 CTs with any land in a C HOLC area (76.0% of the total 474 CTs for the 28 municipalities), 48 (13.3%) were 100% C, 128 (35.6%) were ≥50% and <100% C (mean: 76.3%), and the remaining 184 (51.1%) were 0.01 % to 49.6% C (mean: 19.6%).
-- Hence, among those CT assigned a C grade (≥50% C; 48 + 128 = 176, equal to 48.5% of the total 363 CTs), 27.3% were homogenous and 72.7% were mixed.

D grade: Among the 194 CTs with any land in a D HOLC area (40.9% of the total 474 CTs for the 28 municipalities), 29 (15.0%) were 100% D, 49 (25.3%) were ≥50% and <100% D (mean: 74.9%), and the remaining 116 (59.8%) were 0.04 % to 48.8% D (mean: 17.1%).
-- Hence, among those CT assigned a D grade (≥50% D; 29 + 49 = 78, equal to 40% of the total 195 CTs), 37.2% were homogenous and 62.8% were mixed

HOLC area, but no grade ≥50%: Among the 39 CTs with land in a HOLC area but no HOLC grade ≥50% of this land (8.2% of the total 474 CTs for the 28 municipalities), the range of the percent of HOLC grade was as follows: **grade A:** 0% to 44.3%; **grade B:** 0% to 46.2%; **grade C:** 0% to 49.4%; and **grade D:** 0% to 48.8%.

U grade: Among the 334 CTs with any land in a U HOLC area (70.5% of the total 474 CTs for the 28 municipalities), 17 (5.1%) were 100% U, 121 (36.3%) were ≥50% and <100% U (mean: 75.9%), and the remaining 196 (58.7%) were 0.1 % to 48.9% U (mean: 20.9%).
-- Hence, among those CT assigned a U grade (≥50% U; 3 + 125 = 128, equal to 39.5% of the total 324 CTs), 2.4% were homogenous and 97.6% were mixed

Together, these data show that among those CT assigned grade A-D where the percent of land in that grade was ≥50% and <100%, the mean percent in the designated grade ranged from 67.3% for grade B to 76.3% for grade C, and the overall mean for the percent of land in the specified grade (for grades A-D) equaled 74.4%.

Web Table 1. HOLC notes for different types of parcels -- illustrative examples, including sociodemographic descriptions, Massachusetts 1937–1938 HOLC maps			
HOLC grade		**Area**	**Notes: sociodemographic description**
A (green): "Best" **-- The notes document how the HOLC grade A was reserved solely for economically prosperous areas, with no relief families, few if any "foreign-born," and no "Negroes" – i.e., exclusively white, affluent, and almost entirely US-born persons only, with no concern stated about any "undesirable" populations "infiltrating" the area**	Brookline	A1	*Clarifying remarks*: The bulk of these houses have been very recently built and the majority are priced under $25,000. As a whole the houses are well built, but there is some poor construction in the area. *Infiltration*: Desirables *Foreign-born:* none *Negro*: (not filled in) *Relief families*: None *Occupation*: executives - professional men *Estimated annual family income*: $7500 and up *Detrimental influences*: Speculation building
	Newton	A2	*Clarifying remarks:* Many large estates valued at prices not quoted are very well held, and although the houses are not new, the general character of the entire area warrants a first grade rating. New construction ranges up to $30,000 and constitutes about 20% of all the housing in this section. *Infiltration:* High class *Foreign-born:* No *Negro:* (not filled in) *Relief families:* none *Occupation:* executives - professionals *Estimated annual family income:* $10,000 and up *Detrimental influences:* nominal
	Braintree	A1	*Clarifying remarks:* New construction selling $9,500-15,000. *Infiltration:* Desirables *Foreign-born:* none *Negro:* (not filled in) *Relief families:* none *Occupation:* executives - commuters *Estimated annual family income:* $3,000 - $10,000 *Detrimental influences:* Nominal
	Haverhill	A1	*Clarifying remarks:* Two new houses are being erected on Eastland Drive, one costing about $7500, the other at the corner of Eastland Drive & Kenoza Ave. and costing about $4500. The older houses are occupied by owners and property is well kept up. *Infiltration:* Hebrew *Foreign-born:* (not filled in) *Negro:* (not filled in) *Relief families:* None *Occupation:* (not filled in) *Estimated annual family income:* $3,500 and up *Detrimental influences:* Poor condition of streets
	Arlington	A1	Clarifying comments: There is a large farm in the center of this section which is open for development when land prices improve. *Infiltration:* Desirables *Foreign-born:* none *Negro:* (not filled in) *Relief families:* none *Occupation:* business and professional *Estimated annual family income:* $2,500 to $4,000 *Detrimental influences:* Distance to grade school. Approach to the area is unattractive.

B (blue): "Still Desirable" **-- The notes document how the HOLC grade B was applied to economically stable areas with white collar workers, with few relief families, some "foreign-born," and no "Negroes" – i.e., exclusively white, economically stable, with potentially a mix of US-born and foreign born – and with little worry about any "undesirable" populations "infiltrating" the area**	Boston	B4	*Clarifying remarks:* There are a few houses along the parkway which are valued up to $60,000 although some doubt exists as to their ever selling at that price. *Infiltration:* None *Foreign-born:* (not filled in) *Negro:* (not filled in) *Relief families:* Nominal *Occupation:* Business – professional men *Estimated annual family income:* $3000-$10,000 *Detrimental influences:* Some older houses interspersed through the area
	Newton	B2	*Clarifying remarks:* The original high class owner is slowly moving out of this area and many of the large old houses are for sale at bargain prices. The smaller houses are concentrated near the Newton Line with the larger units principally below Boylston St. New construction located on and near Ellot St. *Infiltration:* Jewish *Foreign-born:* no *Negro:* (not filled in) *Relief families:* none *Occupation:* executives – professional men *Estimated annual family income:* $4000 and up *Detrimental influences:* nominal other than size of houses
	Arlington	B5	*Clarifying remarks:* Foreign infiltration will come in the next five years from the adjoining third grade area. Properties backing up to railroad tracks are considered less desirable. Infiltrating: foreign threatening *Foreign-born:* mixture *Negro:* (not filled in) *Relief families:* moderate *Occupation:* clerks *Estimated annual family income:* $1,500 - $3,000 *Detrimental influences:* Changing ownership
	Haverhill	B4	*Clarifying remarks:* About 20 years ago the area was extensively developed and good types of cottage houses were built containing 6-7-8 rooms. These houses were well built and sold at that time for about $5,000 average. The area is not being developed today but is very close to an "A" rating. Fernwood Ave is easily the best street in the neighborhood, being 100% owner occupied, and there have been fewer transfers here than any street in all of Haverhill. Infiltrating: Mixed *Foreign-born:* (not filled in) *Negro:* (not filled in) *Relief families:* very few *Occupation:* Business men & skilled workers *Estimated annual family income:* $2400 *Detrimental influences:* Slight infiltration of foreign element
	Brookline	B1	*Clarifying remarks:* The original high class owner is slowly moving out of this area and many of the large old houses are for sale at bargain prices. The smaller houses are concentrated near the Newton Line with the larger units principally below Boylston St. New construction located on and near Ellot St. *Infiltration:* Jewish *Foreign-born:* no *Negro:* (not filled in) *Relief families:* none *Occupation:* executives – professionals *Estimated annual family income:* $4,000 and up *Detrimental influences:* nominal other than size of houses.

C (yellow): "Definitely Declining" **-- The notes document how the HOLC grade C was applied to economically stable areas with both white collar and working class residents, with few relief families, some to many "foreign-born," and few if any "Negroes" – i.e., predominantly white, economically stable, with potentially a mix of US-born and foreign born – and with explicit concern about "undesirable" populations "infiltrating" the area**	Boston	C9	*Clarifying remarks:* A very conservative area with all convenience which is feeling a return of many buyers who moved into other sections. The southern end of this section has the newer singles and two-family units of smaller size and is the more active portion, although the entire area enjoys a fairly good reputation locally. *Infiltration:* Jewish threatening *Foreign-born:* nominal *Negro:* (not filled in) *Relief families:* Nominal *Occupation:* white collar class *Estimated annual family income:* $2000-$5000 *Detrimental influences:* nominal. Obsolescence.
	Malden	C2	*Clarifying remarks:* The eastern end of the section becomes poorer gradually and financing is difficult to obtain. Less desirables encroaching slowly. *Infiltration:* Jewish *Foreign-born:* Italian *Negro:* (not filled in) *Relief families:* moderate *Occupation:* clerks – skilled labor *Estimated annual family income*: $1,500 - $2,500 *Detrimental influences:* obsolescence
	Haverhill	C2	*Clarifying remarks:* The area is a part of the section of Haverhill called Bradford. The older part is west of Pine Street, the more or less undeveloped portion to the eastward. Houses to a great extent have been converted to 2-family throughout the area. Character is spotty with a few good streets. There will probably be no development here for years to come. *Infiltration:* (not filled in) *Foreign-born:* 20% *Negro:* (not filled in) *Relief families:* quite a few *Occupation:* mechanics and workers *Estimated annual family income:* 1500 *Detrimental influences:* Streets in very bad condition excepting the two man thorofares. Section is old and southeast portion undeveloped.
	Cambridge	C4	*Clarifying remarks:* Apartments on Prescott are fairly high class. Harvard St. is better than the balance of the area. Houses have value for rooming house purposes, being near Harvard University. Majority of 2 family units are converted singles. A few negro families have moved in on Dame St. and threaten to spread. *Infiltration:* lower class *Foreign-born:* Italian *Negro:* 1% *Relief families:* moderate *Occupation:* clerks – skilled labor – white collar *Estimated annual family income:* $1,200 - $3,000 *Detrimental influences:* Obsolescense. Heavy traffic through area.
	Somerville	C2	*Clarifying remarks:* The eastern end of the section becomes poorer gradually and financing is difficult to obtain. Less desirables encroaching slowly. *Infiltration:* Jewish *Foreign born:* Italian *Negro:* (not filled in) *Relief families:* moderate *Occupation:* clerks – skilled labor *Estimated annual family income:* $1,500 - $2,000 *Detrimental influences:* obsolescence

D (red): "Hazardous" **-- The notes document how the HOLC grade D was applied to economically impoverished areas, with mainly laborers and unemployed persons, many relief families, many "foreign-born," and anywhere from no to many "Negroes" – i.e., predominantly mixed, impoverished, foreign-born with often a high concentration of "Negroes," and with explicit concern about "undesirable" populations "infiltrating" the area**	Boston	D1	*Clarifying remarks:* Negro concentrated around Empire St. The better 3-family units are heated, and rentals include such. Some very poor tenement houses are scattered throughout. Originally a good section with some large houses now given over to rooming house use. *Infiltration:* foreign – negro *Foreign-born:* Italian *Negro:* 5% *Relief families:* heavy *Occupation:* labor – relief *Estimated annual family income:* $600 - $1,500 *Detrimental influences:* mixture of houses and business. Low class occupants.
		D9	*Clarifying remarks:* Negro heavily concentrated north of Ruggles St. on the west side of Washington. Jewish centered near Columbus Square. A large territory with some streets showing better experiences than the balance of the section. *Infiltration:* foreign - negro Foreign born: mixture *Negro:* 25% *Relief families:* heavy *Occupation:* clerks – labor – relief *Estimated annual family income:* $600-$1,500 *Detrimental influences:* congested. Heavy traffic. Large assessments. Unimproved property. Poor housing. Cosmopolitan population. Obsolescence.
	Everett	D1	*Clarifying remarks:* This neighborhood is a fair 4th grade section which is declining constantly in desirability as lower type buyers come into the section. *Infiltration:* Foreign *Foreign-born:* Italian *Negro:* 2% *Relief families:* fairly heavy *Occupation:* labor - industrial *Estimated annual family income:* $1,000 - $1,500 *Detrimental influences:* obsolescence. Low class occupants. Poor housing. Adjacent to dumping grounds of city of Maldea.
		D3	*Clarifying remarks:* Section south of the railroad is a slum area. Negro is concentrated around the West St. Station. The bulk of the two-family houses are converted singles. *Infiltration:* foreign Foreign-born families: Italian *Negro:* 2% *Relief families:* very heavy *Occupation:* labor - relief *Estimated annual family income:* $600 - $1,500 *Detrimental influences:* poor housing. Heavy obsolescence. Objectionable gas refining fumes.
	Malden	D1	*Clarifying remarks:* Negro is concentrated around Sherman St. This neighborhood was heavily vacant until summer of 1937 - since then vacancy has decreased considerably. *Infiltration:* Italian *Foreign-born:* Italian *Negro:* 2% *Relief families:* heavy *Occupation:* labor – relief *Estimated annual family income:* $500 - $1,500 *Detrimental influences:* obsolescence. Lower class populace. Poor housing.
		D2	*Clarifying remarks:* Foreclosures were heavy in this area with the bulk of the overhand still held by banks. The neighborhood has been largely Jewish for some time with lower class Jews moving into the section at present.

			Infiltration: Jewish *Foreign-born:* Italian *Negro:* 2% *Relief families:* very heavy *Occupation:* labor – relief *Estimated annual family income:* $600 - $1,500 *Detrimental influences:* poor housing. Lower class occupants. Adjacent to Malden City Dump.
	Cambridge	D1	*Clarifying remarks:* (not filled in) *Infiltration:* negro *Foreign-born:* Italian *Negro:* 70% *Relief families:* heavy *Occupation:* labor - relief *Estimated annual family income:* $600 - $1,500 Detrimental influences Obsolescence. Low class occupants. Poor housing. Congested area.
		D2	*Clarifying remarks:* (not filled in) *Infiltration:* negro *Foreign-born:* Italian *Negro:* 10% *Relief families:* heavy *Occupation:* labor - relief *Estimated annual family income:* $600 - $1,500 *Detrimental influences:* (not filled in)
	Haverhill	D1	*Clarifying remarks:* Main thorofare to Lawrence running through the area; is all zoned for business and industrial usage although there has been little development in this direction. Area is built up of small cheap singles. Poor planning a characteristic, and no pride of ownership. *Infiltration:* (not filled in) *Foreign born:* (not filled in) *Negro:* (not filled in) *Relief families:* many *Occupation:* mixed foreign descent *Estimated annual family income:* $700 - 1000 *Detrimental influences:* Low, subject to floods. In 1936 it was 4 to 5 feet under water
		D2	*Clarifying remarks:* One of the poor areas of the city. Has been taken over by foreign element. Property generally has little care. It is an area the banks feel quite discouraged about due to poor collections and inability to dispose of acquisitions. *Infiltration:* (not filled in) *Foreign born:* (not filled in) *Negro:* (not filled in) *Occupation:* low class factory workers *Relief families:* Many *Estimated annual family income:* 700 - 1000 *Detrimental influences:* Railroad, Little River, cemetery in northern part, very old section marked by dilapidation

Web Table 2. Distribution of incident cases of primary invasive lung, colorectal, breast and cervical cancer by cases' sociodemographic and census tract (CT) characteristics, among the population in the catchment area of the 474 CTs in the 28 Massachusetts municipalities[a] with 1930s HOLC rankings: 2001–2015

Characteristic	Cancer type[b]							
	Lung			Colorectal			Breast	Cervical
	Total	Women	Men	Total	Women	Men	Women	Women
Case characteristics								
No. of cases	18537	9561	8974	12977	6518	6457	20808	874
Age at diagnosis (continuous; year): mean (SD)	69.2 (11.1)	69.4 (11.2)	69.1 (11.0)	68.4 (14.3)	70.0 (14.8)	66.8 (13.7)	61.7 (14.4)	51.5 (15.7)
[unknown/missing: N (%)]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]
Age at diagnosis (categorical; year): %								
< 25 y	11 (0.1)	8 (0.1)	3 (0.0)	22 (0.2)	14 (0.2)	8 (0.1)	28 (0.1)	12 (1.4)
25-34 y	49 (0.3)	21 (0.2)	28 (0.3)	161 (1.2)	83 (1.3)	78 (1.2)	461 (2.2)	119 (13.6)
35-44 y	278 (1.5)	142 (1.5)	136 (1.5)	555 (4.3)	270 (4.1)	285 (4.4)	2020 (9.7)	197 (22.5)
45-54 y	1613 (8.7)	855 (8.9)	758 (8.4)	1704 (13.1)	772 (11.8)	931 (14.4)	4484 (21.5)	205 (23.5)
55-64 y	3999 (21.6)	2019 (21.1)	1978 (22.0)	2435 (18.8)	1060 (16.3)	1374 (21.3)	4944 (23.8)	147 (16.8)
65-74 y	6093 (32.9)	3111 (32.5)	2982 (33.2)	2968 (22.9)	1313 (20.1)	1655 (25.6)	4392 (21.1)	112 (12.8)
75-84 y	5238 (28.3)	2727 (28.5)	2511 (28.0)	3468 (26.7)	1926 (29.5)	1542 (23.9)	3300 (15.9)	61 (7.0)
≥85 y	1256 (6.8)	678 (7.1)	578 (6.4)	1664 (12.8)	1080 (16.6)	584 (9.0)	1179 (5.7)	21 (2.4)
[unknown/missing: N (%)][c]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]
Race/ethnicity								
non-Hispanic (NH) White	15692 (84.8)	8324 (87.2)	7366 (82.2)	10377 (80.6)	5229 (80.8)	5146 (80.3)	16710 (80.7)	495 (57.1)
NH Black	1590 (8.6)	738 (7.7)	852 (9.5)	1310 (10.2)	676 (10.4)	634 (9.9)	2057 (9.9)	161 (18.6)
NH Asian + Pacific Islander	761 (4.1)	268 (2.8)	493 (5.5)	614 (4.8)	286 (4.4)	328 (5.1)	902 (4.4)	73 (8.4)
NH American Indian + Alaska Native	16 (0.1)	14 (0.1)	2 (0.0)	10 (0.1)	6 (0.1)	4 (0.1)	17 (0.1)	1 (0.1)
Hispanic	432 (2.3)	194 (2.0)	238 (2.7)	550 (4.3)	269 (4.2)	281 (4.4)	965 (4.7)	133 (15.3)
NH other race/ethnicity	16 (0.1)	6 (0.1)	10 (0.1)	18 (0.1)	6 (0.1)	12 (0.2)	51 (0.2)	4 (0.5)
[unknown/missing: N (%)][c]	[30 (0.2)]	[17 (0.2)]	[13 (0.1)]	[98 (0.8)]	[46 (0.7)]	[52 (0.8)]	[106 (0.5)]	[7 (0.8)]
Sex/gender:								
Women	9561 (51.6)	9561 (100.0)	0 (0.0)	6518 (50.2)	6518 (100.0)	0 (0.0)	20808 (100.0)	874 (100.0)
Men	8974 (48.4)	0 (0.0)	8974 (100.0)	6457 (49.8)	0 (0.0)	6457 (100.0)	0 (0.0)	0 (0.0)
Transsexual	2 (0.0)	0 (0.0)	0 (0.0)	2 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Other	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
[unknown/missing: N (%)][c]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]
Stage:								
Local	4698 (25.3)	2670 (27.9)	2027 (22.6)	5657 (43.6)	2753 (42.2)	2903 (45.0)	14510 (69.7)	454 (51.9)
Regional	4651 (25.1)	2383 (24.9)	2267 (25.3)	4828 (37.2)	2516 (38.6)	2311 (35.8)	5372 (25.8)	311 (35.6)
Distant	9188 (49.6)	4508 (47.1)	4680 (52.2)	2492 (19.2)	1249 (19.2)	1243 (19.3)	926 (4.5)	109 (12.5)
[unknown/missing: N (%)][c]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]
Cases' residential census tract characteristics[d]								
No. of cases per CT:								
mean (SD)	39.1 (23.7)	20.2 (13.4)	18.9 (11.5)	27.4 (16.2)	13.8 (8.8)	13.6 (8.4)	43.9 (25.8)	1.8 (1.7)
median	36	18	18	25	13	12.5	41	1
min, max	0, 132	0, 80	0, 71	0, 98	0, 54	0, 45	0, 138	0, 9
[missing: N (%)][c]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]
Census tract poverty (continuous):								
mean (SD)	12.7 (9.9)	12.2 (9.6)	13.1 (10.2)	12.4 (10.0)	12.4 (10.1)	12.4 (9.9)	11.7 (9.8)	16.0 (11.7)
[missing: N (%)][c]	[24 (0.13)]	[12 (0.13)]	[12 (0.13)]	[15 (0.12)]	[7 (0.11)]	[8 (0.12)]	[17 (0.08)]	[0]
Census tract poverty (categorical):								

< 5%	3865 (20.9)	2052 (21.5)	1813 (20.2)	3028 (23.4)	1535 (23.6)	1493 (23.2)	5439 (26.2)	112 (12.8)
5 - <10%%	5345 (28.9)	2888 (30.2)	2455 (27.4)	3711 (28.6)	1907 (29.3)	1803 (28.0)	5981 (28.8)	202 (23.1)
10 - <20%	5979 (32.3)	3045 (31.9)	2934 (32.7)	3925 (30.3)	1915 (29.4)	2010 (31.2)	6187 (29.8)	313 (35.8)
≥ 20%	3324 (18.0)	1564 (16.4)	1760 (19.6)	2298 (17.7)	1154 (17.7)	1143 (17.7)	3184 (15.3)	247 (28.3)
[missing: N (%)][c]	[24 (0.13)]	[12 (0.13)]	[12 (0.13)]	[15 (0.12)]	[7 (0.11)]	[8 (0.12)]	[17 (0.08)]	[0]
ICE for homeownership (continuous):								
mean (SD)	0.04 (0.44)	0.06 (0.44)	0.02 (0.45)	0.05 (0.45)	0.05 (0.45)	0.05 (0.45)	0.09 (0.45)	-0.10 (0.43)
[missing: N (%)][c]	[24 (0.13)]	[12 (0.13)]	[12 (0.13)]	[15 (0.12)]	[7 (0.11)]	[8 (0.12)]	[17 (0.08)]	[0]
ICE for income (continuous):								
mean (SD)	0.12 (0.25)	0.13 (0.25)	0.11 (0.25)	0.13 (0.26)	0.13 (0.26)	0.13 (0.25)	0.17 (0.26)	0.04 (0.27)
[missing: N (%)][c]	[24 (0.13)]	[12 (0.13)]	[12 (0.13)]	[15 (0.12)]	[7 (0.11)]	[8 (0.12)]	[17 (0.08)]	[0]
ICE for race/ethnicity (continuous):								
mean (SD)	0.56 (0.38)	0.58 (0.37)	0.54 (0.39)	0.56 (0.40)	0.56 (0.40)	0.56 (0.40)	0.58 (0.39)	0.40 (0.46)
[missing: N (%)][c]	[24 (0.13)]	[12 (0.13)]	[12 (0.13)]	[15 (0.12)]	[7 (0.11)]	[8 (0.12)]	[17 (0.08)]	[0]
ICE for race/ethnicity + income (continuous):								
mean (SD)	0.22 (0.19)	0.23 (0.19)	0.21 (0.19)	0.23 (0.20)	0.23 (0.20)	0.23 (0.20)	0.25 (0.21)	0.15 (0.21)
[missing: N (%)][c]	[24 (0.13)]	[12 (0.13)]	[12 (0.13)]	[15 (0.12)]	[7 (0.11)]	[8 (0.12)]	[17 (0.08)]	[0]
HOLC ranking (categorical):[e]								
Green	198 (1.1)	113 (1.2)	85 (0.9)	184 (1.4)	100 (1.5)	85 (1.3)	315 (1.5)	6 (0.7)
Blue	1359 (7.3)	739 (7.7)	620 (6.9)	1064 (8.2)	548 (8.4)	516 (8.0)	2029 (9.8)	59 (6.8)
Yellow	7126 (38.4)	3612 (37.8)	3514 (39.2)	4928 (38.0)	2481 (38.1)	2447 (37.9)	7534 (36.2)	371 (42.4)
Red	2536 (13.7)	1242 (13.0)	1294 (14.4)	1607 (12.4)	764 (11.7)	842 (13.0)	2331 (11.2)	139 (15.9)
Mixed, no category ≥ 50%	1859 (10.0)	983 (10.3)	876 (9.8)	1333 (10.3)	695 (10.7)	637 (9.9)	2342 (11.3)	77 (8.8)
No grade assigned	5459 (29.4)	2872 (30.0)	2585 (28.8)	3860 (29.7)	1930 (29.6)	1930 (29.9)	6257 (30.1)	222 (25.4)
[missing: N (%)][c]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]

Notes:

[a] Cases resided, at time of diagnosis, within CTs in the following 28 MA municipalities with HOLC maps: Arlington, Belmont, Boston, Braintree, Brockton, Brookline, Cambridge, Chelsea, Chicopee, Dedham, Everett, Haverhill, Holyoke, Lexington, Malden, Medford, Melrose, Milton, Needham, Newton, Quincy, Revere, Saugus, Somerville, Waltham, Watertown, Winchester, Winthrop.

[b] Cases were categorized using ICD-O-3 codes C500-506, C508-509 (breast cancer); C530-531, C538-539 (cervical cancer); C180-189, C199, C209, C260 (colorectal cancer); C340-C349 (lung cancer).

[c] Percent missing based on total; otherwise, distributions are based on observed cases only.

[d] CT ICE and poverty measures are based on ACS 5-year annual estimates centered in the years 2008 (2006-2010), 2009 (2007-2011), 2010 (2008-2012), 2011 (2009-2013), 2012 (2010-2014), 2013 (2011-2015), 2014 (2012-2016), and 2015 (2013-2017). CT ICE and poverty measures for the years 2001-2007 are based on linear interpolation using decennial census data for the year 2000 and census and the ACS 5-year estimate for 2008 (2006-2010) as anchors. All decennial census and ACS data are normalized to 2010 CT boundaries.

[e] Operational definition: “Green” = census tracts whose land area is [(100% A OR (≥ 50% A and <100% A)]; “Blue” = census tracts whose land area is [100% B OR (≥ 50% B and <100% B)]; “Yellow” = census tracts whose land area is [(100% C) OR (≥ 50% C and <100%C)]; “Red” = census tracts whose land is [(100% D) OR (≥ 50% D and <100% D)]; “Mixed, no category ≥50%” = census tracts with ≥ 50% of land area in areas with HOLC grades but with no HOLC grade accounting for ≥50% of the total land area; and “No grade assigned” = census tracts whose land area is ≥ 50% unknown (in relation to HOLC grade).

Web Table 3. Data sources and methods for generating the census tract characteristics, for Index of Concentration at the Extremes and the poverty level

Index of Concentration at the Extremes (ICE), using US decennial census data (2000) and American Community Survey (ACS) data (5 year estimate, for 2008 through 2015), using 2010 census tract normalized boundaries

The formula for the Index of Concentration at the Extremes (ICE) is as follows:

$ICE_i = (A_i - P_i)/T_i$

where A_i, P_i and T_i correspond, respectively, to the number of persons in the ith geographic area who are categorized as belonging to: the most privileged extreme, the most deprived extreme, and the total population whose privilege level was measured. For example, for the ICE for racialized economic segregation, A_i = number of non-Hispanic white persons in the top income households (80th percentile) in neighborhood i; P_i = number of non-Hispanic black persons in the bottom income households (20th percentile) in neighborhood i; and T_i = total population across all income percentiles in neighborhood i. The ICE thus ranges from -1 to 1, delineating areas in which 100% of the population is in the most extreme group for deprivation to 100% in the most extreme group for privilege.

Components	ICE measure			
	Racialized economic segregation (ICE for race/ethnicity + income)	**Household income (ICE for income)**	**Race/ethnicity (ICE for race/ethnicity)**	**Home ownership (ICE for home ownership)**
Privileged group	Non-Hispanic white high-income households (upper 80th percentile for US household income)	High-income households (upper 80th percentile for US household income)	Non-Hispanic white	Owner-occupied housing units
Deprived group	Non-Hispanic black low-income households (bottom 20th percentile for US household income)	Low-income households (bottom 20th percentile for US household income)	Non-Hispanic black	Renter-occupied housing units
Census source:	**ICE formula, with census variable categories**			
Census 2000	P151A_(013 + 014 + 015 + 016 + 017) - P151B_(002 + 003 + 004)/ P052001	P052_(013 + 014 + 015 + 016 + 017) - (002 + 003 + 004)/ P052001	P007_(003 - 004)/ P006001	H011_(002 - 010)/ H011001
ACS 2006-2010 (2008)	B19001A_(VD14 + VD15 + VD16 + VD17) - B19001B_(VD02 + VD03 + VD04 + VD05)/ B19001(VD01)	B19001_(VD14 + VD15 + VD16 + VD17) - (VD02 + VD03 + VD04)/ B19001(VD01)	B03002_(VD03) - (VD04)/ B03002(VD01)	B25003_(VD02) - (VD03)/ B25003(VD01)
ACS 2007-2011 (2009)	B19001A_(VD14 + VD15 + VD16 + VD17) - B19001B_(VD02 + VD03 + VD04 + VD05)/ B19001(VD01)	B19001_(VD14 + VD15 + VD16 + VD17) - (VD02 + VD03 + VD04)/ B19001(VD01)	B03002_(VD03) - (VD04)/ B03002(VD01)	B25003_(VD02) - (VD03)/ B25003(VD01)
ACS 2008-2012 (2010)	B19001A_(VD14 + VD15 + VD16 + VD17) - B19001B_(VD02 + VD03 + VD04 + VD05)/ B19001(VD01)	B19001_(VD14 + VD15 + VD16 + VD17) - (VD02 + VD03 + VD04)/ B19001(VD01)	B03002_(VD03) - (VD04)/ B03002(VD01)	B25003_(VD02) - (VD03)/ B25003(VD01)
ACS 2009-2013 (2011)	B19001A_(VD14 + VD15 + VD16 + VD17) - B19001B_(VD02 + VD03 + VD04 + VD05)/ B19001(VD01)	B19001_(VD14 + VD15 + VD16 + VD17) - (VD02 + VD03 + VD04)/ B19001(VD01)	B03002_(VD03) - (VD04)/ B03002(VD01)	B25003_(VD02) - (VD03)/ B25003(VD01)
ACS 2010-2014 (2012)	B19001A_(VD14 + VD15 + VD16 + VD17) - B19001B_(VD02 + VD03 + VD04 + VD05)/ B19001(VD01)	B19001_(VD14 + VD15 + VD16 + VD17) - (VD02 + VD03 + VD04)/ B19001(VD01)	B03002_(VD03) - (VD04)/ B03002(VD01)	B25003_(VD02) - (VD03)/ B25003(VD01)
ACS 2011-2015 (2013)	B19001A_(VD14 + VD15 + VD16 + VD17) - B19001B_(VD02 + VD03 + VD04 + VD05)/ B19001(VD01)	B19001_(VD14 + VD15 + VD16 + VD17) - (VD02 + VD03 + VD04)/ B19001(VD01)	B03002_(VD03) - (VD04)/ B03002(VD01)	B25003_(VD02) - (VD03)/ B25003(VD01)
ACS 2012-2016 (2014)	B19001A_(VD14 + VD15 + VD16 + VD17) - B19001B_(VD02 + VD03 + VD04 + VD05)/ B19001(VD01)	B19001_(VD14 + VD15 + VD16 + VD17) - (VD02 + VD03 + VD04)/ B19001(VD01)	B03002_(VD03) - (VD04)/ B03002(VD01)	B25003_(VD02) - (VD03)/ B25003(VD01)

ASC 2013-2017 (2015)	B19001A_(VD14 + VD15 + VD16 + VD17) - B19001B_(VD02 + VD03 + VD04 + VD05)/ B19001(VD01)		B19001_(VD14 + VD15 + VD16 + VD17) - (VD02 + VD03 + VD04)/ B19001(VD01)		B03002_(VD03) - (VD04)/ B03002(VD01)		B25003_(VD02) - (VD03)/ B25003(VD01)	
Tercile cutpoints: for MA distribution	**T1 (highest) – T2**	**T2 – T3 (lowest)**	**T1 (highest) – T2**	**T2 – T3 (lowest)**	**T1 (highest) – T2**	**T2 – T3 (lowest)**	**T1 (highest) – T2**	**T2 – T3 (lowest)**
2001–2015 (based on observed and interpolated annual values for entire study period)	0.34	0.17	0.26	0.03	0.90	0.70	0.57	0.00
Census 2000	0.33	0.17	0.22	-0.01	0.93	0.78	0.55	-0.05
ACS 2006-2010 (2008)	0.32	0.15	0.24	0.02	0.90	0.69	0.59	0.03
ACS 2007-2011 (2009)	0.33	0.16	0.26	0.03	0.90	0.69	0.59	0.03
ACS 2008-2012 (2010)	0.34	0.17	0.27	0.05	0.89	0.69	0.58	0.00
ACS 2009-2013 (2011)	0.35	0.17	0.28	0.04	0.88	0.68	0.56	0.00
ACS 2010-2014 (2012)	0.36	0.17	0.29	0.06	0.88	0.67	0.56	-0.01
ACS 2011-2015 (2013)	0.36	0.18	0.30	0.07	0.87	0.65	0.56	-0.01
ACS 2012-2016 (2014)	0.37	0.18	0.32	0.08	0.86	0.64	0.56	-0.01
ASC 2013-2017 (2015)	0.39	0.20	0.35	0.12	0.86	0.63	0.57	0.00

Census tract poverty: percent of persons below the federal poverty line (2010 normalized boundaries)

Census source	**Census poverty variables**	
	Persons below poverty	**Denominator (all persons for whom poverty level ascertained)**
2000 decennial census	P087002	P087001
ACS 2006-2010 (2008)	B17001_(VD02)	B17001_(VD01)
ACS 2007-2011 (2009)	B17001_(VD02)	B17001_(VD01)
ACS 2008-2012 (2010)	B17001_(VD02)	B17001_(VD01)
ACS 2009-2013 (2011)	B17001_(VD02)	B17001_(VD01)
ACS 2010-2014 (2012)	B17001_(VD02)	B17001_(VD01)
ACS 2011-2015 (2013)	B17001_(VD02)	B17001_(VD01)
ACS 2012-2016 (2014)	B17001_(VD02)	B17001_(VD01)
ASC 2013-2017 (2015)	B17001_(VD02)	B17001_(VD01)

Note:
1) CT ICE and poverty measures for the years 2001-2007 are based on linear interpolation using 2000 decennial census and the ACS 5-year estimate for 2008 (2006-2010) as anchors.
2) Tercile cut-points are based on the full distribution of CTs for MA, and the cut-points for the 2001–2015 terciles (corresponding to the entire study period) are based on all MA CT values for this span of years (justified by similarity of tercile cut-points in each observed time point)

Web Table 4. Massachusetts census tract (CT) characteristics for the 28 Massachusetts municipalities with 1930s Home Owners' Loan Corporation (HOLC) maps and for the Massachusetts cities and towns without these HOLC maps (*n* = 1004), selected years, 2001–2015[a]

CT Characteristic[b]	Time Period							
	2001–2015		2000		2010		2015	
	HOLC	No HOLC	HOLC	No HOLC	HOLC	No HOLC	HOLC	No HOLC
No. of CTs	474	1004	474	1004	474	1004	474	1004
% of MA population	29.6	70.4	29.5	70.5	29.5	70.5	30.0	70.0
Census tract poverty (continuous):								
mean (SD)	15.4 (12.7)	10.2 (9.9)	13.7 (11.2)	8.9 (9.2)	16.1 (13.6)	10.7 (10.8)	14.9 (12.2)	11.0 (10.6)
[missing: N (%)][c]	[5 (1.1)]	[9 (0.90)]	[0]	[2 (0.20)]	[8 (1.7)]	[10 (1.0)]	[6 (1.3)]	[9 (0.90)]
Census tract poverty (categorical):								
< 5%	80 (17.1)	372 (37.4)	102 (21.5)	451 (45.0)	81 (17.4)	366 (36.8)	92 (19.7)	335 (33.7)
5 - <10%	120 (25.6)	325 (32.7)	135 (28.5)	290 (28.9)	110 (23.6)	301 (30.3)	109 (23.3)	311 (31.3)
10 - <20%	146 (31.1)	150 (15.1)	135 (28.5)	150 (15.0)	140 (30.0)	174 (17.5)	144 (30.8)	193 (19.4)
≥ 20%	123 (26.2)	148 (14.9)	102 (21.5)	111 (11.1)	135 (29.0)	153 (15.4)	123 (26.3)	156 (15.7)
[missing: N (%)][c]	[5 (1.1)]	[9 (0.90)]	[0]	[2 (0.20)]	[8 (1.7)]	[10 (1.0)]	[6 (1.3)]	[9 (0.90)]
ICE for home ownership (continuous):								
mean (SD)	-0.07 (0.46)	0.36 (0.46)	-0.10 (0.46)	0.35 (0.46)	-0.06 (0.47)	0.36 (0.47)	-0.06 (0.47)	0.34 (0.48)
[missing: N (%)][c]	[6 (1.3)]	[9 (0.90)]	[0]	[2 (0.20)]	[9 (1.9)]	[10 (1.0)]	[8 (1.7)]	[9 (0.90)]
ICE for income (continuous):								
mean (SD)	0.10 (0.27)	0.15 (0.27)	0.05 (0.27)	0.11 (0.29)	0.10 (0.28)	0.16 (0.27)	0.20 (0.28)	0.21 (0.27)
[missing: N (%)][c]	[6 (1.3)]	[9 (0.90)]	[1 (0.21)]	[2 (0.20)]	[9 (1.9)]	[10 (1.0)]	[8 (1.7)]	[9 (0.90)]
ICE for race/ethnicity (continuous):								
mean (SD)	0.50 (0.42)	0.79 (0.24)	0.57 (0.43)	0.84 (0.21)	0.49 (0.42)	0.77 (0.25)	0.45 (0.41)	0.74 (0.26)
[missing: N (%)][c]	[4 (0.84)]	[7 (0.70)]	[0]	[2 (0.20)]	[7 (1.5)]	[7 (0.70)]	[5 (1.1)]	[7 (0.70)]
ICE for race/ethnicity + income (continuous):								
mean (SD)	0.20 (0.21)	0.28 (0.17)	0.20 (0.21)	0.28 (0.21)	0.20 (0.22)	0.28 (0.18)	0.24 (0.22)	0.31 (0.18)
[missing: N (%)][c]	[6 (1.3)]	[6 (0.60)]	[1 (0.21)]	[2 (0.20)]	[9 (1.9)]	[10 (1.0)]	[8 (1.7)]	[9 (0.90)]
HOLC grade (categorical)[d]:								
Green	5 (1.1)	-	5 (1.1)	-	5 (1.1)	-	5 (1.1)	-
Blue	38 (8.0)	-	38 (8.0)	-	38 (8.0)	-	38 (8.0)	-
Yellow	176 (37.1)	-	176 (37.1)	-	176 (37.1)	-	176 (37.1)	-
Red	78 (16.5)	-	78 (16.5)	-	78 (16.5)	-	78 (16.5)	-
Mixed, no category ≥ 50%	39 (8.2)	-	39 (8.2)	-	39 (8.2)	-	39 (8.2)	-
No grade assigned	138 (29.1)	-	138 (29.1)	-	138 (29.1)	-	138 (29.1)	-
[missing: N (%)][c]	[0]		[0]		[0]		[0]	

Notes:

[a] The following 28 MA municipalities had HOLC maps and together contained 474 CTs (2010 normalized boundaries): Arlington, Belmont, Boston, Braintree, Brockton, Brookline, Cambridge, Chelsea, Chicopee, Dedham, Everett, Haverhill, Holyoke, Lexington, Malden, Medford, Melrose, Milton, Needham, Newton, Quincy, Revere, Saugus, Somerville, Waltham, Watertown, Winchester, Winthrop. All maps were dated 1938, except for three from 1937 (Haverhill, Chicopee, Brockton) and one map had no date visible (Holyoke).

[b] CT ICE and poverty measures are based on decennial census data for the year 2000 and ACS 5-year annual estimates centered in the years 2008 (2006-2010), 2009 (2007-2011), 2010 (2008-2012), 2011 (2009-2013), 2012 (2010-2014), 2013 (2011-2015), 2014 (2012-2016), and 2015 (2013-2017). CT ICE and poverty measures for the years 2001-2007 are based on linear interpolation using 2000 decennial census and the ACS 5-year estimate for 2008 (2006-2010) as anchors.

[c] Percent missing based on total; otherwise, distributions are based on observed cases only.

[d] Operational definition: "Green" = census tracts whose land area is [(100% A OR (≥ 50% A and <100% A)]; "Blue" = census tracts whose land area is [100% B OR (≥ 50% B and <100% B)]; "Yellow" = census tracts whose land area is [(100% C) OR (≥ 50% C and <100%C)]; "Red" = census tracts whose land is [(100% D) OR (≥ 50% D and <100% D)]; "Mixed, no category ≥50%" = census tracts with ≥ 50% of land area in areas with HOLC grades but with no HOLC grade accounting for ≥50% of the total land area; and "No grade assigned" = census tracts whose land area is ≥ 50% unknown (in relation to HOLC grade).

Web Table 5. Correlations among census tract (CT) characteristics (2001–2015) for the 28 Massachusetts municipalities with the 1937–1938 HOLC maps					
	Pearson correlation coefficient (all p-values <0.001; 2-sided tests)				
	Variable				
Variable	Percent below poverty	ICE for race/ ethnicity	ICE for income	ICE for race/ ethnicity + income	ICE for home ownership
Percent below poverty	1.000				
ICE for race/ethnicity	-0.570	1.000			
ICE for income	-0.840	0.653	1.000		
ICE for race/ethnicity + income	-0.737	0.848	0.914	1.000	
ICE for home ownership	-0.770	0.479	0.749	0.652	1.000
Note: correlation matrices for the individual years available upon request (Decennial 2000 census data; ACS 5-year estimates for 2008 through 2015)					

Web Table 6. Univariable and multivariable analysis of age-standardized risk of cancer stage at diagnosis, for late (regional + distant) versus early stage, by cases' current sociodemographic and census tract (CT) characteristics at time of diagnosis, and HOLC grade: Massachusetts Cancer Registry data, primary invasive cases of lung, colorectal, breast, and cervical cancer, 2001–2015, for the 28 municipalities with HOLC rankings

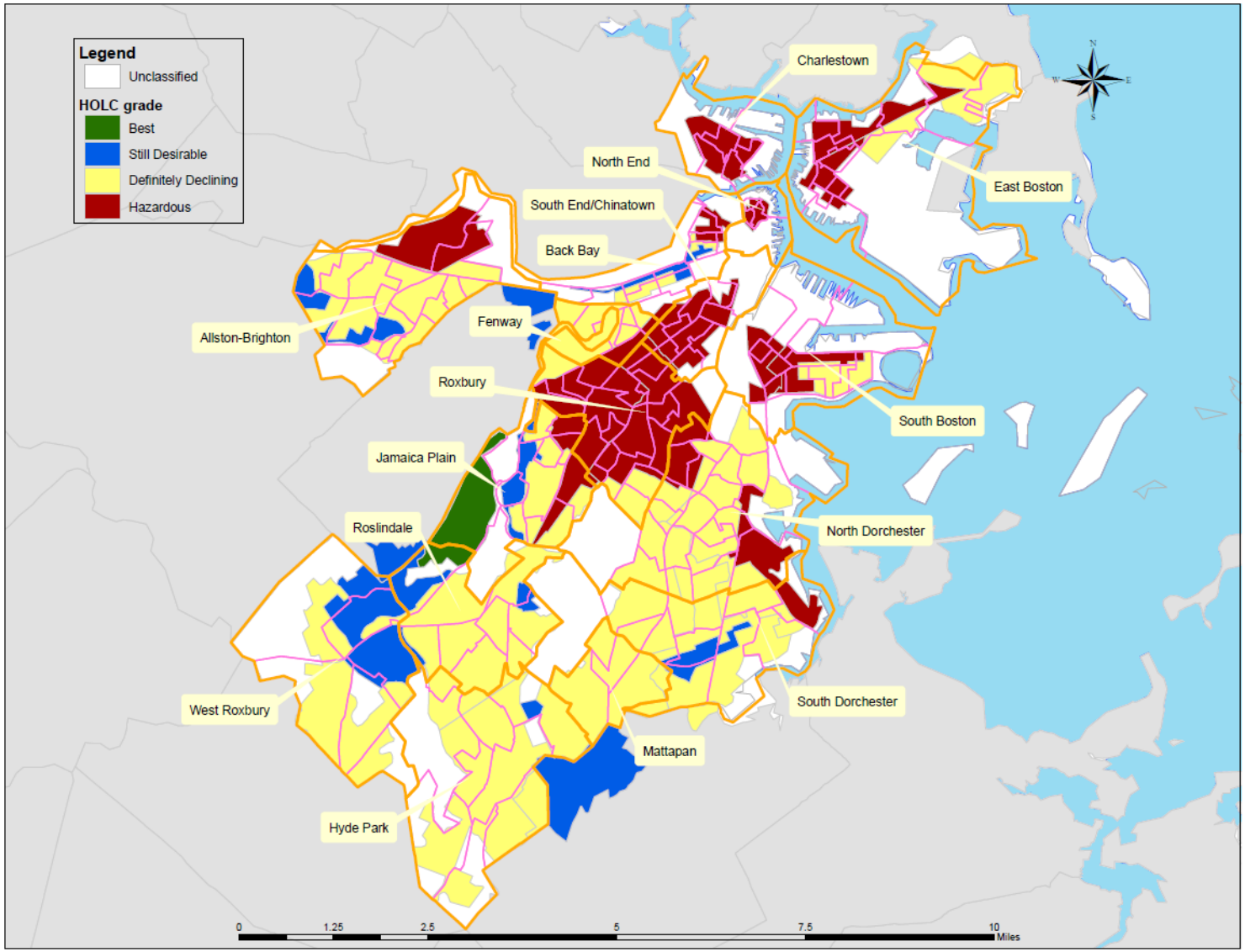
Characteristic	Risk Ratio (RR) and 95% confidence interval (CI)							
	Lung			Colorectal			Breast	Cervical
	Total	Women	Men	Total	Women	Men	Women	Women
	RR (95% CI)	RR (95% CI)	RR (95% CI)	RR (95% CI)	RR (95% CI)	RR (95% CI)	RR (95% CI)	RR (95% CI)
UNIVARIABLE								
Case characteristics								
Sex/gender								
Women (reference)	1.00	-	-	1.00	-	-	-	-
Men	1.07 (1.06, 1.09)	-	-	0.96 (0.93, 0.99)	-	-	-	-
Race/ethnicity								
Non-Hispanic (NH) White (reference)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
NH Black	1.01 (0.98, 1.05)	0.99 (0.94, 1.03)	1.03 (0.99, 1.07)	1.02 (0.97, 1.07)	1.00 (0.94, 1.07)	1.04 (0.96, 1.11)	1.26 (1.18, 1.34)	1.00 (0.84, 1.20)
NH Asian + Pacific Islander	1.06 (1.01, 1.10)	0.99 (0.92, 1.07)	1.07 (1.02, 1.12)	0.96 (0.89, 1.03)	0.97 (0.87, 1.07)	0.95 (0.85, 1.05)	0.97 (0.88, 1.07)	0.82 (0.61, 1.07)
Hispanic	1.03 (0.98, 1.09)	0.99 (0.90, 1.08)	1.06 (0.99, 1.13)	1.00 (0.93, 1.08)	0.95 (0.85, 1.05)	1.06 (0.95, 1.17)	1.05 (0.95, 1.15)	0.90 (0.72, 1.10)
Cases' residential CT characteristics								
HOLC ranking								
Green + blue (best off; reference)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Yellow	1.04 (1.01, 1.07)	1.02 (0.97, 1.07)	1.06 (1.01, 1.11)	1.01 (0.96, 1.07)	1.00 (0.93, 1.08)	1.03 (0.95, 1.11)	1.09 (1.02, 1.17)	0.90 (0.69, 1.18)
Red	1.04 (1.00, 1.08)	1.00 (0.95, 1.06)	1.08 (1.02, 1.13)	1.02 (0.96, 1.09)	0.99 (0.90, 1.08)	1.06 (0.96, 1.16)	1.11 (1.02, 1.21)	0.96 (0.71, 1.30)
Mixed (no area ≥50%)	1.02 (0.98, 1.06)	1.01 (0.96, 1.07)	1.03 (0.98, 1.09)	0.97 (0.91, 1.04)	1.00 (0.91, 1.10)	0.94 (0.85, 1.05)	1.01 (0.93, 1.11)	0.76 (0.54, 1.08)
No grade assigned	1.02 (0.99, 1.05)	0.99 (0.94, 1.04)	1.05 (1.00, 1.10)	1.00 (0.95, 1.06)	0.98 (0.91, 1.06)	1.02 (0.94, 1.11)	1.03 (0.96, 1.11)	0.87 (0.66, 1.17)
ICE for home ownership								
T1 (best off; reference)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
T2	0.99 (0.96, 1.01)	1.00 (0.96, 1.04)	0.97 (0.93, 1.00)	1.02 (0.98, 1.07)	0.98 (0.92, 1.04)	1.07 (1.00, 1.14)	1.08 (1.02, 1.15)	1.10 (0.84, 1.46)
T3	1.01 (0.99, 1.04)	1.02 (0.98, 1.06)	0.99 (0.96, 1.03)	1.05 (1.01, 1.10)	1.00 (0.94, 1.06)	1.10 (1.03, 1.17)	1.18 (1.11, 1.25)	1.11 (0.86, 1.45)
ICE for income								
T1 (best off; reference)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
T2	1.04 (1.02, 1.06)	1.05 (1.02, 1.09)	1.02 (0.99, 1.05)	1.02 (0.98, 1.06)	1.02 (0.97, 1.07)	1.01 (0.96, 1.07)	1.11 (1.06, 1.17)	1.25 (1.02, 1.54)
T3	1.05 (1.03, 1.07)	1.05 (1.02, 1.09)	1.04 (1.01, 1.07)	1.04 (1.00, 1.08)	1.03 (0.98, 1.09)	1.05 (0.99, 1.11)	1.22 (1.16, 1.28)	1.22 (1.01, 1.47)
ICE for race/ethnicity								
T1 (best off; reference)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
T2	1.00 (0.97, 1.03)	1.00 (0.96, 1.05)	0.99 (0.95, 1.03)	1.05 (0.99, 1.10)	1.00 (0.93, 1.08)	1.09 (1.01, 1.18)	0.92 (0.86, 1.00)	0.98 (0.73, 1.35)
T3	1.00 (0.97, 1.03)	0.98 (0.94, 1.03)	1.01 (0.97, 1.05)	1.05 (1.00, 1.11)	1.01 (0.94, 1.08)	1.09 (1.01, 1.18)	1.05 (0.98, 1.13)	1.14 (0.87, 1.53)
ICE for race/ethnicity + income								
T1 (best off; reference)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
T2	1.04 (1.02, 1.07)	1.04 (1.00, 1.07)	1.04 (1.01, 1.08)	1.02 (0.98, 1.06)	1.04 (0.98, 1.10)	0.99 (0.94, 1.05)	1.12 (1.06, 1.18)	1.26 (1.01, 1.58)
T3	1.05 (1.02, 1.07)	1.04 (1.01, 1.07)	1.05 (1.02, 1.08)	1.04 (1.00, 1.08)	1.03 (0.98, 1.09)	1.04 (0.99, 1.10)	1.20 (1.15, 1.27)	1.25 (1.03, 1.53)
Percent below poverty								
< 5% (best off; reference)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
5 - < 10%	1.02 (1.00, 1.05)	1.03 (0.99, 1.06)	1.02 (0.99, 1.06)	1.03 (0.99, 1.08)	1.04 (0.98, 1.10)	1.02 (0.96, 1.09)	1.07 (1.01, 1.14)	1.13 (0.88, 1.46)
10% - <20%	1.02 (0.99, 1.04)	1.01 (0.98, 1.05)	1.02 (0.99, 1.05)	1.02 (0.98, 1.06)	1.04 (0.98, 1.10)	0.99 (0.93, 1.06)	1.12 (1.06, 1.19)	1.19 (0.94, 1.51)
≥ 20%	1.05 (1.02, 1.08)	1.05 (1.00, 1.09)	1.04 (1.00, 1.08)	1.06 (1.01, 1.11)	1.03 (0.97, 1.10)	1.09 (1.02, 1.17)	1.19 (1.11, 1.27)	1.23 (0.97, 1.57)

MULTIVARIABLE									
HOLC ranking: total effect (sex/gender and race/ethnicity-adjusted)[a]									
HOLC ranking									
Green + Blue (best off; reference)		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Yellow		1.04 (1.00, 1.07)	1.02 (0.97, 1.07)	1.06 (1.01, 1.11)	1.01 (0.96, 1.07)	1.00 (0.93, 1.08)	1.03 (0.95, 1.11)	1.07 (0.99, 1.15)	0.90 (0.69, 1.19)
Red		1.03 (1.00, 1.07)	1.00 (0.95, 1.06)	1.07 (1.02, 1.13)	1.02 (0.96, 1.09)	0.99 (0.91, 1.09)	1.05 (0.96, 1.16)	1.07 (0.98, 1.17)	0.98 (0.73, 1.33)
Mixed HOLC grades		1.02 (0.98, 1.06)	1.01 (0.96, 1.07)	1.03 (0.98, 1.09)	0.97 (0.91, 1.04)	1.00 (0.91, 1.10)	0.94 (0.85, 1.04)	1.02 (0.93, 1.11)	0.76 (0.54, 1.09)
No grade assigned		1.02 (0.98, 1.05)	0.99 (0.94, 1.04)	1.05 (1.00, 1.10)	1.00 (0.95, 1.06)	0.98 (0.91, 1.06)	1.02 (0.94, 1.11)	1.03 (0.96, 1.11)	0.88 (0.66, 1.17)
Interaction of HOLC ranking and ICE tercile for racialized economic segregation (race/ethnicity and sex/gender-adjusted)									
HOLC x ICE tercile for racialized economic segregation									
HOLC	**ICE**								
Green + blue (best off; reference)	T1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Yellow	T1	1.09 (1.03, 1.15)	1.10 (1.01, 1.18)	1.09 (1.01, 1.17)	1.02 (0.93, 1.12)	0.96 (0.85, 1.09)	1.08 (0.94, 1.24)	1.09 (0.97, 1.22)	0.76 (0.43, 1.32)
Red	T1	1.08 (1.00, 1.16)	1.00 (0.89, 1.12)	1.17 (1.06, 1.29)	0.95 (0.81, 1.10)	0.97 (0.79, 1.19)	0.92 (0.73, 1.15)	1.15 (0.95, 1.39)	0.57 (0.21, 1.29)
Mixed	T1	1.03 (0.98, 1.09)	1.02 (0.94, 1.11)	1.04 (0.96, 1.13)	0.99 (0.91, 1.09)	0.95 (0.84, 1.08)	1.04 (0.90, 1.19)	0.98 (0.87, 1.09)	0.89 (0.47, 1.61)
No grade	T1	1.04 (0.99, 1.09)	1.02 (0.95, 1.09)	1.06 (0.99, 1.13)	1.03 (0.96, 1.11)	0.95 (0.85, 1.05)	1.12 (1.01, 1.26)	1.04 (0.94, 1.14)	0.75 (0.48, 1.18)
Green + blue	T2+T3	1.12 (1.05, 1.19)	1.14 (1.04, 1.24)	1.10 (1.01, 1.19)	1.08 (0.97, 1.19)	0.99 (0.86, 1.13)	1.19 (1.02, 1.38)	1.23 (1.08, 1.40)	1.22 (0.73, 2.01)
Yellow	T2+T3	1.08 (1.04, 1.13)	1.07 (1.01, 1.13)	1.10 (1.04, 1.17)	1.04 (0.98, 1.12)	1.00 (0.92, 1.10)	1.09 (0.99, 1.20)	1.15 (1.06, 1.26)	1.01 (0.71, 1.48)
Red	T2+T3	1.08 (1.04, 1.14)	1.07 (1.00, 1.14)	1.11 (1.04, 1.18)	1.07 (0.99, 1.15)	0.99 (0.89, 1.11)	1.14 (1.02, 1.28)	1.16 (1.04, 1.28)	1.12 (0.77, 1.68)
Mixed	T2+T3	1.09 (1.04, 1.15)	1.10 (1.02, 1.19)	1.10 (1.02, 1.18)	1.01 (0.92, 1.10)	1.04 (0.92, 1.18)	0.97 (0.85, 1.11)	1.25 (1.11, 1.41)	0.82 (0.52, 1.30)
No grade	T2+T3	1.08 (1.04, 1.13)	1.06 (0.99, 1.13)	1.11 (1.04, 1.17)	1.03 (0.96, 1.11)	1.00 (0.91, 1.10)	1.06 (0.96, 1.18)	1.16 (1.06, 1.27)	1.06 (0.73, 1.57)

Note: All models were fit using quasipoisson models, with indirect age-standardization and adjustment for sex/gender and race/ethnicity (except for single-variable models). All models use HOLC categories weighted by land area. Terciles cutpoints based on the total distribution of each ICE measure for the MA population (2001–2015) are -0.003 and 0.57 (ICE for home ownership), 0.70 and 0.90 (ICE for race/ethnicity), 0.03 and 0.26 (ICE for income), and 0.17 and 0.34 (ICE for race/ethnicity + income).

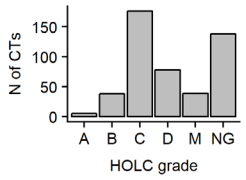
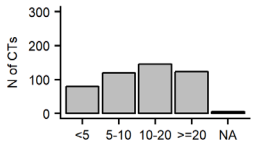
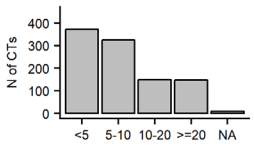
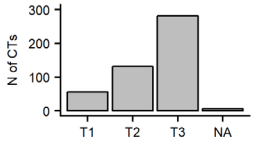
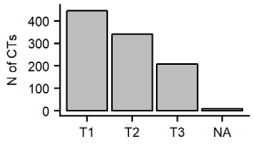
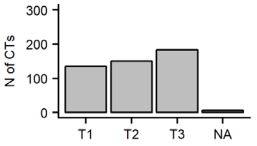
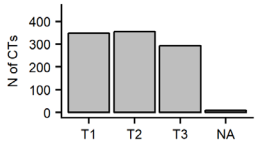
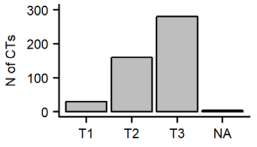
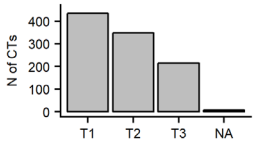
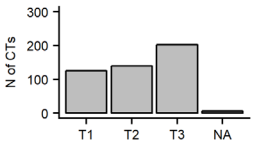
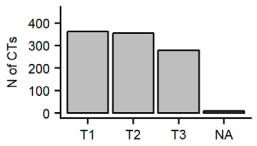
Web Figure 1. Massachusetts municipalities (*n* = 28) with Home Owners' Loan Corporation (HOLC) maps, prepared in 1937–1938

Web Figure 2. Original HOLC area map (1938) and 2010 census tract and neighborhood boundaries: Boston, MA



Note: CT boundaries within neighborhoods are pink, and neighborhood boundaries (which are also CT boundaries) are orange

Web Figure 3. Massachusetts census tract (CT) characteristics for the 28 Massachusetts municipalities with Home Owners Loan Corporation (HOLC) maps, prepared in 1937–1938, and for the Massachusetts cities and towns without these HOLC maps (*n* = 1004), 2001–2015

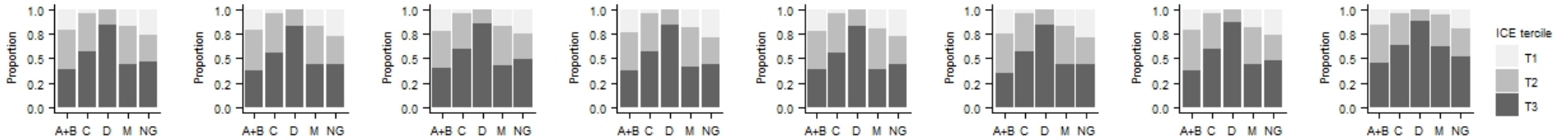
Census tract characteristic (2001–2015)	HOLC maps	No HOLC maps
No. of CTs	474	1004
% of Massachusetts population	29.6	70.4
HOLC grade: A: "Best" (green) B: "Still Desirable" (blue) C: "Definitely Declining" (yellow) D: "Hazardous" (red) M: mixed CT NG: not graded CT		not applicable
CT poverty, categorical		
Index of Concentration at the Extremes (ICE) for home ownership: terciles (CT) T1: most privileged (up to 100% homeowners) T3: least privileged (up to 100% renter)		
Index of Concentration at the Extremes (ICE) for income: terciles (CT) T1: most privileged (up to 100% high income) T3: least privileged (up to 100% low income)		
Index of Concentration at the Extremes (ICE) for race/ethnicity: terciles (CT) T1: most privileged (up to 100% non-Hispanic white) T3: least privileged (up to 100% non-Hispanic black)		
Index of Concentration at the Extremes (ICE) for race/ ethnicity + income: terciles (CT) T1: most privileged (up to 100% non-Hispanic white high-income) T3: least privileged (up to 100% non-Hispanic black low-income)		

-- Operational definition: "A" = census tracts whose land area is [100% A OR (≥ 50% A and <100% A)); "B" = census tracts whose land area is [100% B OR (≥ 50% B and <100% B)]; "C" = census tracts whose land area is [(100% C) OR (≥ 50% C and <100%C)]; "D" = census tracts whose land area is [100% D OR (> =50% D and <100% D)]; "Mixed" = mixed census tracts with ≥ 50% of land area assigned HOLC grades but with no HOLC grade accounting for ≥ 50% of the total land area; "No grade assigned" = census tracts whose land area is ≥ 50% unknown.

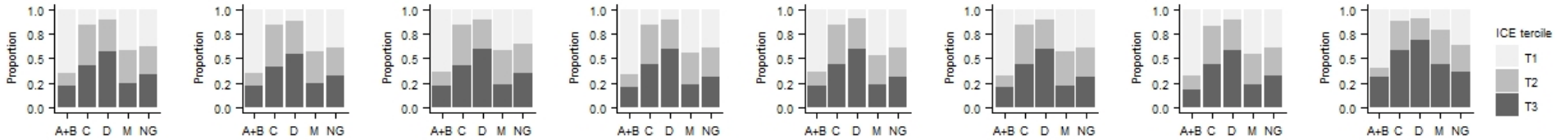
Web Figure 4. Joint distribution of HOLC grades (1937–1938) and current CT characteristics (2001–2015) among the cancer cases, by cancer site (primary invasive lung, colorectal, breast, and cervical cancer), in the 28 MA municipalities with HOLC maps: Massachusetts Cancer Registry data, 2001–2015

Lung cancer			Colorectal cancer			Breast	Cervical	
Total (*n* = 18537 cases)	Women (*n* = 9561 cases)	Men (*n* = 8974 cases)	Total (*n* = 12977 cases)	Women (*n* = 6518 cases)	Men (*n* = 6457 cases)	Women (*n* = 20808 cases)	Women (*n* = 874 cases)	Legend

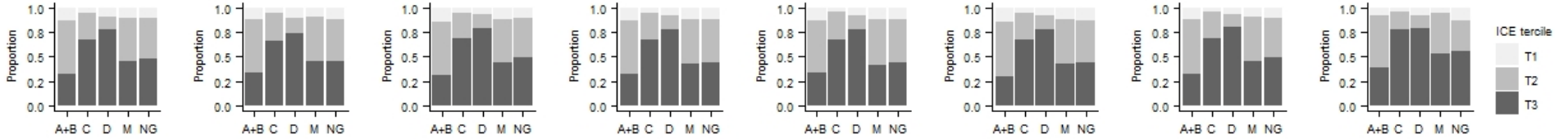
ICE for home ownership



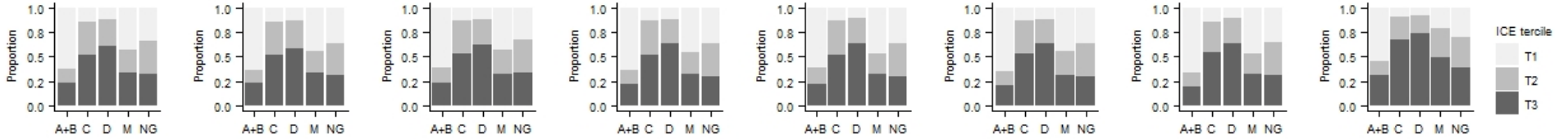
ICE for income



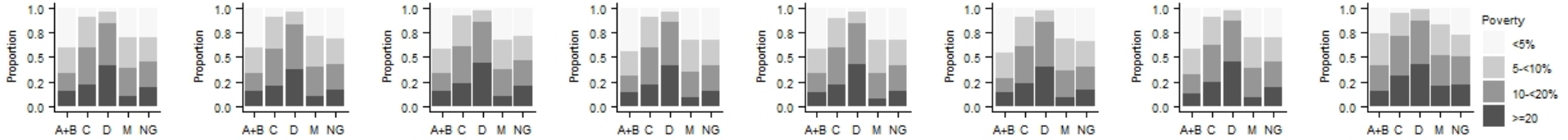
ICE for race/ethnicity



ICE for race/ethnicity + income



Percent below poverty



Abbreviations: CT = census tract; HOLC = Home Owners' Loan Corporation; ICE = Index of Concentration at the Extremes; MA = Massachusetts

--Note: False discovery rate-corrected p-values for Chi square tests all <0.001 (all tests of statistical significance are 2-sided)

-- Operational definition: “A+B” = census tracts whose land area is [(100% A OR (≥ 50% A and <100% A)) OR (100% B OR (≥ 50% B and <100% B))]; “C” = census tracts whose land area is [100% C OR (≥ 50% C and <100%C)]; “D” = census tracts whose land area is 100% D OR (> =50% D and <100% D)]; “M” = mixed census tracts with ≥ 50% of land area assigned HOLC grades but with no HOLC grade accounting for ≥ 50% of the total land area; “NG” = census tracts whose land area is ≥ 50% unknown.