Ethnic Inequalities in Health: A New Ontology of the Geographies of Need in London

Pablo Mateos

Mphil to PhD Upgrade presentation

PhD Supervisors:
Paul Longley and Richard Webber

Camden NHS Primary Care Trust

dti KNOWLEDGE TRANSFER PARTNERSHIPS

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Centre for Advanced Spatial Analysis (CASA)
Department of Geography
University College London
12th October 2005

Contents

1 – Topic Formulation and Justification

1.1- Topic Formulation

Ethnic Inequalities in Health: A New Ontology of the Geographies of Need in London

Social dimension: Ethnicity

Problem: Inequalities / Need

Scope of application: Health

Spatial manifestation: Geographies / London

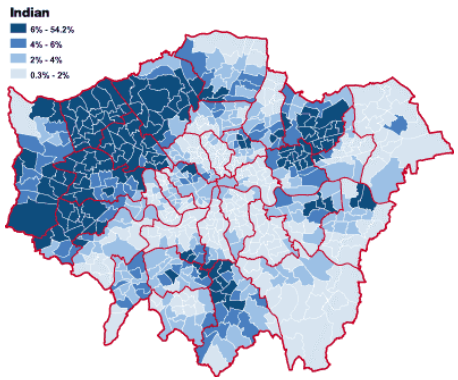
Contribution: New Ontology of Ethnicity

Justification: Ethnic inequalities in health

- Sustained inequalities in health but inability to explain or tackle factors
- A political priority. Health authorities are required to:
  - Prove equity of service provision
  - Combat inequalities in health
- Lack of detailed data & solid research methods on ethnicity
- Immigration & ethnic relations are hot political issues

Justification: Ethnic segregation

- Growing debate on ‘ghettoization’ of Britain, and US abandoning its poor black citizens (*New Orleans)*
- London ethnic minorities represent 40% of total population (UK 12%)
- Need to measure spatial segregation at much finer scales



The Sunday Times

The Sunday Times September 18, 2005

Race chief warns of ghetto crisis

DAVID LEPPARD

BRITAIN'S race relations chief is to warn that the country is "sleepwalking" into New Orleans-style racial segregation, with Muslim and black ghettos dividing cities.

Justification: Ontology of Ethnicity

Ethnicity ≠ Race

Science magazine

"125 big questions that face scientific inquiry over the next quarter-century"

What are human races, and how did they develop?

Anthropologists have long argued that race lacks biological reality. But our genetic makeup does vary with geographic origin and as such raises political and ethical as well as scientific questions.

2 – Literature Review

2.1. Ethnic inequalities in health

2.2. Ontologies and measurement of ethnicity

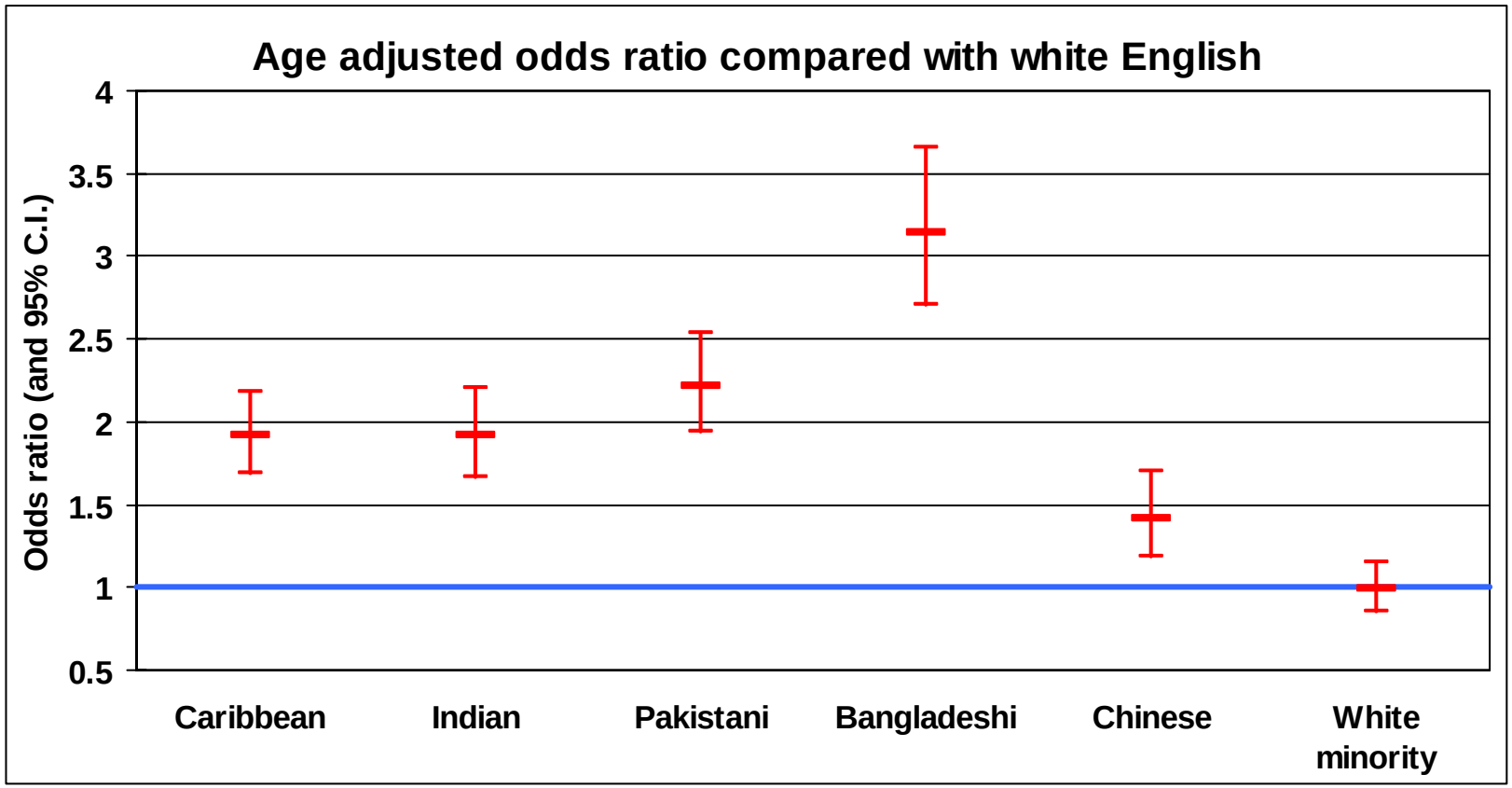
2.3. Names origin & distribution analysis

2.4. Spatial segregation

Ethnic Differences in Health (I)

- One of the two primary goals of US Department of Health "Healthy People 2010" is to *"eliminate health disparities among segments of the population, including; by gender, race of ethnicity, education or income, disability, geographic location or sexual orientation"* (US DoH 2000, 11)
- UK Department of Health - Public Health White paper 2004 "Choosing Health" focuses on tackling health inequalities
- UK Race Relations Amendment Act (2000), explicitly addresses discrimination and racism
- Contemporary societies are composed of increasingly **diverse cultural groups**. Local health status reflects global population health needs, migration histories, and cultures

Reported fair or bad health



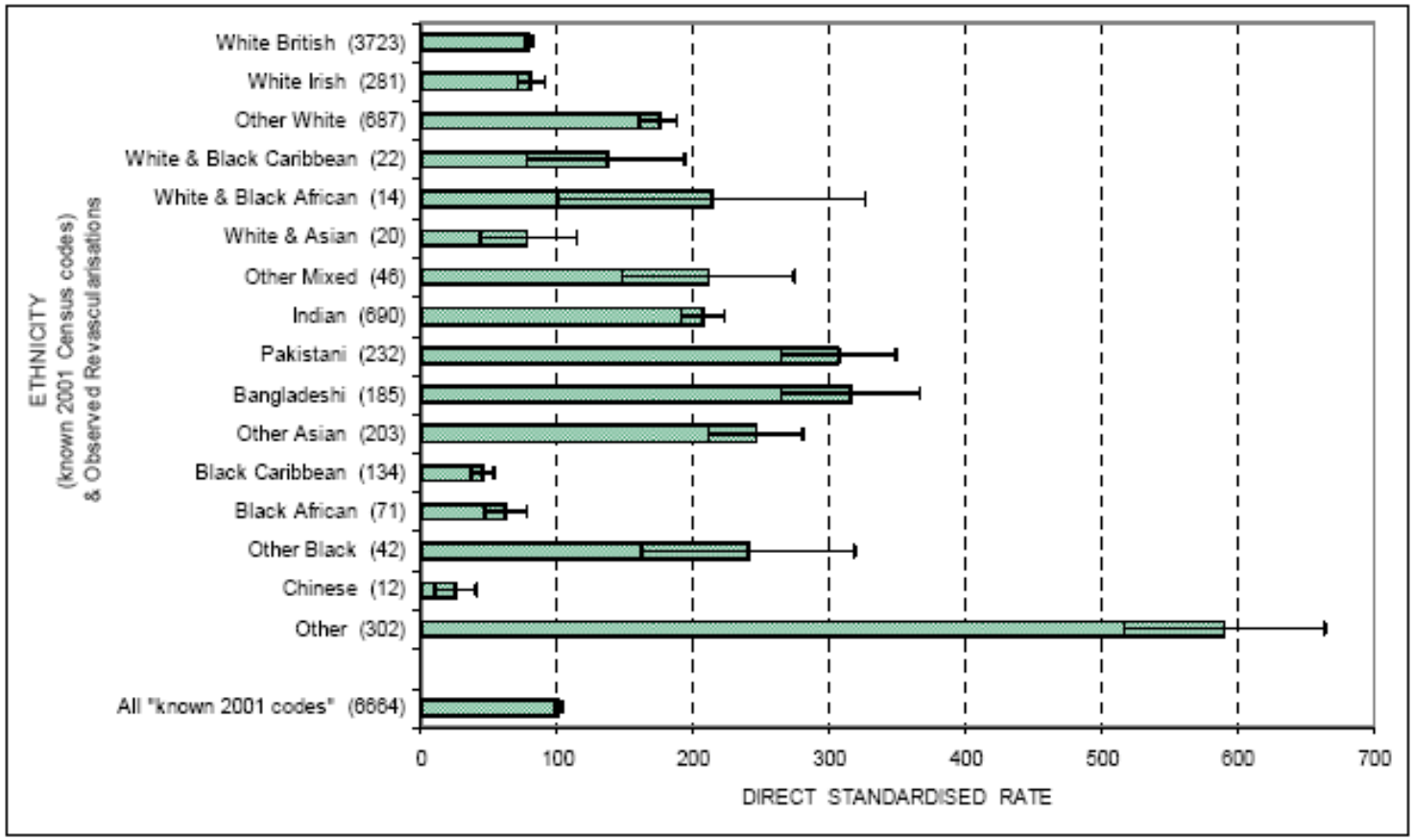
(Nazhroo, 2003 based on Health Survey for England 1999)

Ethnic Differences in Health (II)

- Differences by ethnicity in both the characteristics of populations and their experience of disease have been easy to describe (Senior & Bhopal, 1994)
- Thousands of associations between racial and ethnic groups and disease have been published (Bhopal 1997)
  - E.g. higher risk of:
    - Lung, liver & colon cancer – Caribbean Men
    - Diabetes - Bangladeshis, Black Caribbeans
    - Coronary heart disease – South Asians
- Different use of /access to healthcare services (Cooper et al., 1998)
  - More frequent use of GPs - South Asians
  - Less admissions to hospital – South Asians

Revascularisation by ethnic group in London

(direct standarised rates 2002/03)



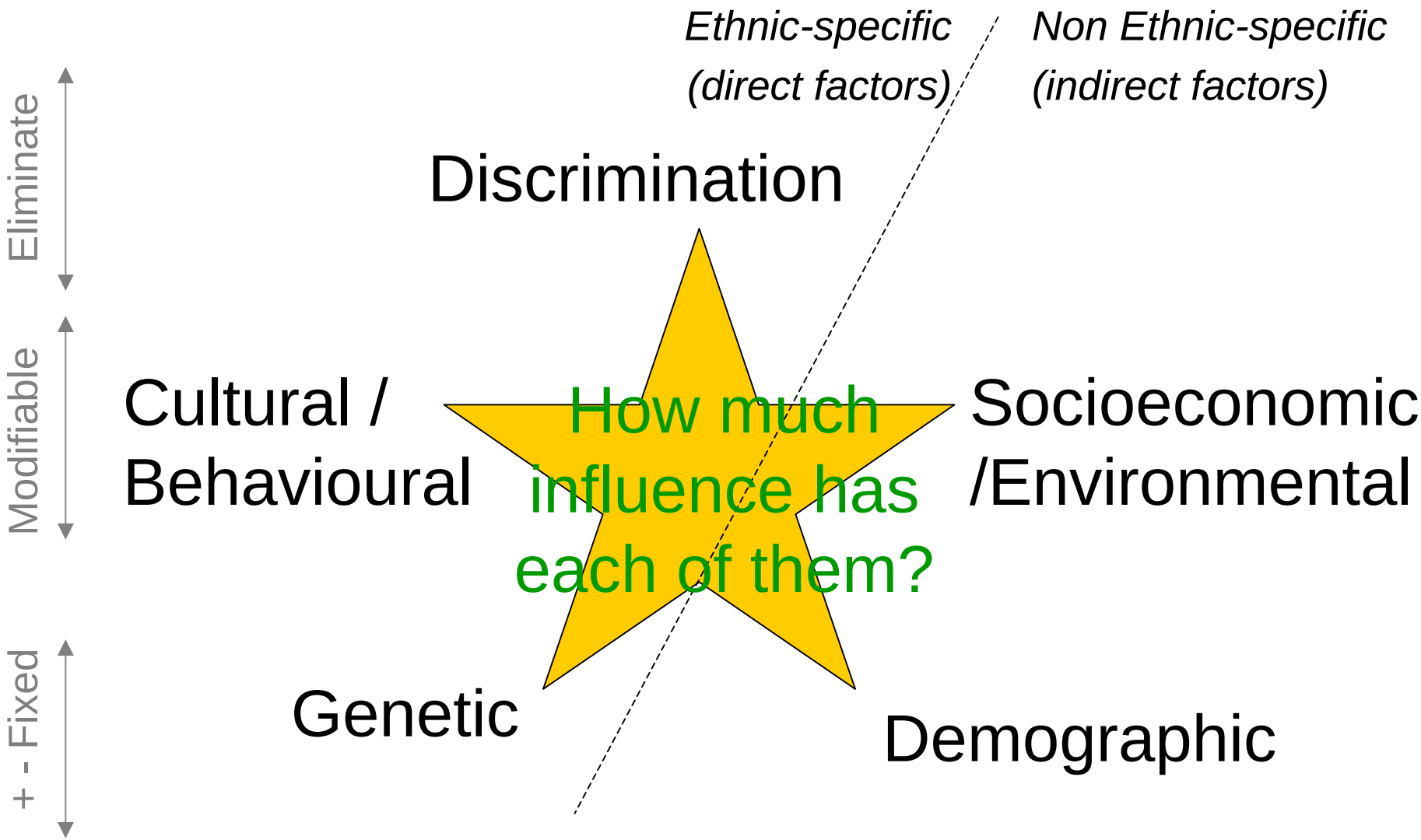
Source: LHO (2005) *Using routine data to measure ethnic differentials in access to revascularisation in London*. Derived from HES data 2002/03

Lack of Causal Explanations

But...

- Factors underlying ethnic differences in health are poorly explained and highly contested. This is considered *"black box" epidemiology* (Skrabanek, 1994)
- There is a classic idea of a package of “specific ethnic diseases”; a racist concept (Bhopal, 1997)
- Very <u>few genetic differences</u> between ethnic groups have been found which directly relate to health (Cooper, 2003)
- Therefore, other environmental, cultural, and socioeconomic and demographic factors should explain those differences

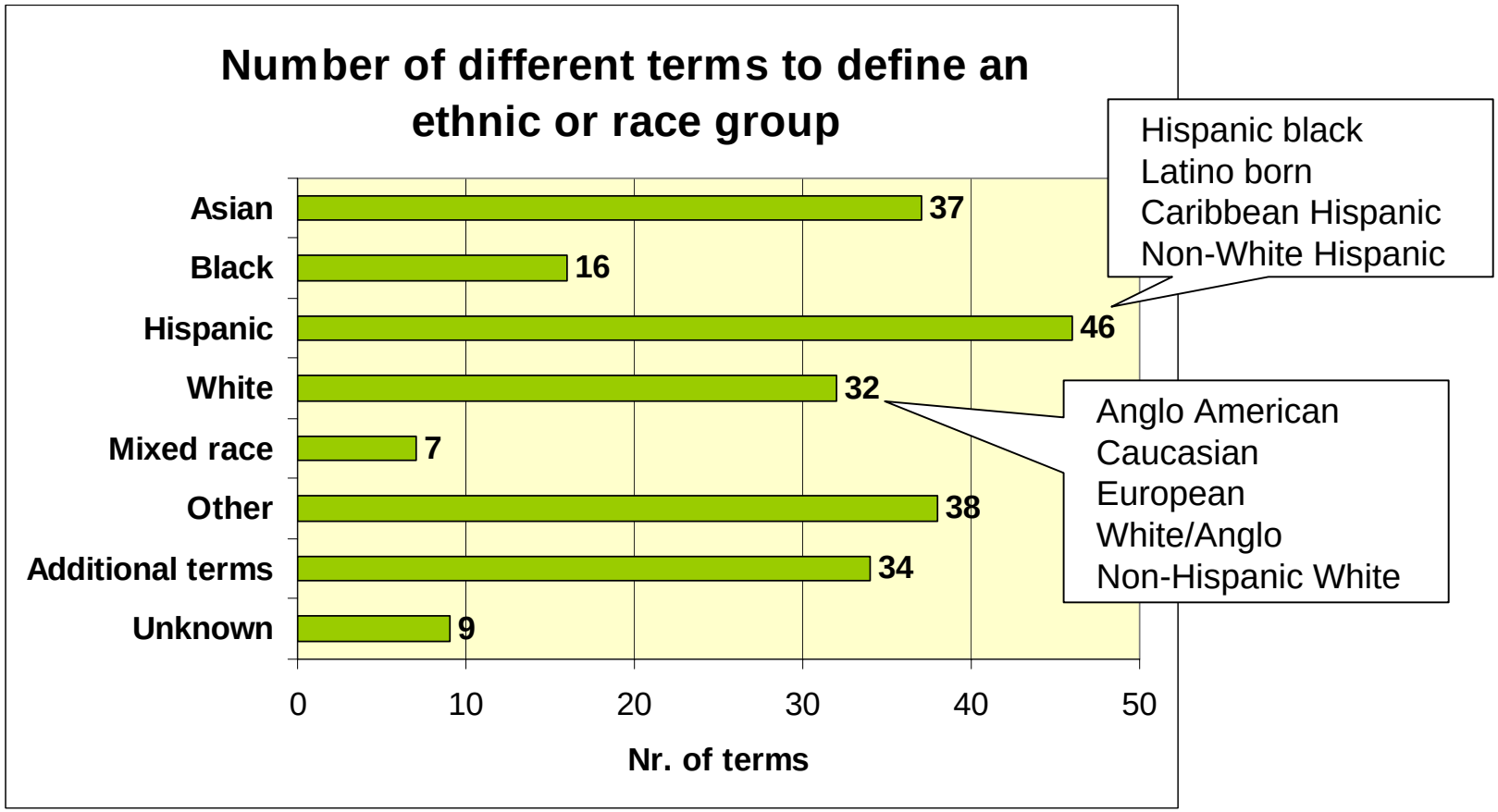
Factors of Ethnic Inequalities in Health



Lack of comparability between studies

- In order to establish some of the causal relationships between these factors and ethnic inequalities in health, a common measurement of ethnicity is required across studies
- Data sources vary enormously in: (Comstock et al, 2004)
  - Definitions of ethnic groups
  - Methods used to ascribe ethnicity to individuals
- Ethnicity has not always been a valuable and sound epidemiological variable due to: (Senior & Bhopal, 1994)
  - Errors of measurement
  - Heterogeneity
  - Ambiguity about the purpose of ethnicity and health research
  - Ethnocentricity
- Contested categories: Who is Asian? (Aspinall, 2003)

Different terms, different ethnicities

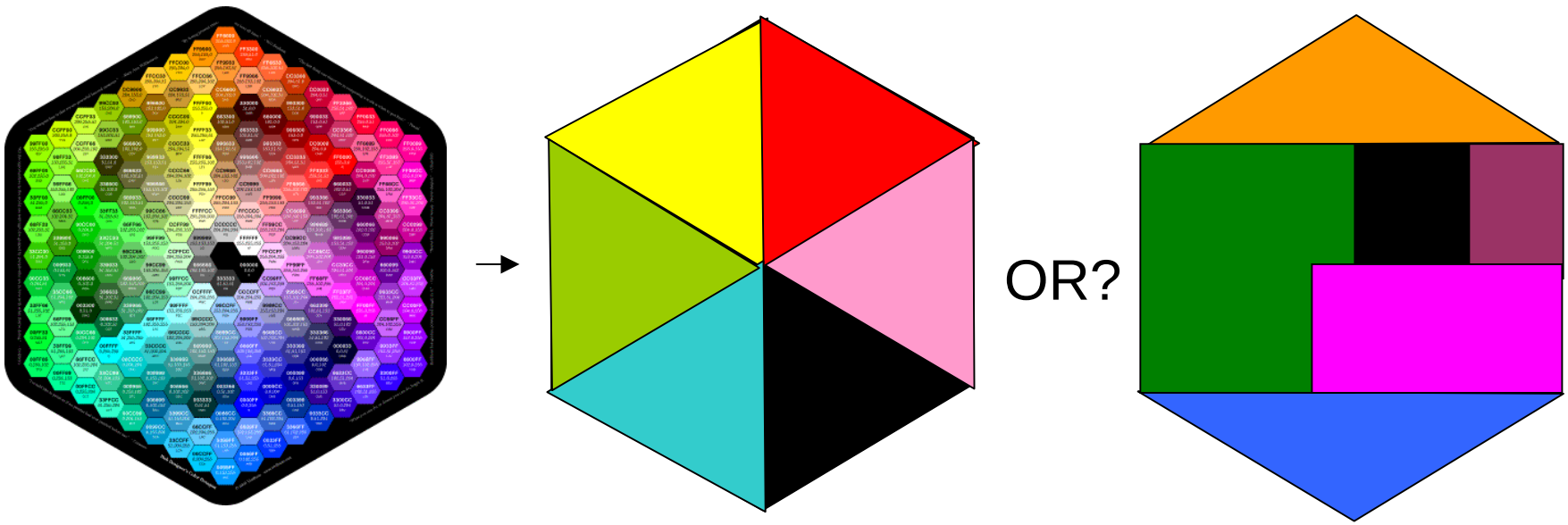


219 terms for 8 'Ethnic Groups' in 1,198 articles published in 2 American epidemiology journals 1996-99

(Comstock et al, 2004)

Defining Consistent Ethnic Groups

The Modifiable Ethnic Unit Problem (**MEUP**)!



Self-identity

Ethnic Groups

2 – Literature Review

2.1. Ethnic inequalities in health

2.2. Ontologies and measurement of ethnicity

2.3. Names origin & distribution analysis

2.4. Spatial segregation

Concepts of Ethnicity and Race

Mitchell's New Primary Geography.

THE SECOND BOOK OF THE SERIES.

THE NEW

PRIMARY GEOGRAPHY:

ILLUSTRATED BY

TWENTY COLORED MAPS

AND

EMBELLISHED WITH A HUNDRED ENGRAVINGS.

The people in the world are divided into five principal races, named according to their color and residence.

What is known of the White race?

The White race is superior to the others, and is found in Europe and America.

Geography of Races (Mitchell, 1868)

An Eurocentric White man view of the world

OF THE PEOPLE WHO INHABIT THE EARTH.

How many people are there in the world?

There are upwards of a billion—(1,000,000,000.)

How are the people in the world divided?

The people in the world are divided into five principal races, named according to their color and residence.

Name the five races in the world.

The five races are the White or Caucasian; the Yellow or Mongolian; the Black or African; the Brown or Malay; and the Red or American.

What is known of the White race?

The White race is superior to the others, and is found in Europe and America.

THE WHITE RACE.

Of the Yellow race?

The Yellow race is found in Asia; the best specimens are in China and Japan.

Of the Black race?

The Black race is found in Africa, and is commonly called the Negro race.

Of the Brown race?

The Brown race inhabits the islands of the Pacific Ocean.

THE YELLOW RACE. (A Chinese Laborer.)

THE BLACK RACE. (An African Chief.)

Of the Red race?

The Red race includes the Indians of North and South America.

THE BROWN RACE. (A New Zealand Chief.)

THE RED RACE. (An Indian Chief.)

Race & Biological Determinism

- 19th century scientists ranked races according to their biological and social worth(Gould, 1984)
- Research in racial theories was used to justify slavery, imperialism, anti-immigration policy, and the social status quo. Biology determined social position- biological determinism. (Bhopal, 1997)
- Eugenics, the improvement of human race. Specially harmful in Germany: the concept of *Racial Hygiene* (Lenz, 1921)

- Race group: *A group perceived as having common inherited and inheritable traits that derive from common descent* (Max Weber 1922)

Concept of Race Today

- The concept of “Race” is socially constructed, and cannot be explained by genetic differences (Olson, 2002)
- None of the numerous racial classifications have stood the test of time (Bhopal, 2004)
- Even though, current ‘race’ classifications are still influenced by ‘biologically rooted’ racial stereotypes
  - Graves (2002) *The Emperor's New Clothes. Biological theories of Race at the Millennium*

Concept of Ethnicity

- The word 'ethnicity' derives from the Greek word *ethnos*, meaning a nation. Thus, the basis of nationalism.
- Ethnic groups (Max Weber 1922)
  - *Those human groups that entertain a subjective belief in their common descent because of similarities of physical type or of customs or both, or because of memories of colonization and migration (...) it does not matter whether or not an objective blood relationship exist*
- Certain shared characteristics are common: (Bhopal, 2004)
  - geographical and ancestral origins
  - cultural traditions and languages (specially)
- A firm belief in group's affinity is required for groups to be defined in opposition to other groups differently perceived and with whom contact is required (Eriksen, 2002)
- The characteristics that define ethnicity are not fixed or easily measured, so ethnicity is imprecise and fluid (Senior & Bhopal, 1994)
- The current preference is for self assessment of ethnicity (*ibid*)

Ethnic diversity as a result of colonial and immigration history

- Ethnic diversity of the population is a dynamic process as old as humankind
- In today's UK current ethnic diversity resembles its imperial past and immigration history


Queen Victoria Presenting a Bible to an Indian subject in 1861

Measuring Ethnic Diversity

White	**91.3%**	
British		87.5%
Irish		1.2%
Other White		2.6%
Mixed	**1.3%**	
White & Black Caribbean		0.5%
White & Black African		0.2%
White & Asian		0.4%
Other Mixed		0.3%
Black or Black-British	**2.2%**	
Black-Caribbean		1.1%
Black-African		0.9%
Black-Other (please describe)		0.2%
Asian or Asian-British	**4.4%**	
Indian		2.0%
Pakistani		1.4%
Bangladeshi		0.5%
Any other Asian background		0.5%
Chinese or other group	**0.9%**	
Chinese		0.4%
Any other ethnic group		0.4%

UK 2001 Census 16+ classification

Total Non- White British	**12.5%**
Poorly Studied Groups	**5.1%**

Q 8 What is Your Ethnic Group
Note: Choose ONE section from A to E, then Ö the appropriate box to indicate your cultural background

- Confusing question!
- Strongly based on a “skin colour problem”
- Represents and reproduces current crude stereotyping of ethnic minorities
- Best used in combination with Country of Birth and Religion

Source: ONS Census 2001 – Great Britain Population

UCL

London 'non-16+ ethnic groups'

(1.2 million people stated 'other' ethnic identities in London 2001 Census)

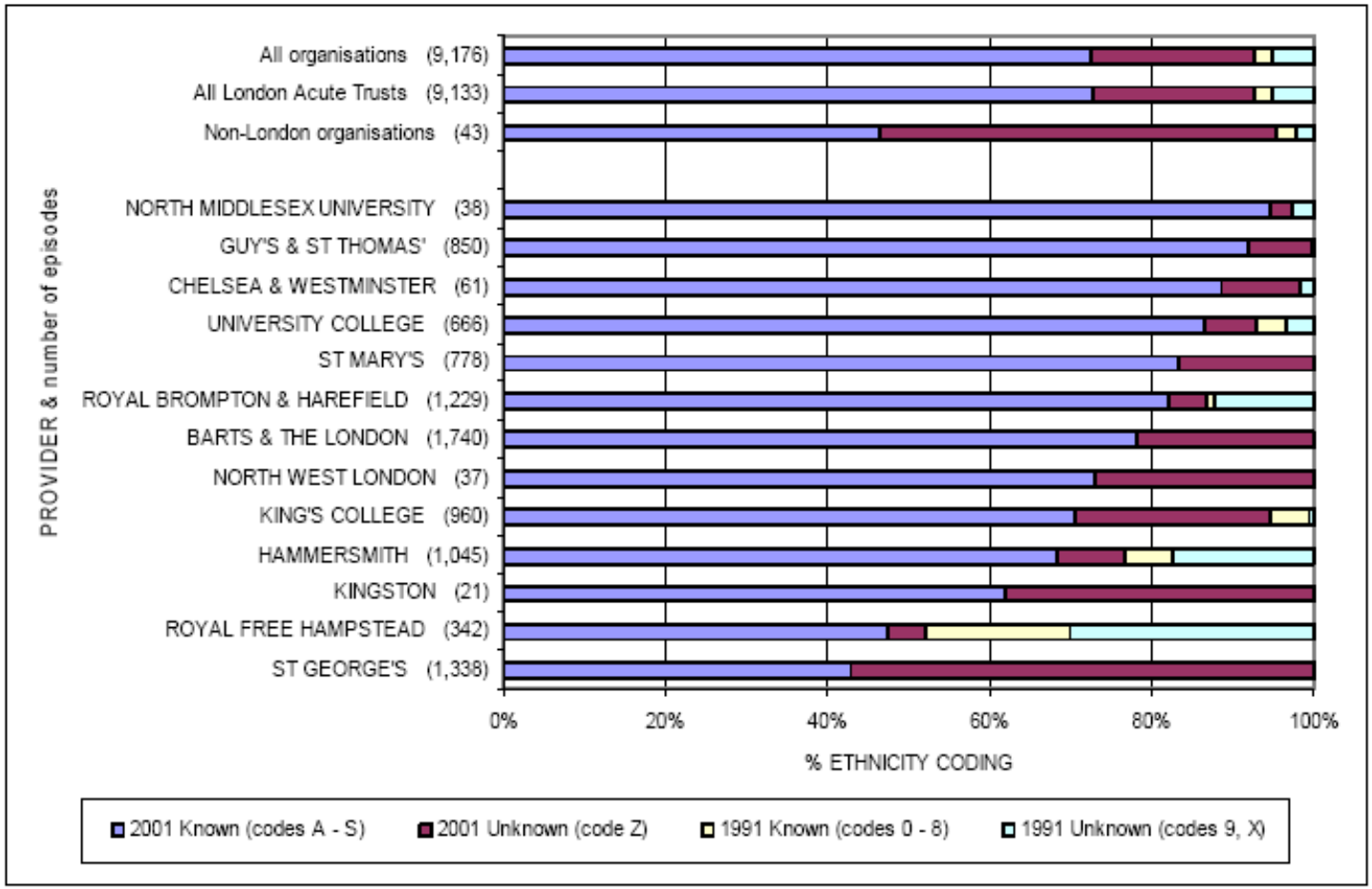
Ethnic Group	Population
Other white European, European Mixed	185,690
Other white, white unspecified	171,744
English	154,203
Sri Lankan	53,307
Black British	46,348
Turkish	37,827
Italian	35,252
Other Mixed, Mixed unspecified	35,027
Any other group	29,469
Greek Cypriot	23,340
Middle Eastern (excluding Israeli, Iranian and 'Arab')	20,537
Arab	20,256
Filipino	19,669
Japanese	19,415
Other mixed white	19,239
Other Asian, Asian unspecified	18,334
Greek	17,888
Iranian	16,494
Multi-ethnic islands	15,952
Polish	15,928
South and Central American	15,607
British Asian	14,625
Turkish Cypriot	14,074

Ethnic Group	Population
Vietnamese	11,719
Commonwealth of (Russian) Independent States	11,606
North African	11,218
Kurdish	9,659
Latin American	9,188
Mixed Black	9,001
Jewish	8,912
Other Black, Black unspecified	8,344
Cypriot (part not stated)	7,360
Mixed: Irish and other white	7,071
Scottish	7,020
Kosovan	6,896
Welsh	6,895
Somali	6,172
East African Asian	5,328
Chinese and White	4,871
Tamil	4,758
Black and White	4,226
Moroccan	4,133
Caribbean Asian	4,070
Black and Asian	3,946
Malaysian	3,384
Albanian	3,226
Sikh	2,814

Source: 2001 Census GLA commissioned tables

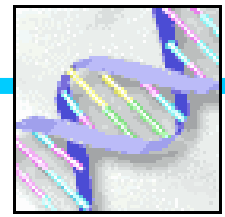
(.../...)

Ethnicity coding in London Hospitals



Source: LHO (2005) *Using routine data to measure ethnic differentials in access to revascularisation in London*. Derived from HES data 2002/03

Human Genetic Diversity



- Human Genome Diversity Project (from 1991)
  - Map differences in genetic markers across populations
- Populations are defined according to linguistic groups, the only objective division that reflects common descent (M'charek 2005)
- Linguistics and evolution are tightly linked

  If we possessed a perfect pedigree of mankind, a genealogical arrangement of the races of man would afford the best classification of the various languages now spoken throughout the world

  Darwin, C (1859) On The Origin of Species, Chapter 13
- If skin pigmentation is ignored, we never find two 'races' totally different, not even for one gene (Cavalli-Sforza, 1995). Most genetic differences occur between individuals.
- However, 'race targeting' of drugs has just started (BiDil for blacks)

  Highly contested genetic studies (Singer, 2005, Wadman, 2005, Kahn, 2005)

2 – Literature Review

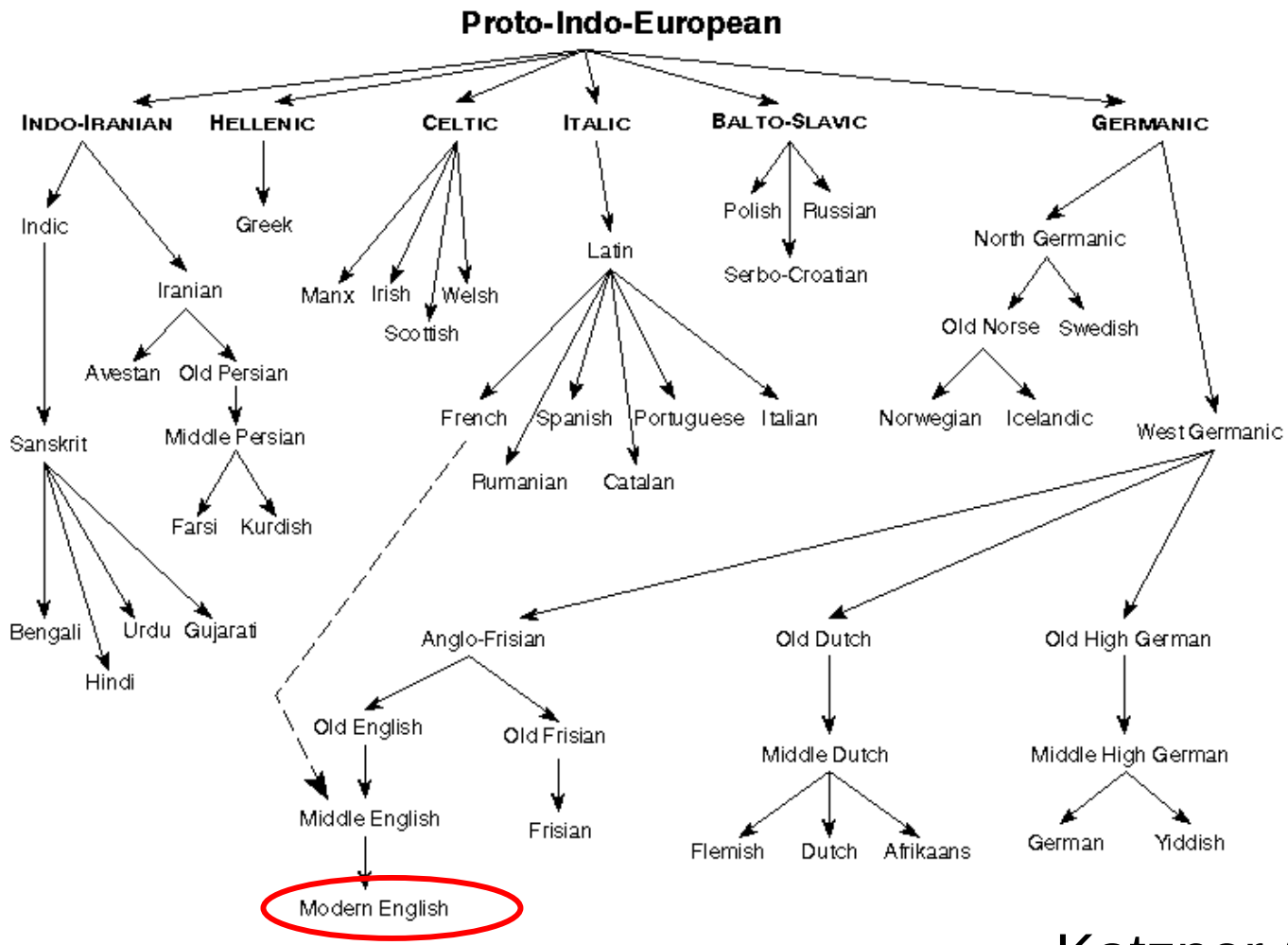
2.1. Ethnic inequalities in health

2.2. Ontologies and measurement of ethnicity

2.3. Names origin & distribution analysis

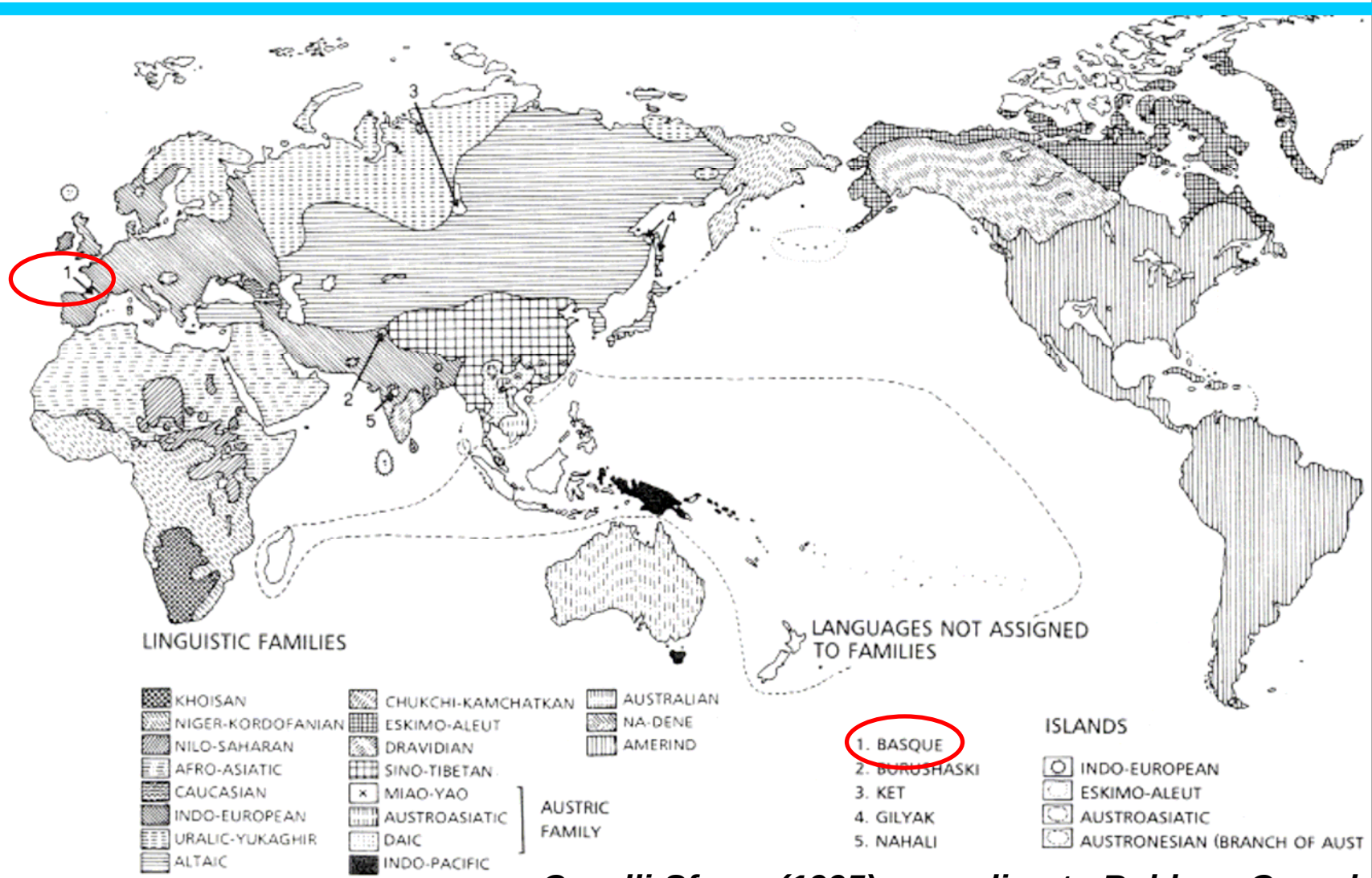
2.4. Spatial segregation

Linguistic Taxonomy



Katzner (2002)

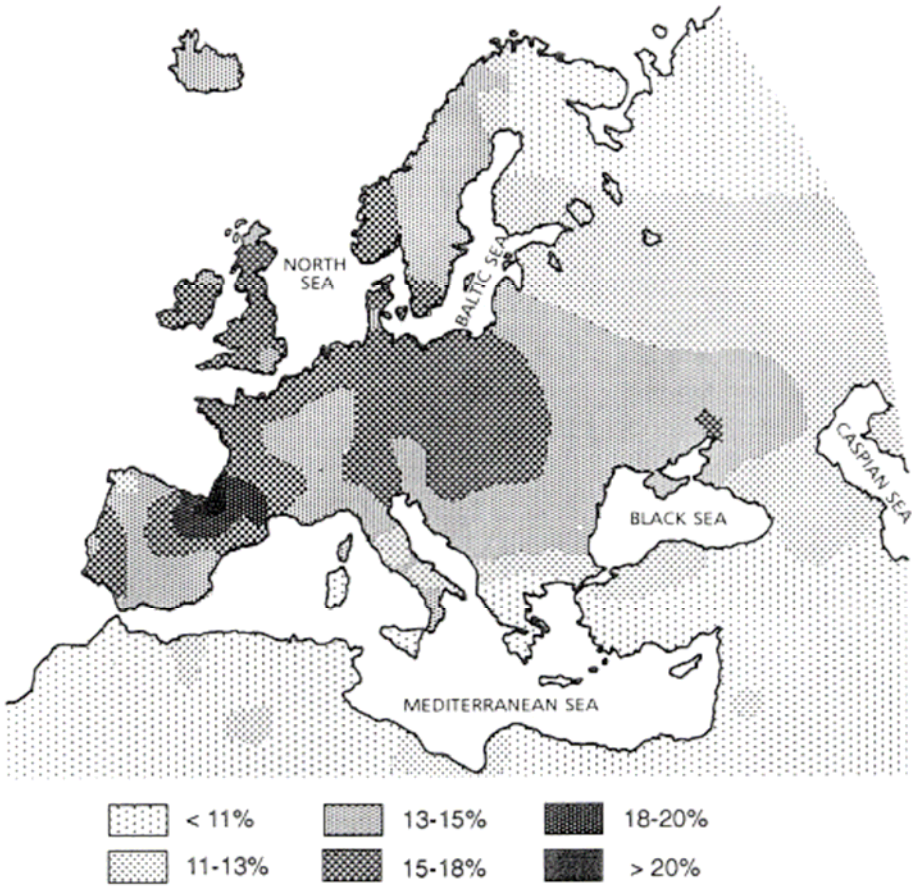
Linguistic Taxonomy



Cavalli-Sforza (1995) according to Ruhlen - Greenberg

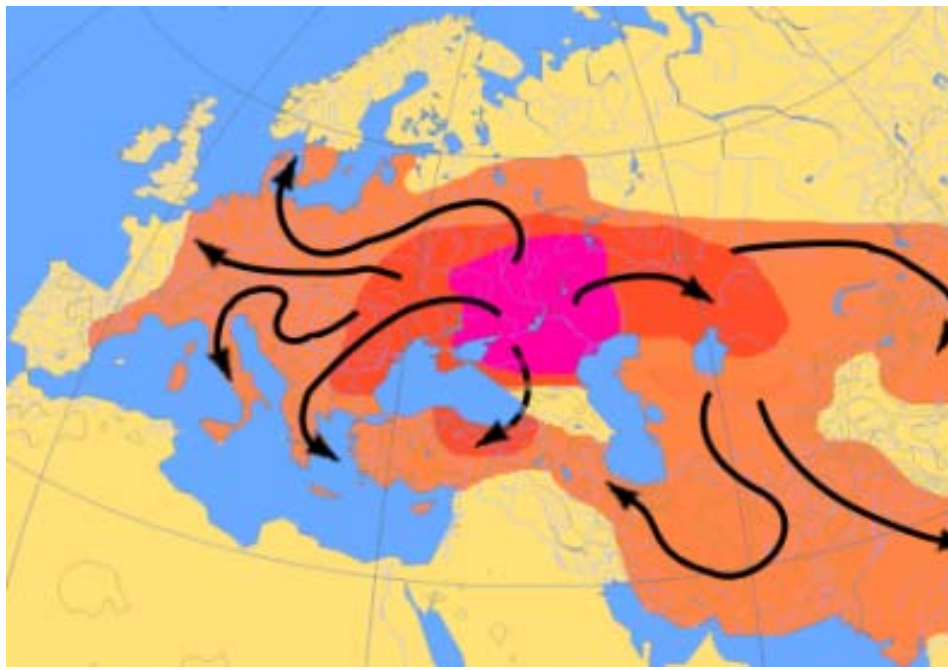
Linguistic & Genetics

Genetic frequency map of Europe (Rh-)



Cavalli-Sforza (1995) Great Human Diasporas

Indo-European Language Family Expansion



Source: Wikipedia

a similar map appears in Cavalli-Sforza (1995)

Name Analysis in Genetic Research

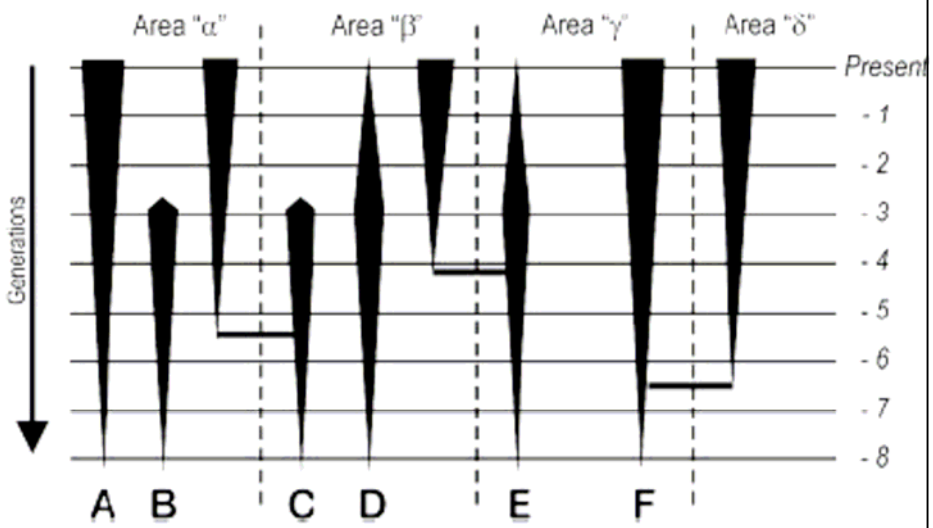
- Surnames generally adopted in the Middle Ages (Europe)
- Surnames in genetic studies dates back to 1875; George Darwin (son of Charles Darwin) used surname frequency to study populaton inbreeding
- Today surnames are used to study ancient patrilineal population structures (Manni et al 2005)

Assumptions:

- Low intermarriage
- Low infidelity
- Common origin (monophyletic)
- Low name change rate

Name Geographic Distribution

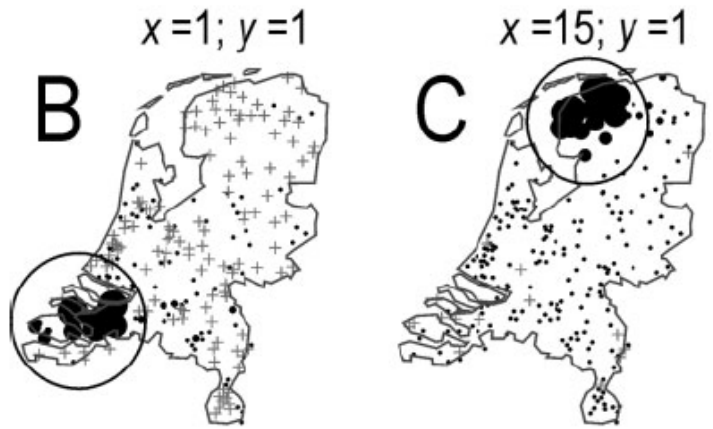
- Surnames current geographical distribution reflect region of origin
- Clusters of surnames “emerge”



Surname evolution

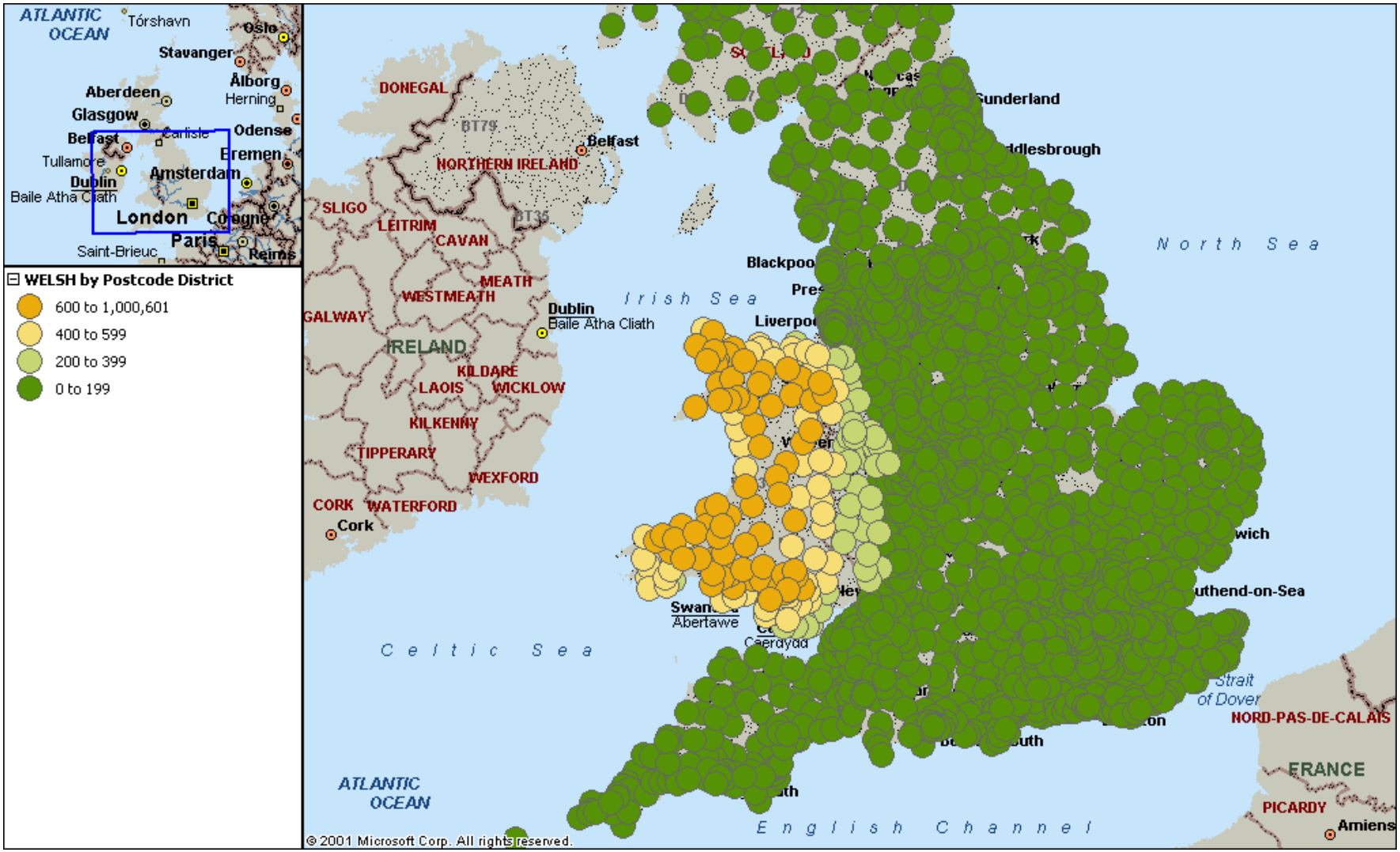
Surname Geographic Origin

Kohonen Self Organising Maps

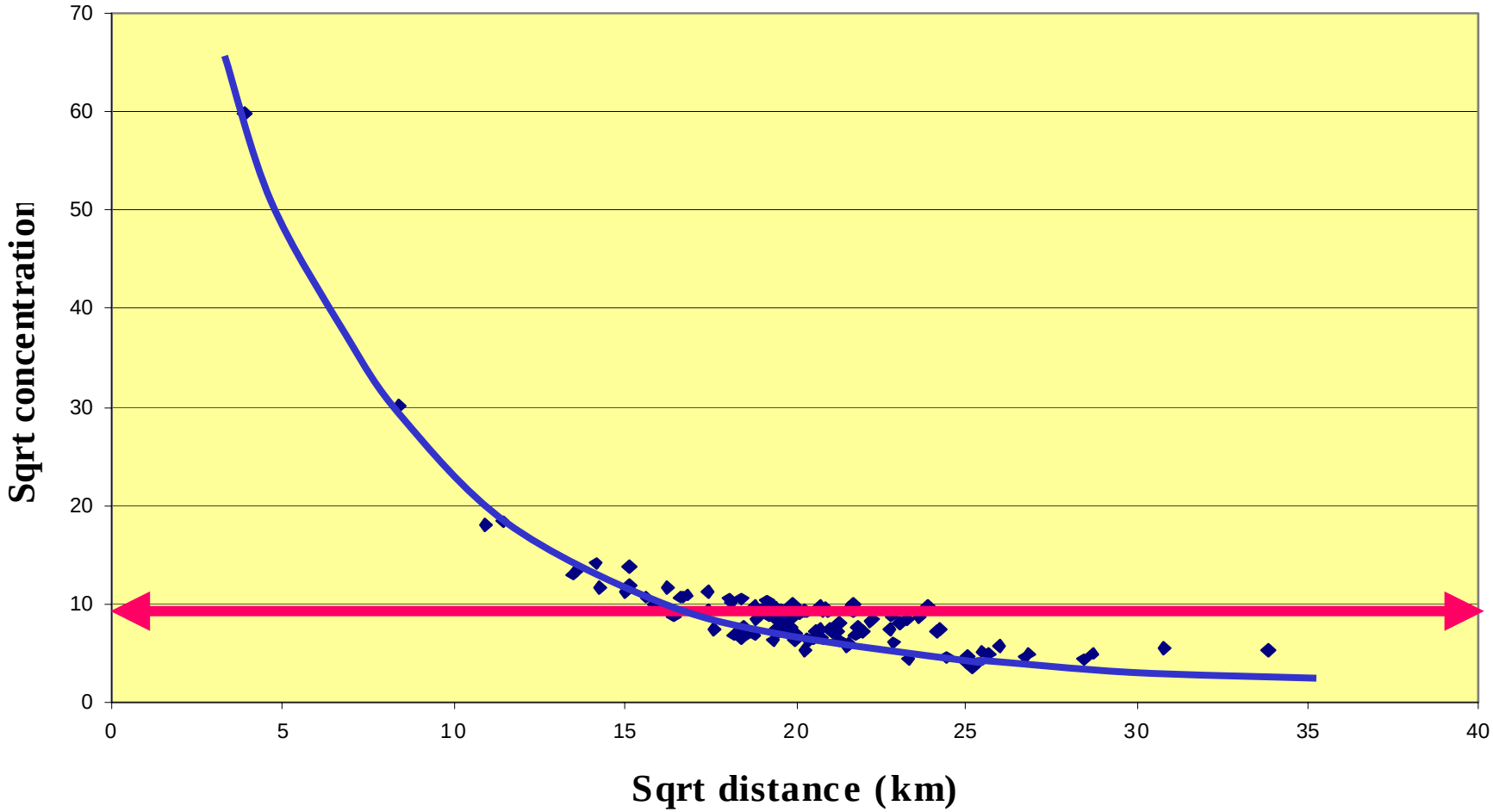


(Manni et al 2005)

Electors with ‘Welsh’ surnames
(Webber, 2005)



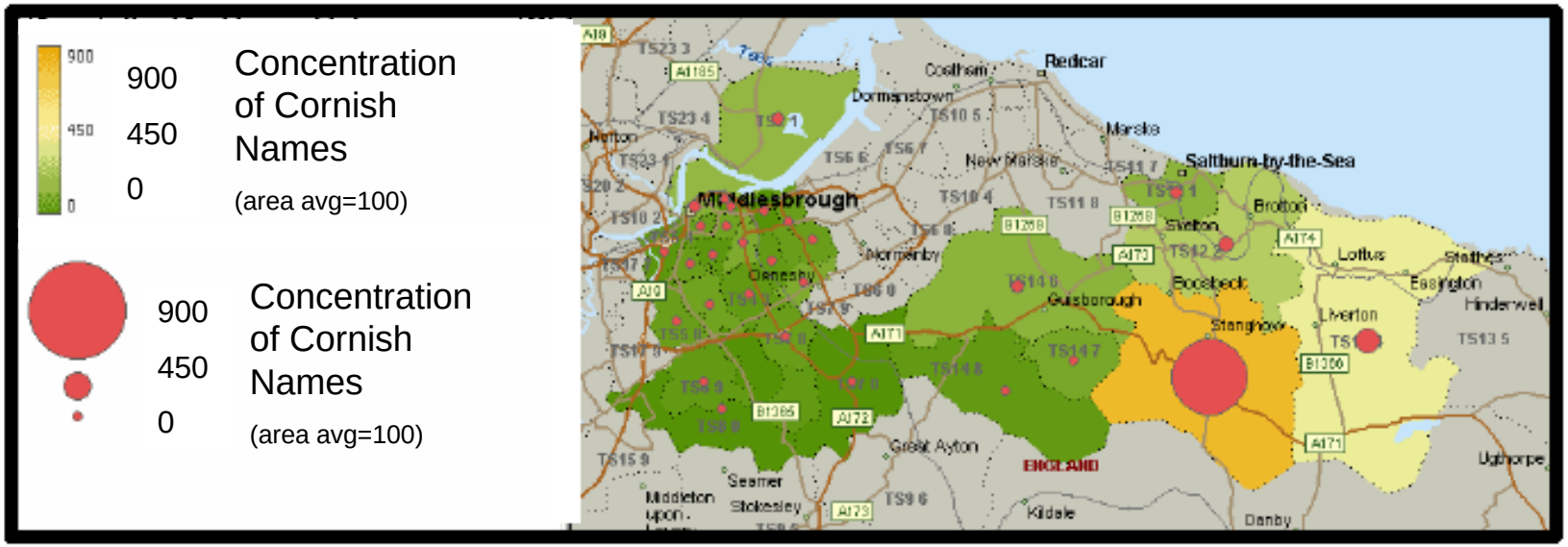
Concentration of ‘Cornish’ names declines with distance from Cornwall



(Webber, 2005)

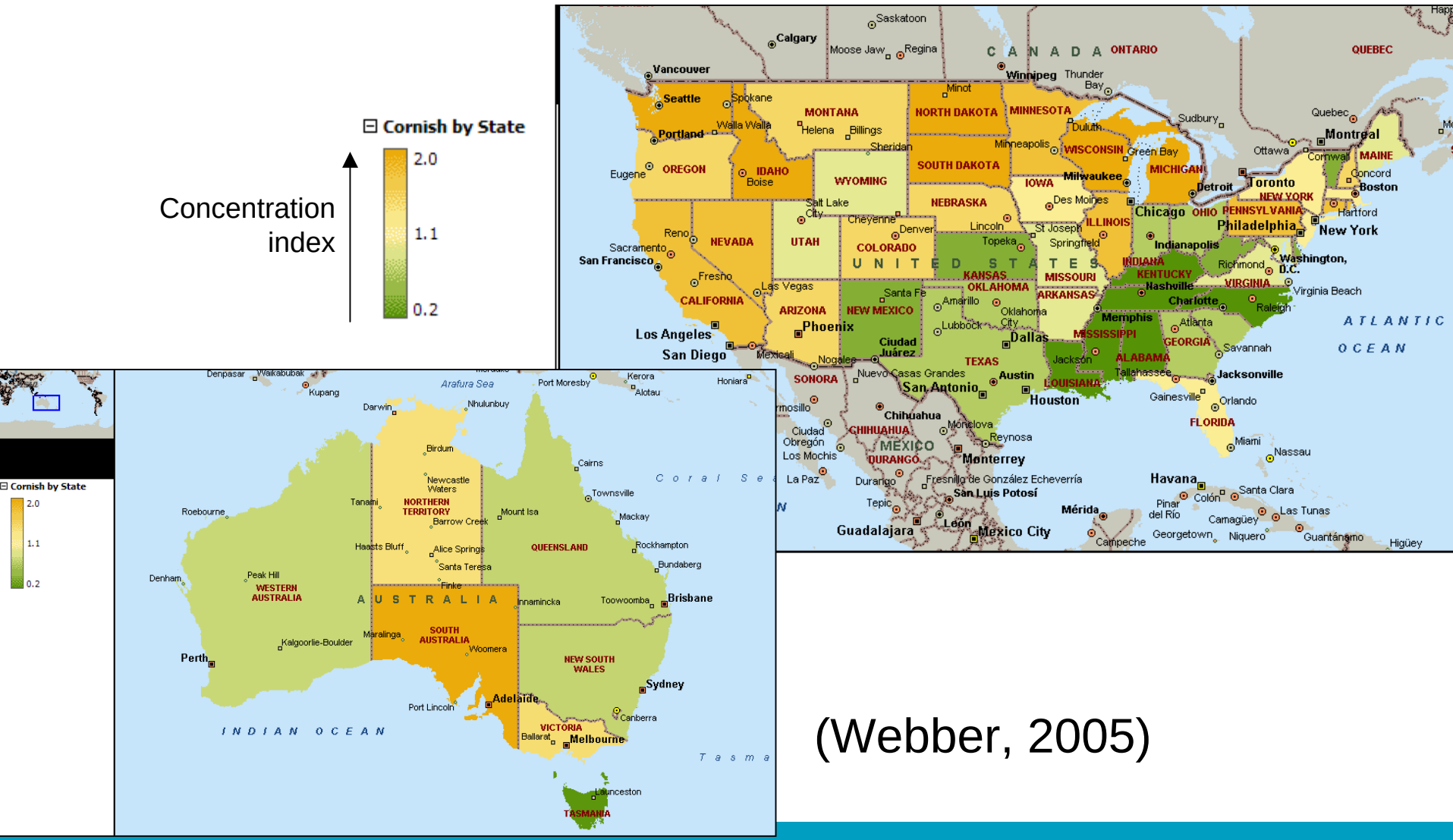
Names Origin and Migration

- High concentration of ‘Cornish’ names in Middlesbrough (North of England) today, due to 19th century miners migration



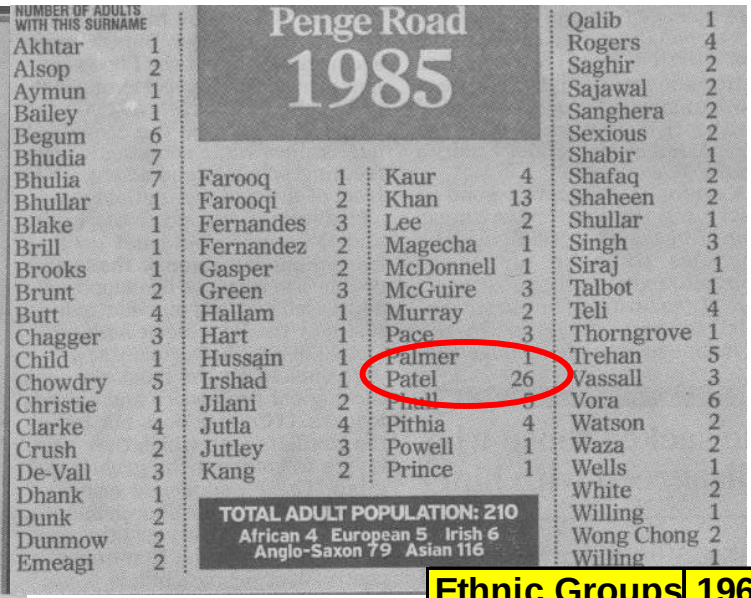
(Webber, 2004)

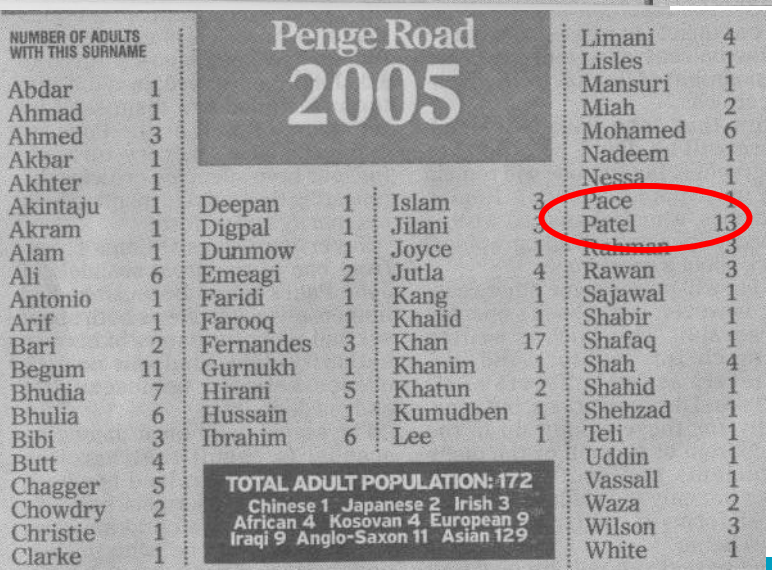
‘Cornish’ names & Anglosaxon diaspora



(Webber, 2005)

Ethnicity change at street level 1965-05





Distribution of Names in Penge Road East London

(Mail on Sunday, 2005)

Ethnic Groups	1965	1985	2005
Anglo-saxon	171	79	11
Jewish	5		
Irish	17	6	3
European	4	5	9
African		4	4
Asian		116	129
Chinese			1
Japanese			2
Kosovan			4
Iraqui			9
TOTAL	**197**	**210**	**172**

Names & Ethnicity in Epidemiology

- Identity, though complex, can be encoded in a name (Seeman, 1980)
- Names can potentially provide information about:

Aspect	Etimology/ Onomastics	Space-time Distribution
Surname & Firstname	**Language**	**Geographic Origin**
	Religion	**Migration flows**
Firstname	**Gender**	**Age**

- In epidemiological studies surnames have been used since the 1950s to subdivide populations when ethnicity data is not recorded (Word & Perkins 1996)

Names & Ethnicity in Epidemiology

- 12 main name analysis methodological papers have been reviewed
- Only some ethnic minorities in the host country have been studied in the US, Canada, UK, Netherlands & Germany:
  a) South Asians
  b) Chinese
  c) Other East and South-east Asians
  d) Hispanics
  e) Turkish f) Moroccans

- Accuracy level 80%-99%
- Most studies use a binary name search: Belong / Not belong (to an ethnic minority), when reality is rather a continuum of name frequencies across many ethnic groups (Word & Perkins 1996)

Literature on Names & Ethnicity

Paper	Geographical area		Ethnic Minorities (EM)	Allocation system	Diccionary Nr Surnames	Nr. Persons coded
	Ctry	Area of study				
Lauderdale & Kestenbaum (2000)	US	National	Chinese, Japanese, Filipino, Korean, Indian, & Vietnamese	Automatic	27,000	1,900,000
Word & Perkins (1996)	US	National	Hispanic	Automatic	25,276	
Razum, Zeeb, & Akgun (2001)	Germany	Rhineland-Palatinate lander	Turkish	Automatic	12,188	4,000,000
Nanchahal, *et al* (2001)	UK	London, W.Midlands, Glasgow	South Asian	Automatic	9,422	130,993
Harding, Dews, & Simpson (1999)	UK	Bradford	South Asian + Hindu, Muslim & Sikh	Automatic	2,995	275,353
Cummins, *et al* (1999)	UK	Thames, Trent, W.Midlands & Yorkshire	South Asian	Automatic	2,995	
Coldman, Braun & Gallagher (1988)	Canada	British Columbia	Chinese	Automatic	544	155,629
Choi, et al (1993)	Canada	Ontario	Chinese	Automatic	427	270,139
Martineau & White (1998)	UK	Newcastle (4 GPs)	Bangladeshi, Pakistani, Indian Muslims, Non-South Asian Muslims, Sikh, Hindu, White, Other	Manual Expert	N/A	137
Bouwhuis &. Moll (2003)	Netherland	Rotterdam (1 Hospital)	Turkish, Moroccan, Surinamese	Manual Expert	N/A	335
Nicoll, Bassett, & Ulijaszek (1986)	UK	Selected areas	South Asian	Manual Expert	N/A	846
Harland, White & Bhopal (1997)	UK	Newcastle	Chinese	Manual Expert	N/A	129,914

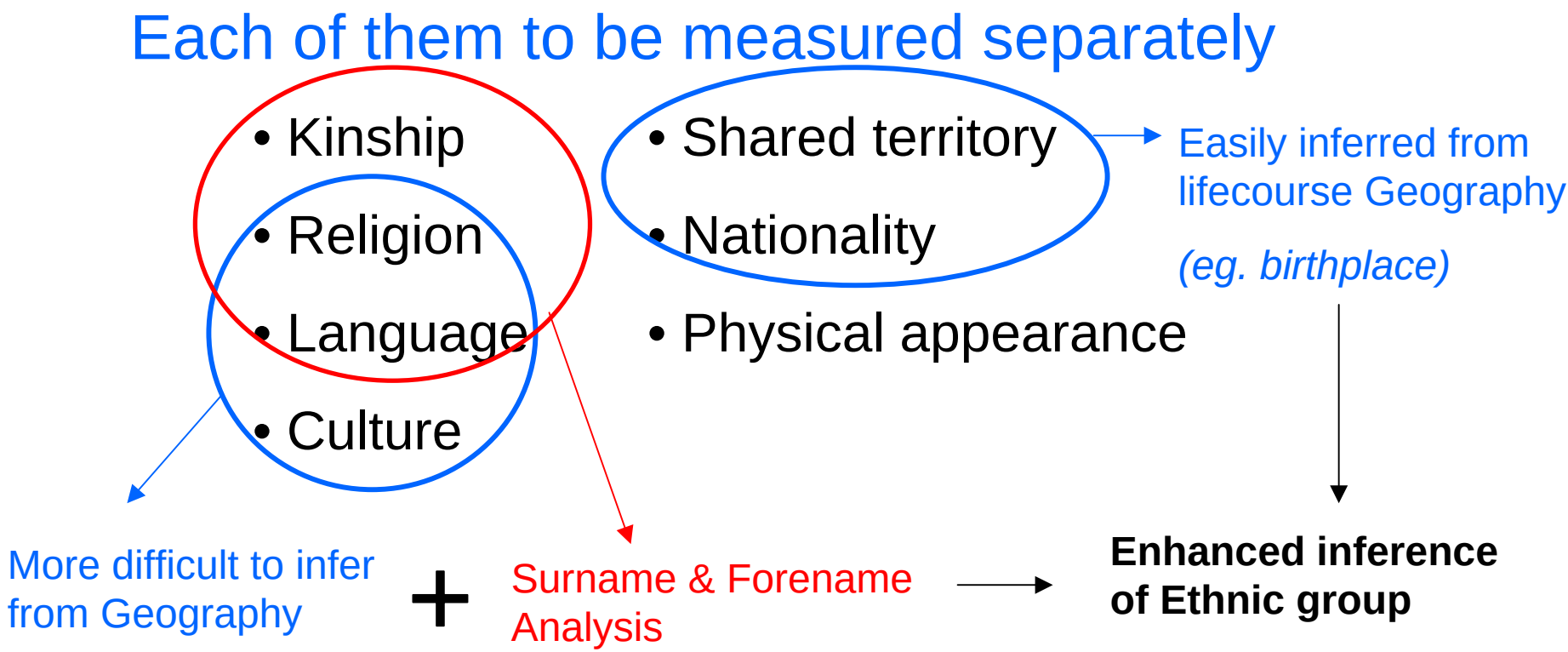
Issues with Names Analysis

- Only reflects patrilineal heritage
- Different history of surname adoption, naming conventions & surname change
- Name normalisation is required
- Family/Household Autocorrelation
- Limited names lists, due to temporal & regional differences in name distribution
- Lack of consistency in self-conceived identity

(Senior & Bhopal, 1994; Martineau 1998, Word & Perkins, 1996; Jobling 2001)

A new ontology of ethnicity

- **<u>Ethnicity</u>: A multi-dimensional concept that encompasses different aspects of identity:**



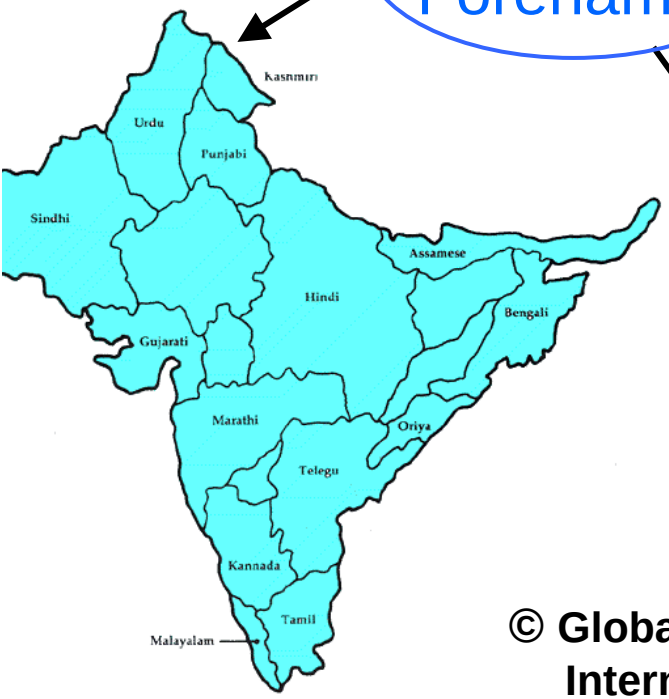
Cultural Ethnic Linguistic (CEL) types

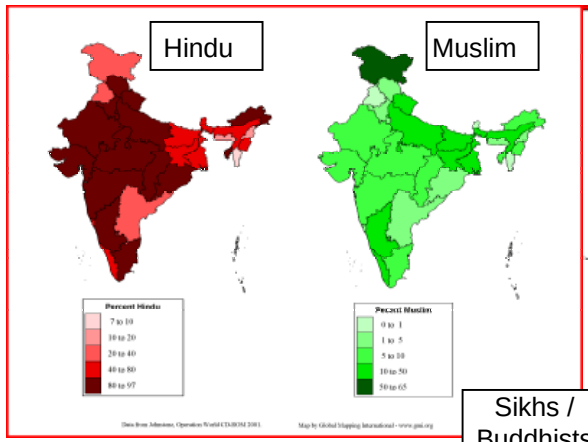
Religion

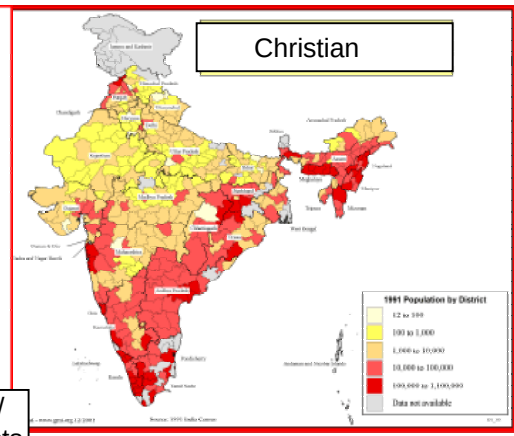
Language

Surname & Forename

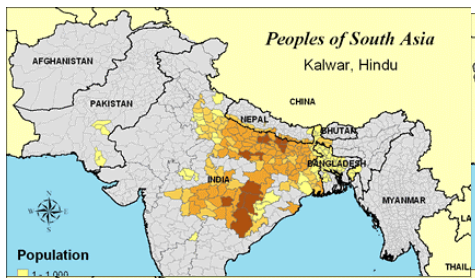
Ethnicity/ Territory

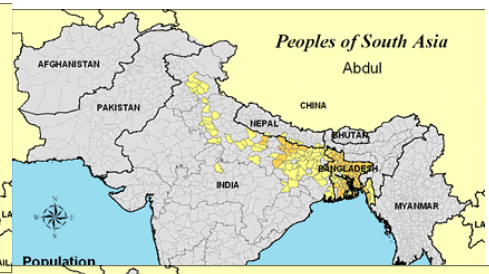


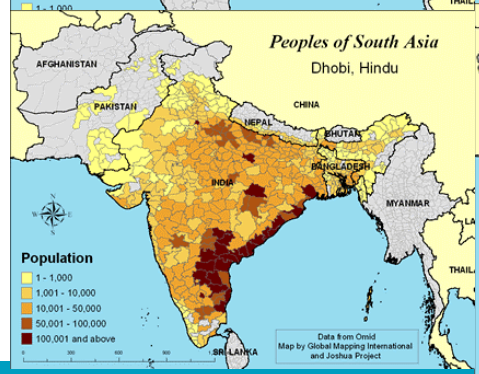


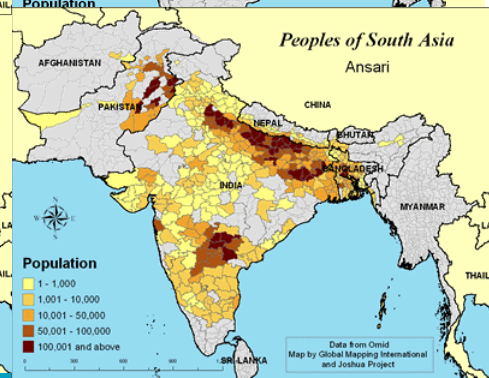


Sikhs / Buddhists









2 – Literature Review

2.1. Ethnic inequalities in health

2.2. Ontologies and measurement of ethnicity

2.3. Names origin & distribution analysis

2.4. Spatial segregation

Ethnic Residential Segregation

- Very high in US cities but typically low in Europe
- Debate on 'ghettoization' of Britain re-opened 19th Sept. 2005 by Trevor Phillips, chairman of the Commission for Racial Equality (CRE) (Leppard, 2005)

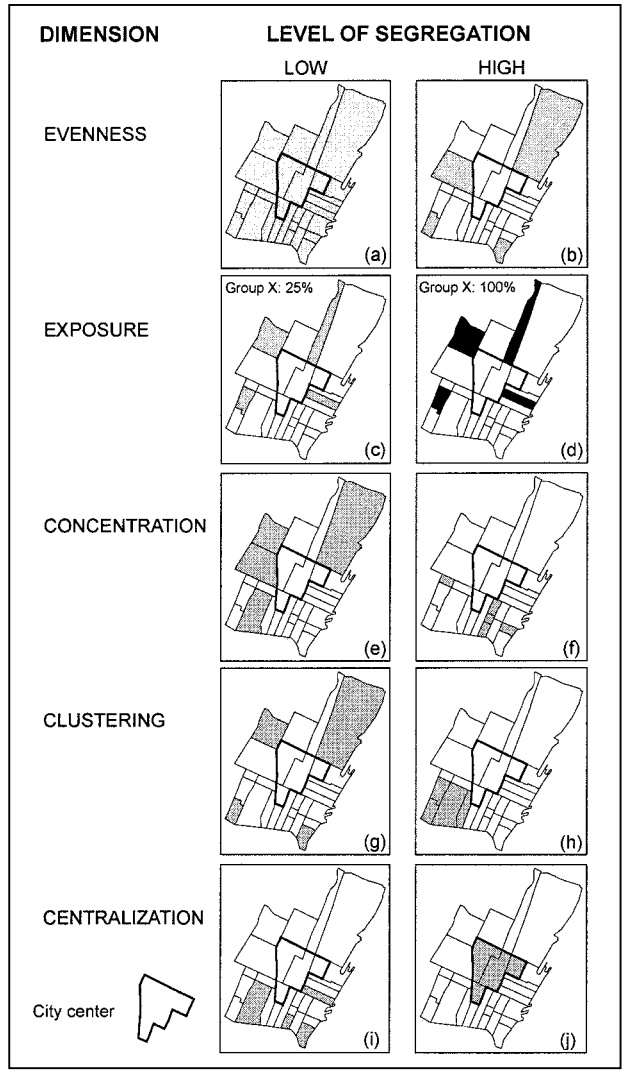
  " the July terror attacks have exposed a racial "nightmare" where some districts are becoming "fully-fledged ghettos — literal black holes" where people fear to go. (...) the country is "sleepwalking" into New Orleans-style racial segregation, with Muslim and black ghettos dividing cities. (...)
- But he added

  "there are also concerns about white working-class ghettos in places such as Barking, Essex, and parts of Yorkshire"
- The real threat is the growing divide between rich and poor

  Dorling (cited by The Observer, 2005) and The Economist (2005)

Spatial Segregation

- Broadly studied since the 1970's
- 5 Dimensions of Spatial Segregation (Massey and Denton 1988)
- Typical spatial analysis issues:
  - Contiguity
  - Connectivity
  - Concentration/Dispersion
  - Centrality

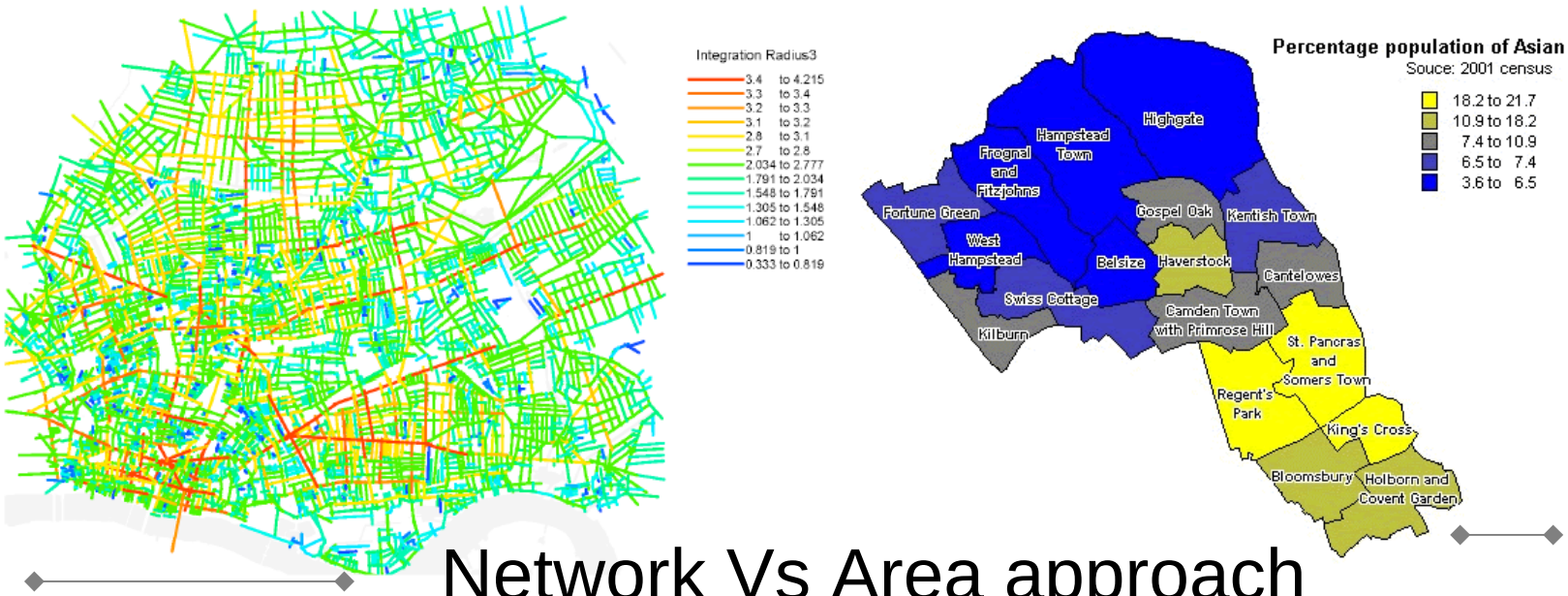


Spatial Segregation Indices

- 3 Types of indices
  - One-group: 1 group -> Entire pop.
  - Inter-group: 1 group -> 1 group
  - Multi-group: Several groups
- Main Multi-group Indices
  - Global measures
    - Spatial multi-group proximity index (Grannis, 2002)
    - Spatial multi-group dissimilarity index (Wong, 1998)
    - Spatial exposure index (Morgan, 1983)
    - Standard deviational ellipse index (Wong, 2002)
  - Local measures
    - Multi-group local entropy index (Wong, 2002)
    - Multi-group Local Getis index (Wong, 2002)

Issues of Scale & Spatial Representation

- Geographic data used is typically aggregated to coarse areas (eg. Wards or OAs)
- Street spatial configuration is ignored (Vaughan 2005)



Network Vs Area approach

Camden ethnic groups- Highgate

CEL/ COB Map

Very fine detail map at individual level is presented here.

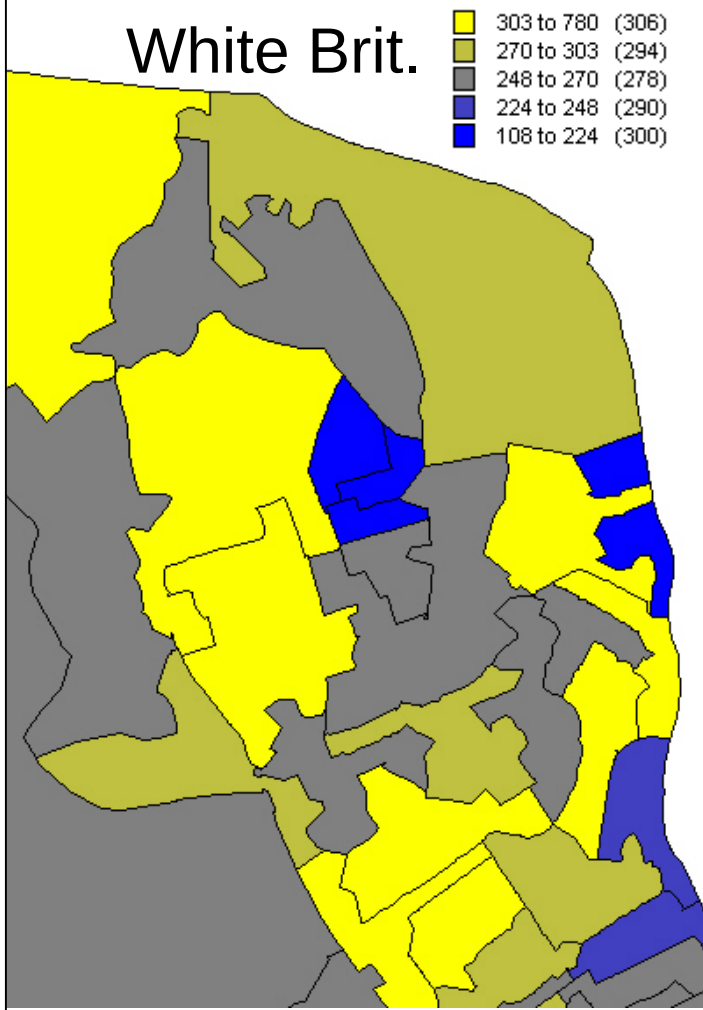
Removed from the Handout version due to confidentiality issues.

Ethnic Groups (COB aggreg.)

Africa;Black African	(4495)
Africa;South Cone	(1995)
Anglophone	(6000)
Anglophone;Caribbean	(368)
Asia	(1222)
Asia;Hindi or Sikh	(1834)
East Asian;CHINESE	(2121)
East Asian;Indochina	(334)
East Asian;Korean	(357)
Eastern Europe	(1309)
Europe;Dutch	(485)
Europe;French	(2076)
Europe;German	(2006)
Europe;Greek	(1441)
Europe;Irish	(3941)
Europe;Italian	(2196)
Europe;Mix	(5771)
Europe;Nordic	(1189)
Europe;Slav	(2041)
Hispanic;Brazilian	(838)
Hispanic;Latin America	(772)
Hispanic;Portugese	(986)
Hispanic;Spain	(1392)
Hispanic;Spanish_World	(1029)
JEWISH	(866)
Muslim;Black African	(2997)
Muslim;Eurasia	(512)
MUSLIM;Middle East	(2285)
Muslim;North African	(1136)
Muslim;Otoman	(660)
MUSLIM;South Asia	(7428)

Census Map (OAs)

White Brit.



3 – Research Questions

3- Research Questions (I)

- What would be an appropriate typology of ethnic groups to study ethnic inequalities in health in London at the individual level?
- Are birthplace and name origin data valid proxies to allocate a probability of ethnicity at individual level? Can they contribute to longitudinal analysis of social mobility and migration history?
- Is there evidence of ethnic residential segregation in inner London, at what scales is it manifested, and how does it differ from the geography of social deprivation?

3- Research Questions (II)

- What alternative methods to traditional segregation indices are more efficient in analysing a large number of ethnic groups at the individual/household level?

  (e.g. a network approach vs. the traditional clustering of adjacent bounded spaces)

- How do differences in ethnic inequalities in health vary across London, and Camden in particular, and between or within ethnic groups, using these methods at different spatial and temporal units of analysis?

4 – Research Design

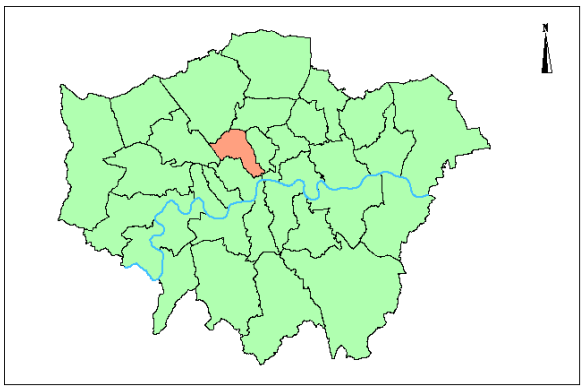
Study Area

- London, a World City

  40% pop. ethnic minorities

  46% of UK ethnic minorities

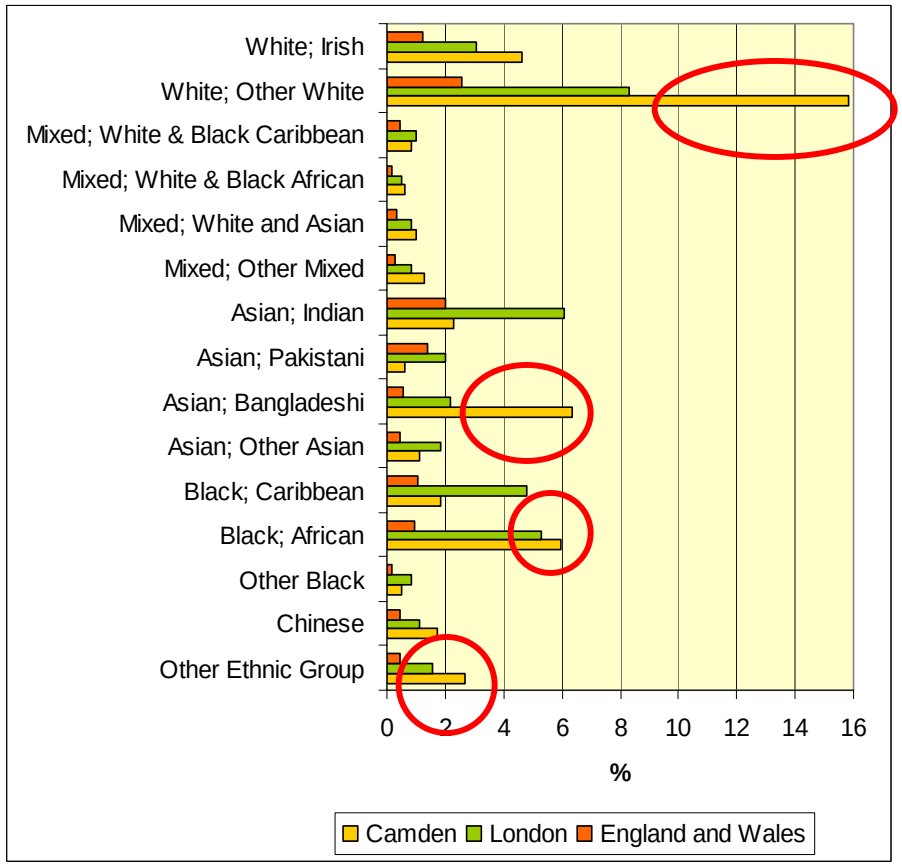
- Camden, a Borough of stark inequalities



210,800 people

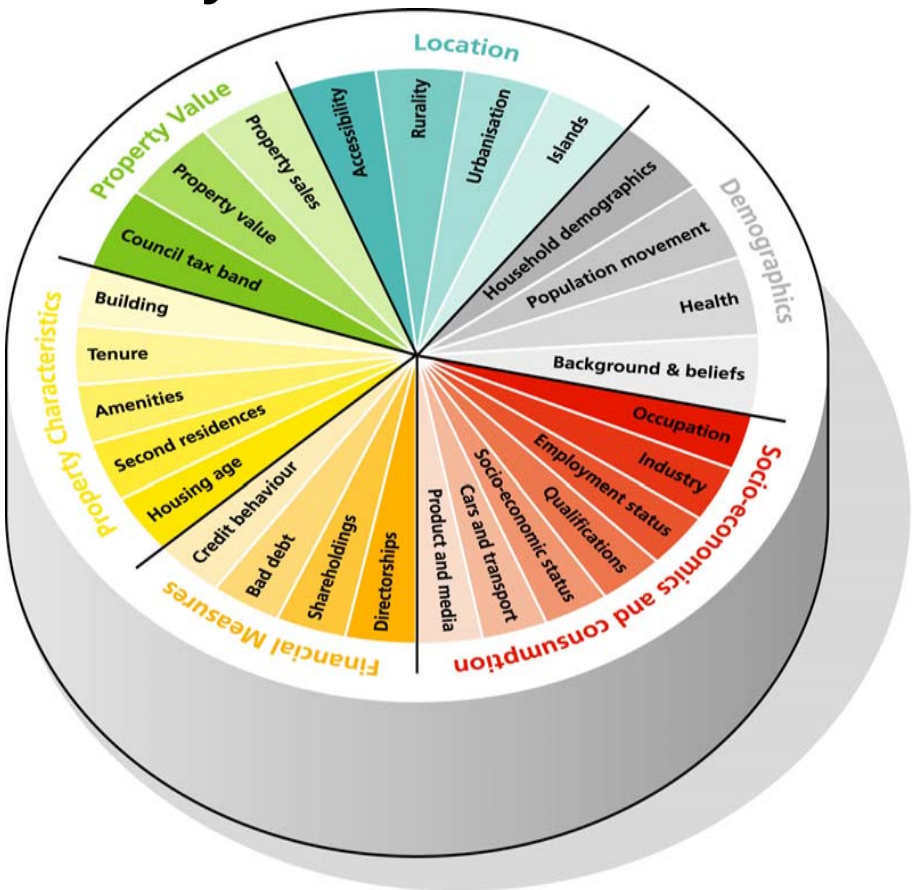
19th most deprived LA

Camden Ethnic Minorities



Geodemographics

Small area measures of socioeconomic, demographic & lifestyle characteristics



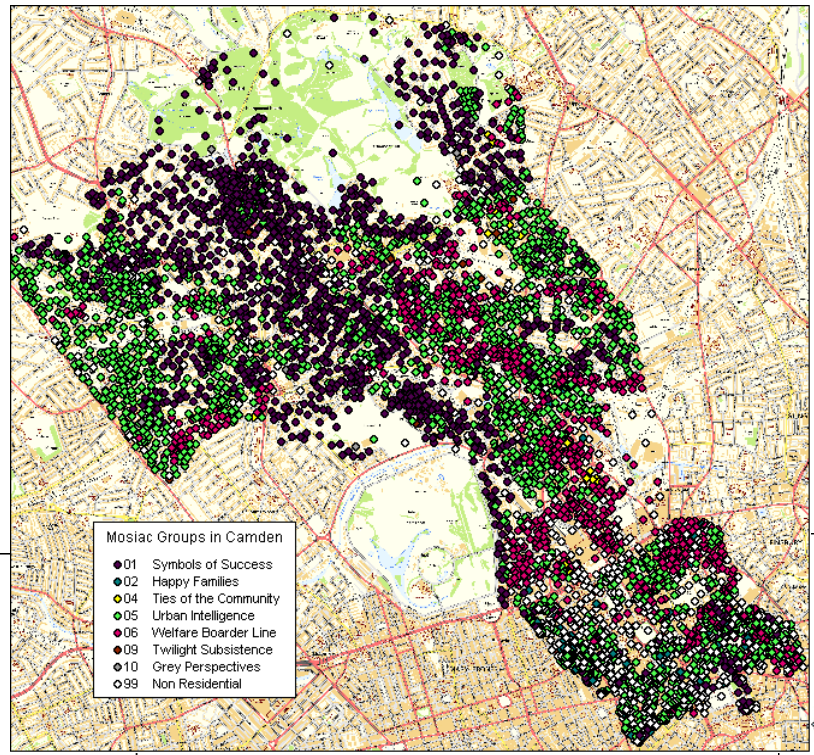
mosaic™

- Postcode unit level
- Classifies UK~1.6million postcodes into:
  - 61 types
  - 11 groups

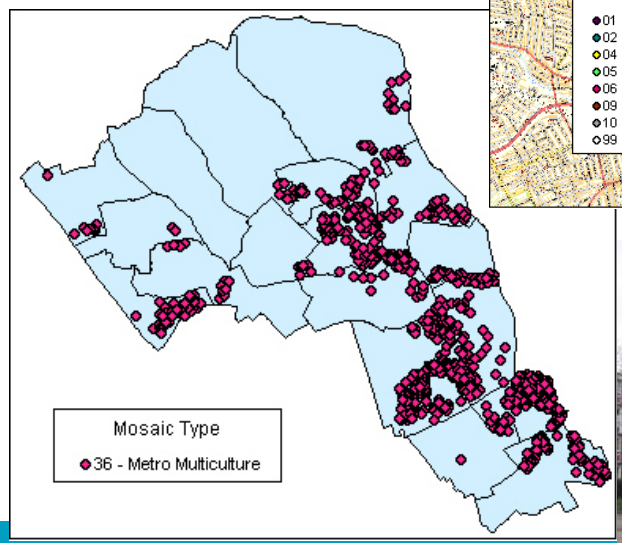
UCL

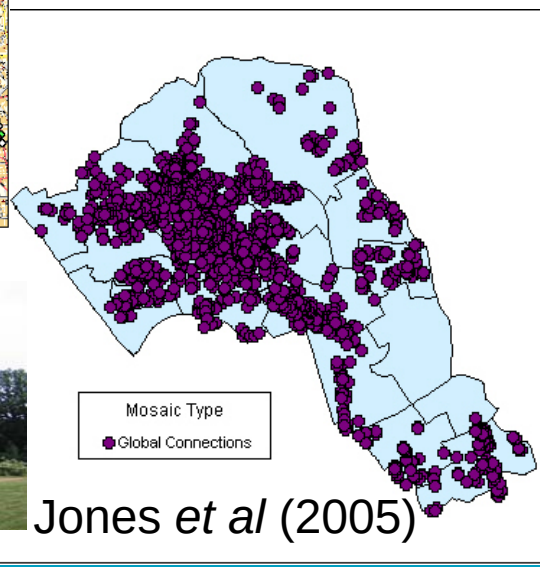
Geodemographic Groups in Camden

Metro Multiculture



Global Connections





Jones *et al* (2005)

Research Methods

- Develop a name-based ethnicity classification system:
  - At individual level (Surname + Forename)
  - With a rich taxonomy of Cultural, Ethnic & Linguistic (CEL) groups
- Use clustering methods to group CELs following different criteria, and other variables (birthplace)
- Measure Spatial segregation:
  - Area Indices – Postcode unit level
  - Network approach – Address level
- Evaluate methods with health inequalities data

Main Datasets Required

Scope	Dataset	Year	Spatial Resolution	Coverage Required	Name Data	Status
Population and Geodemographic data	Electoral Roll	2004	Unit Postcode	London	CEL	Obtained
	MOSAIC (Geodemographic classif.)	2004-06	Unit Postcode	London	CEL count	Obtained
	Census Key Statistics & Migration Data	1991 & 2001	ED/ Output Area	London		Obtained
	Neighbourhood Stats	2001-06	Super Output Area	London		Obtained
	ONS Longitudinal Study	1971-2001	GOR	UK		Evaluating
	Sample of Anonymised Records (SAR)	2001		UK		Evaluating
NHS Datasets	Patient Register	2004-06	Full Address	Camden + Islington?	Name	Obtained
	Birth & Death Registers	1999-06	Full Address	Camden	Name	Obtained
	Hospital Episode Statistics	1999-06	Full Address	Camden	Name	Obtained
	Local public health service uptake	2004-06	Full Address	Camden	Name	In progress
	Local Land & Property Gazetter	2004	Full Address	Camden		Obtained
Surveys	Health Survey for England	2002	Mosaic Type	UK		In progress
	TGI Consumer Survey	2003	Mosaic Type	UK	CEL?	Obtained
Name Resources	**Name-to-CEL database**	**2004**		**UK**		**Obtained**
	Telephone Directory		Full Address	Europe	Name	In progress
GI Infrastucture	Census & Administrative Geography	2001-06	Output Area	London		Obtained
	Street & Transport Network	2003	Street Segment	Camden		In progress
	Local Land & Property Gazetter	2004	Household	Camden		Obtained

CEL - Cultural Ethnic & Linguistic type

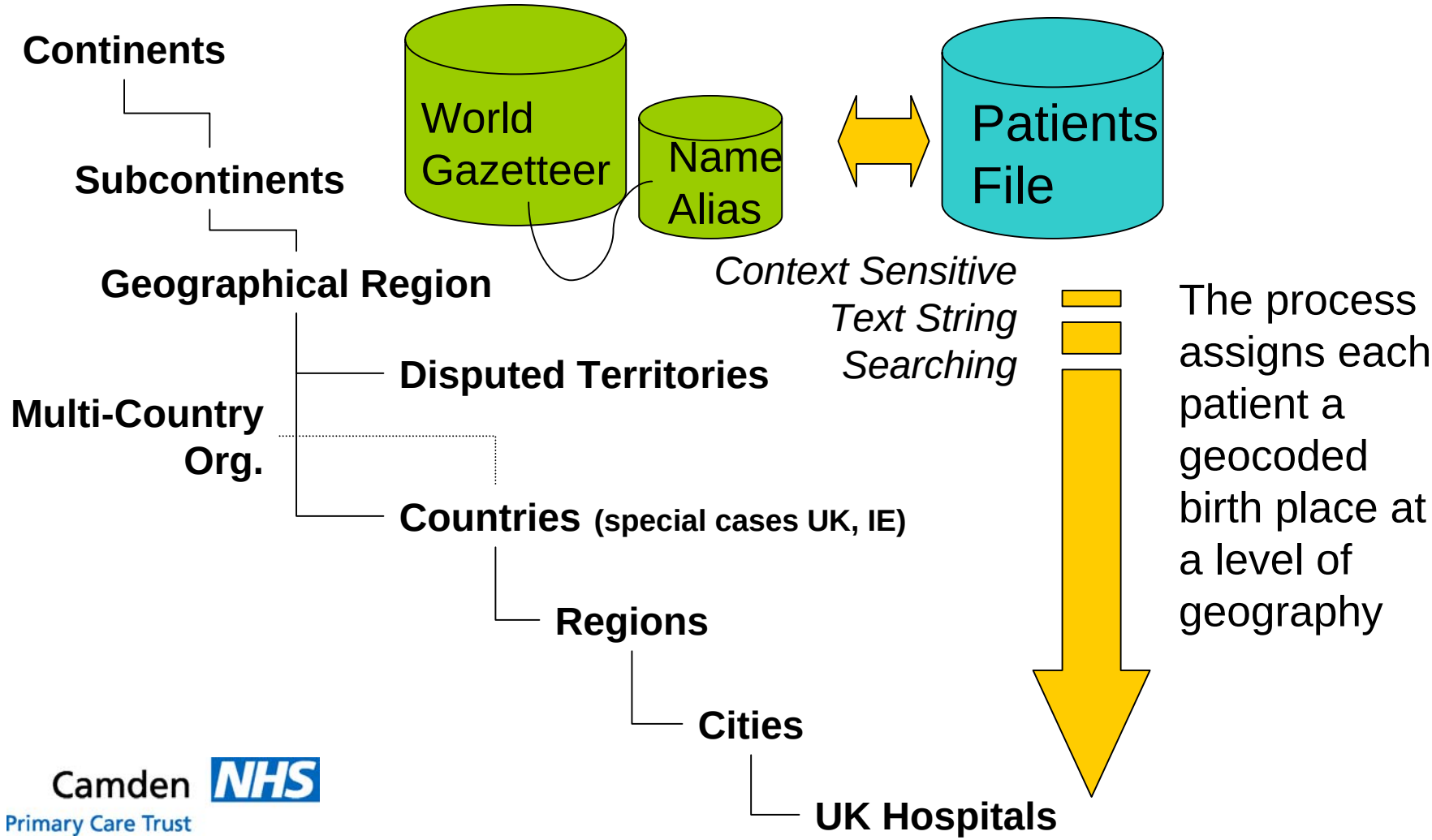
TGI- Target Group Index - A rolling consumer survey

5- Preliminary and Anticipated Results

Country of Birth Analysis

- Country of Birth has been widely used as a proxy for ethnicity (Webb et al 2004)
- Patient Registers have several advantages over Electoral Roll or other population registers
- Opportunity:
  - Underutilized “Birth Place” field in the NHS patient register (NHAIS Exeter)
  - Need to track Camden’s rapidly changing population born abroad

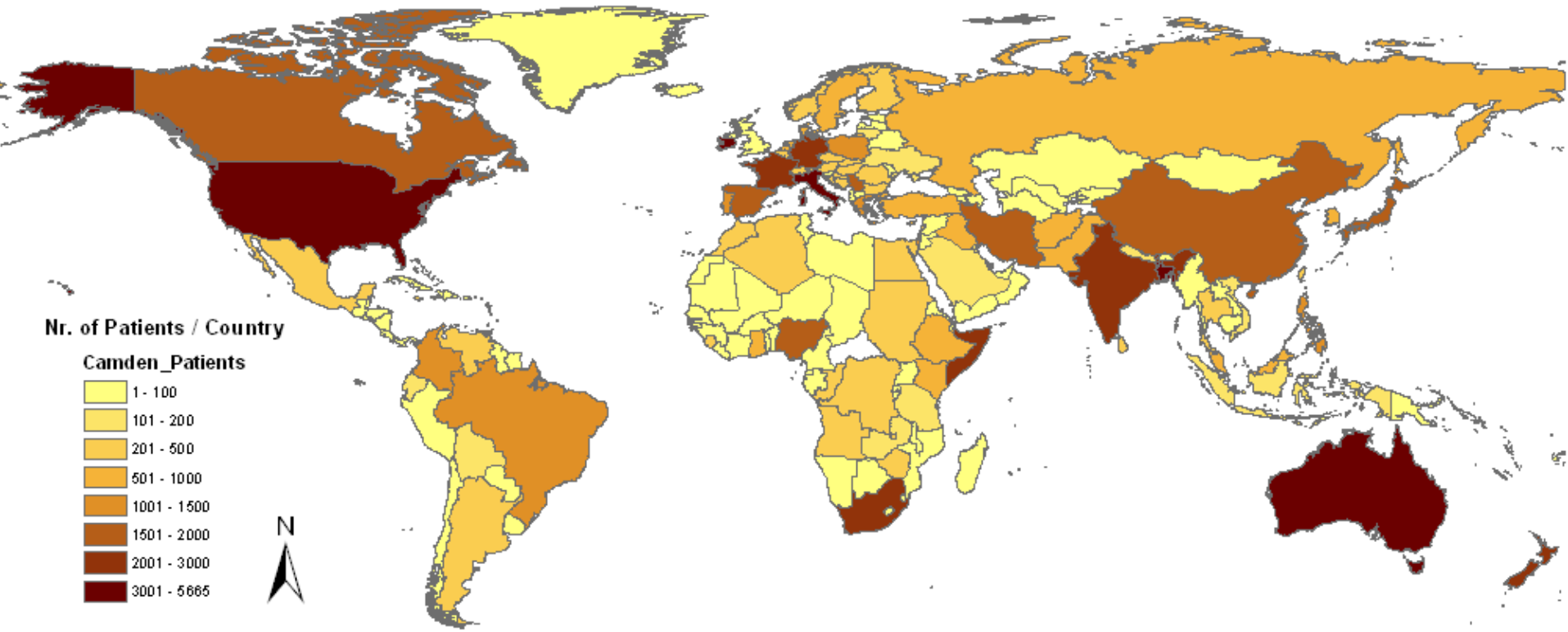
Birthplace Geocoder



Camden NHS Primary Care Trust

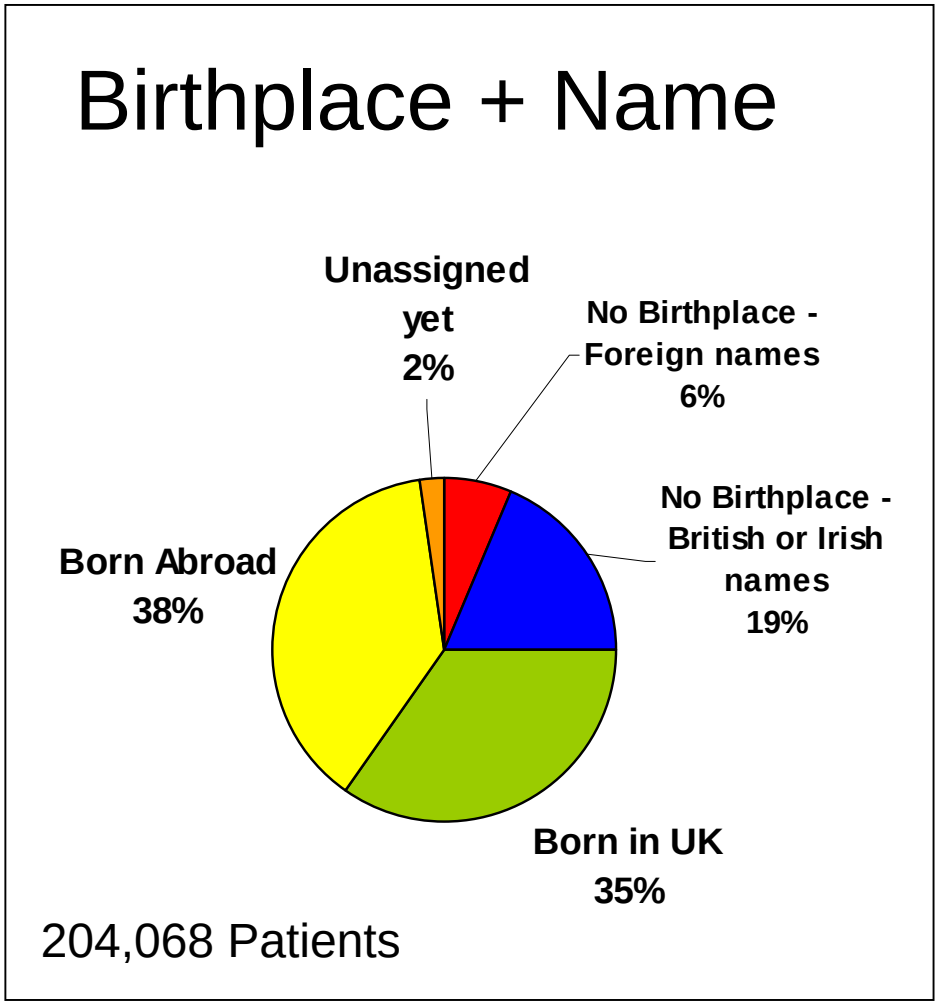
UCL

World map of Camden population



Camden 2004 population born abroad by country of birth

Camden PCT birthplaces & names

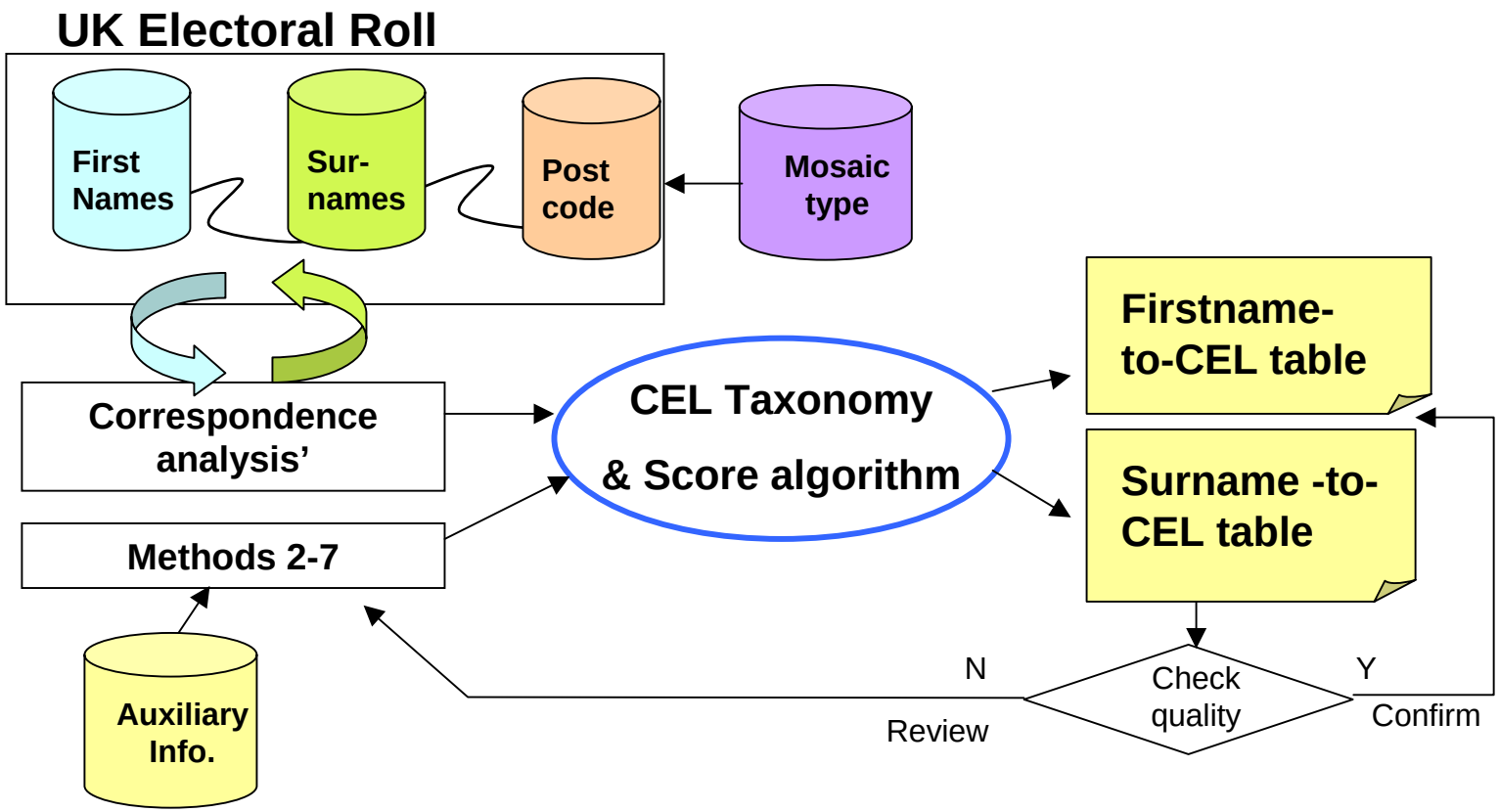


- Issues with COB
  - 2nd & 3rd generation immigrants
  - 'White British' born abroad.
  - Cascade migration

  (Senior & Bhopal 1994)

  - Patient records with no COB (25% in Camden)

Building a Name-to-CEL Diccionary



218,000 Surnames & 100,000 Firstnames coded to 128 CELs

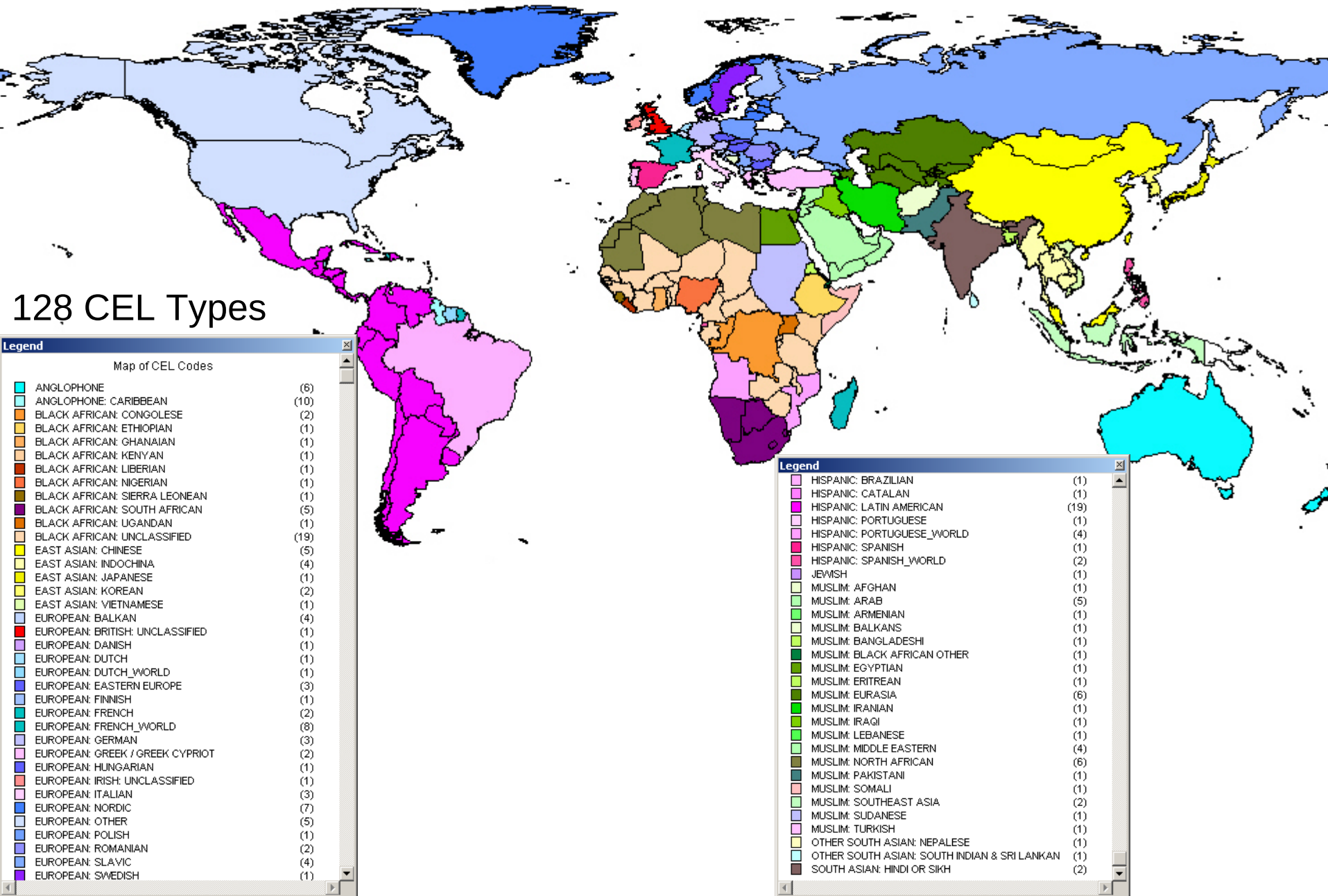
Main methods used to classify names

1. ‘Correspondence analysis’ between first names and surnames
2. Birthplace origin
3. Geodemographic name distribution (Mosaic)
4. Postcode geography
5. Text String mining
6. Lists of names by country from the web
7. ‘Googling’ individual names

Example of Name-to-CEL table

SURNAME	CEL-TYPE	Top Mosaic Type UK		Freq GB 1881	Freq GB 1998	GB 1881 Top area	GB 1996 Top area	British first name
WEINSTEIN	JEWISH;JEWISH	2	Cultural Leadership	22	156		NW	87.01
WOOLF	JEWISH;JEWISH	1	Global Connections	893	1700	E	NW	94.43
WEINER	JEWISH;JEWISH	1	Global Connections	25	260	WC	NW	90.26
WEISZ	JEWISH;JEWISH	2	Cultural Leadership	0	102		NW	73.33
GORSIA	JEWISH;JEWISH	1	Global Connections	19	218		HA	93.15
HALAI	JEWISH;JEWISH	1	Global Connections	18	161		HA	93.58
BUX	JEWISH;JEWISH	3	Corporate Chieftains	28	272	E	IG	88.12
JANJUA	JEWISH;JEWISH	1	Global Connections	146	635	EC	WC	85.88
SAMAD	Muslim;Bangladeshi	26	South Asian Industry	0	236		NW	28.67
HUQ	Muslim;Bangladeshi	29	City Adventurers	0	141		NW	36.63
BHOJANI	Muslim;Bangladeshi	26	South Asian Industry	1	421		E	34.15
KHALIL	Muslim;Bangladeshi	26	South Asian Industry	21	104		E	26.72
SAMAD	Muslim;Bangladeshi	26	South Asian Industry	0	216		E	16.80
KADRI	Muslim;Bangladeshi	26	South Asian Industry	0	115		E	22.76
KANBI	MUSLIM;Bangladeshi		#N/A	0	246		HA	15.15
MENDIS	Muslim;Bangladeshi	20	Asian Enterprise	2	373		HA	20.61
SALEM	MUSLIM;Egyptian	1	Global Connections	11	394		NW	62.41
KHATRI	MUSLIM;Egyptian	1	Global Connections	0	174		EC	52.00
BAH	MUSLIM;Egyptian	26	South Asian Industry	3	157		N	44.44
SHABBIR	Muslim;Egyptian	1	Global Connections	0	105		WC	74.65
BAPU	Muslim;Eritrean	26	South Asian Industry	0	316		IG	24.88

World map of CEL types



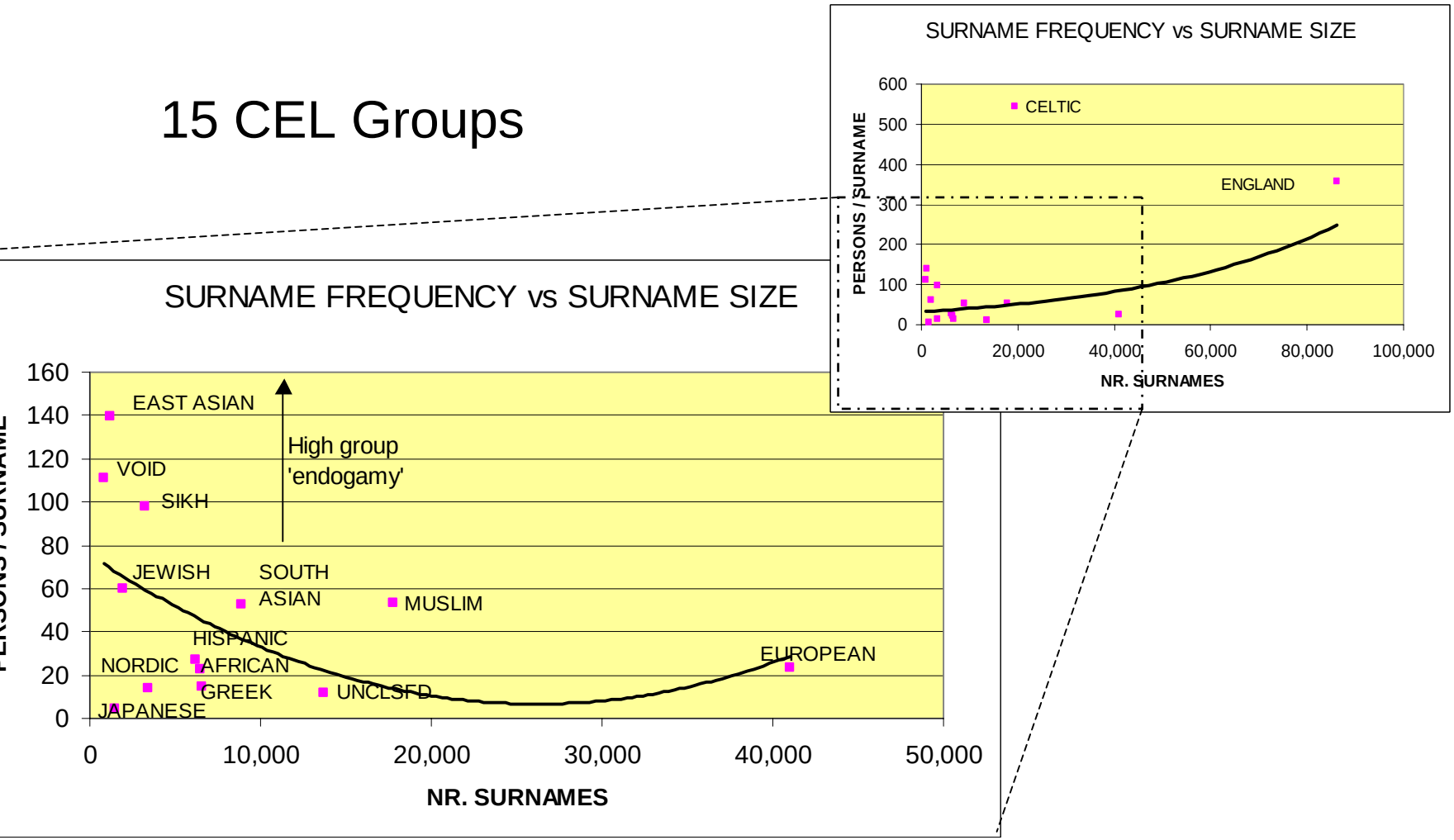
Summary of CEL Groups

128 CEL Types aggregated into 15 CEL Groups

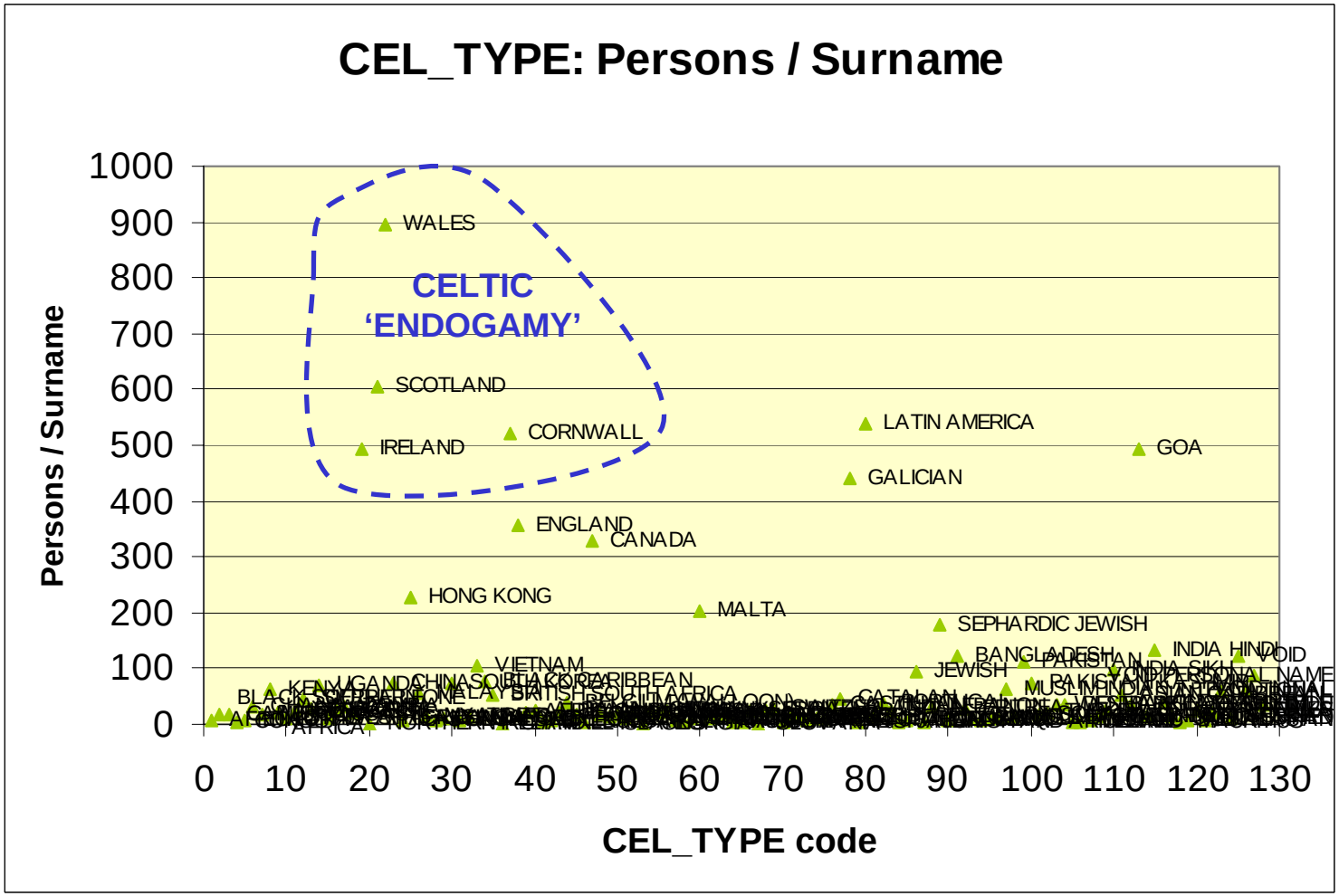
CEL_GROUP	NR. CEL_TYPES	SURNAMES	PERSONS	PERSONS/ SURNAME
ENGLAND	5	86,289	30,856,110	358
CELTIC	5	19,356	10,548,055	545
EUROPEAN	32	41,035	973,590	24
MUSLIM	14	17,758	952,146	54
SOUTH ASIAN	12	8,904	467,455	52
SIKH	1	3,237	316,337	98
EAST ASIAN	11	1,218	170,032	140
HISPANIC	10	6,180	169,258	27
UNCLASSIFIED	2	13,696	155,330	11
AFRICAN	17	6,441	144,540	22
JEWISH AND ARMEN	5	1,960	118,099	60
GREEK ORTHODOX	3	6,609	98,958	15
VOID	4	811	89,872	111
NORDIC	6	3,416	46,536	14
JAPANESE	1	1,482	6,322	4
TOTAL	**128**	**218,392**	**45,112,640**	**207**

Surname Frequency and Size

15 CEL Groups



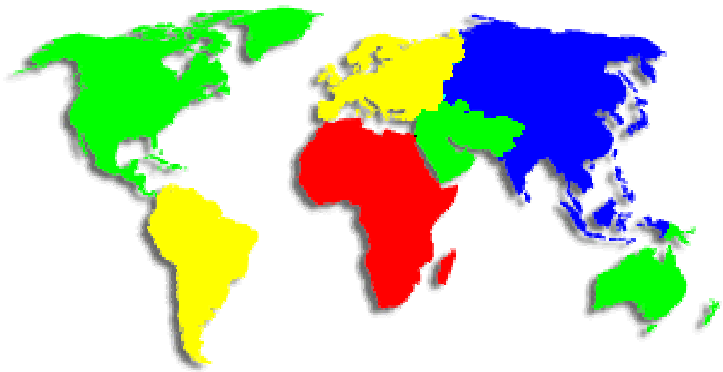
Surname Frequency and Size



Different criteria to aggregate CELs

Geographical Contiguity Vs Cultural Links

Geographical Regions

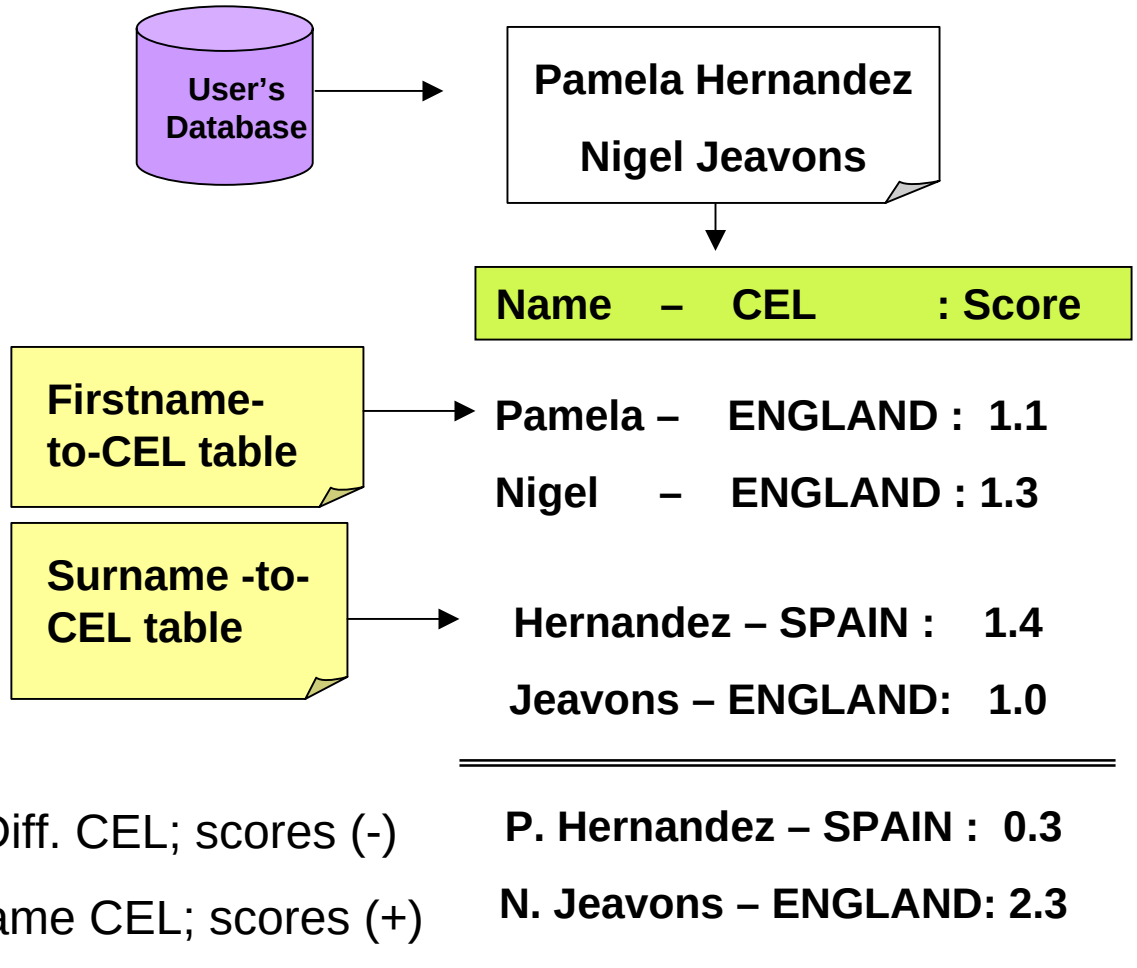


Excolonial Links

Predominant Religions

Major Languages

Coding Name Databases



- 85% of cases the 2 CELs in a name are the same.
- >99% hit rate using both when we might get only 90% hit rate using just one.

User can choose which score threshold suits their purpose, dropping weakest CEL assignments

Household ethnicity analysis

- Patient's Address Geocoded to a UPRN
  (Unique Property Reference Number from a Local Property Gazetteer)

UPRN	SURNAME	AGE	GENDER	COB
123456	Soandso	1	M	UNITED KINGDOM
123456	Soandso	5	F	UNITED KINGDOM
123456	Soandso	8	F	ALBANIA
123456	Soandso	33	F	ALBANIA
123456	Soandso	52	M	ALBANIA
654321	Z1	8	F	UNITED KINGDOM
654321	Z1	15	F	AUSTRALIA
654321	Z1	16	F	
654321	Z1	18	M	SUDAN
654321	Z2	40	F	SUDAN

Household Most Likely CEL

Albanian (3 out of 5)

Sudanese (2 out of 4)

Refining household ethnicity to CEL assignment

• Establish relationships between household members to propose a household structure that explains the CEL of the ancestors

- Differences in :
  - Age
  - Gender
  - Title
- Relationships between:
  - Surnames
  - COB
  - Name-to-CEL prior analysis
  - GP & Registration Date

- Further Analysis:
  - Children School Language
  - People /Household ratio
  - Inter-CEL relationships

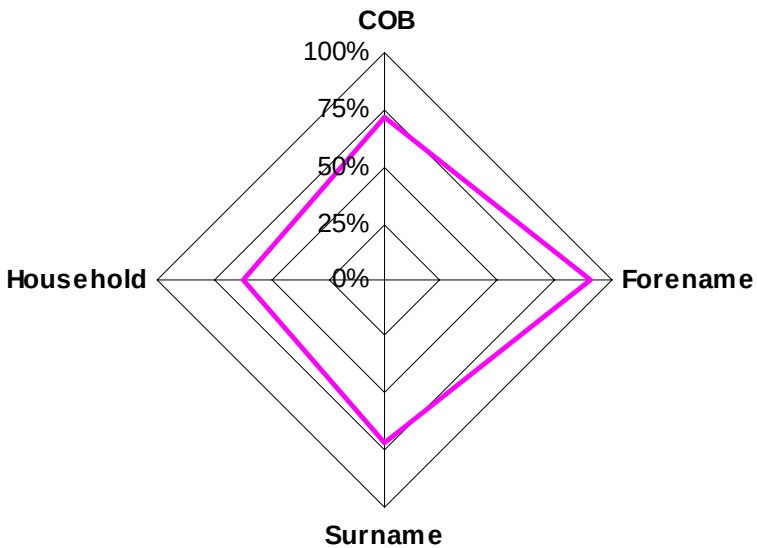
Compiling a CEL enhaced model

Compilation of potential CEL groups per person

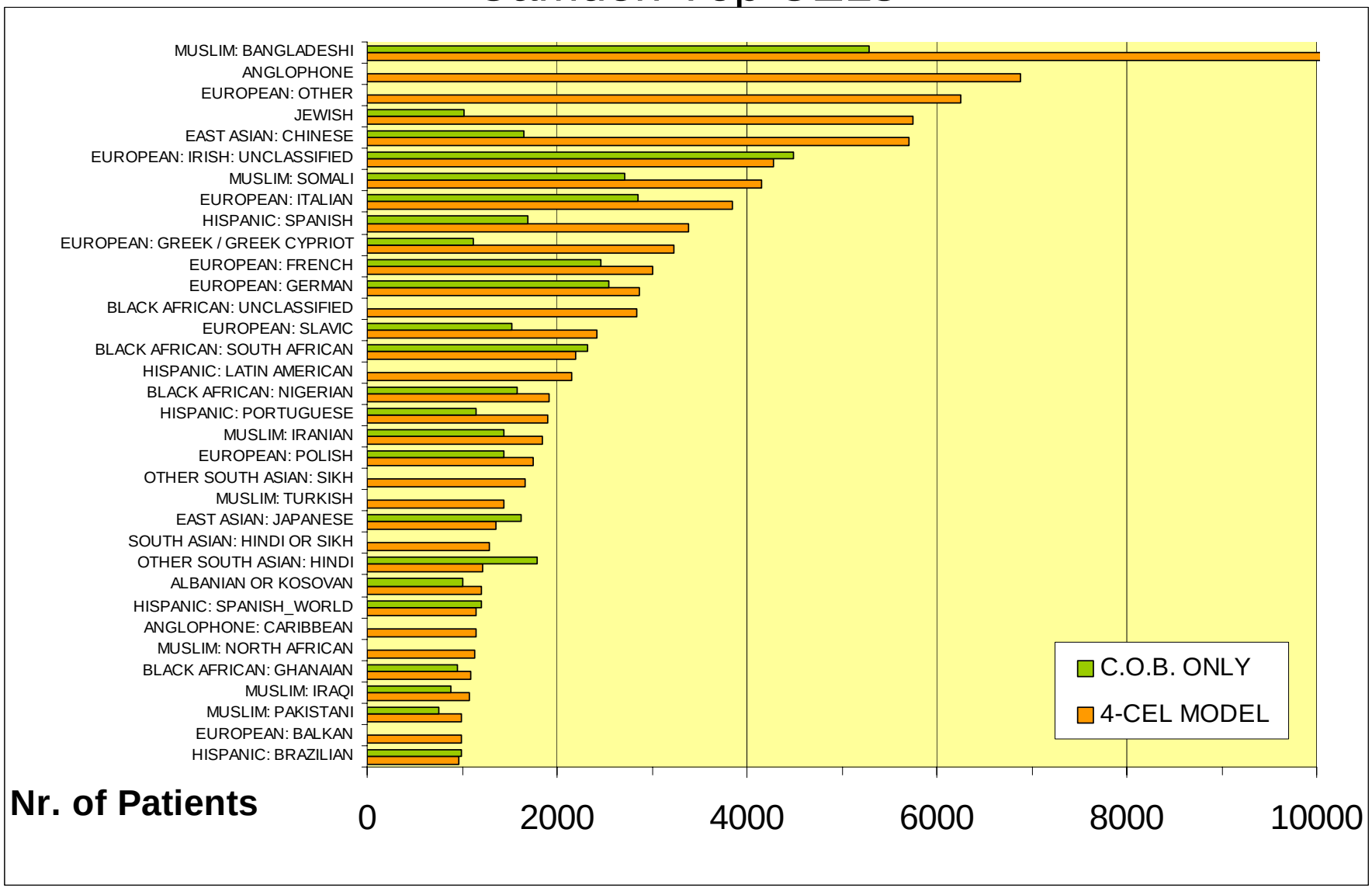
- Country of Birth
- Firstname & Surname
- Household

4-CELs coverage & match in Camden

CEL Matched	Total Available	Patients	%	% Cumm.
4	4	5234	2.6%	2.6%
3	4	8284	4.1%	6.6%
3	3	5421	2.7%	9.3%
2	4	40200	19.8%	29.1%
2	3	51330	25.2%	54.3%
2	2	7696	3.8%	58.1%
1	1	9128	4.5%	62.6%
1	2	41390	20.3%	82.9%
1	3	26228	12.9%	95.8%
1	4	8484	4.2%	100.0%



COB Vs CEL model



Evaluating the Model

- Evaluation of the CEL model through self-reported ethnicity from Hospital Episode Statistics
  - 40,714 patients (20% of total) matched to a unique true ethnic code (1991 Census categories)
  - Problem of bad quality HES data

		Actual Ethnicity from HES data											
Predicted by CEL		0	1	2	3	4	5	6	7	8	Total	Sensitivity	Specificity	PPV
0	White	**24,656**	624	652	331	88	23	388	46	2,499	29,307	0.92	0.67	0.84
1	Black - Caribbean	35	**147**	3	15	3			1	35	239	0.17	1.00	0.62
2	Black - African	385	44	**1,948**	174	47	11	22	5	438	3,074	0.67	0.97	0.63
3	Black - Other										0	0.00	1.00	
4	Indian	426	15	17	8	**333**	16	12	2	150	979	0.13	0.99	0.44
5	Pakistani	19	1	3		22	**75**	11		29	160	0.32	1.00	0.47
6	Bangladeshi	96	5	59	37	132	75	**2,672**	1	292	3,369	0.84	0.98	0.79
7	Chinese	126	2	12	2	6	1	1	**272**	94	516	0.73	0.99	0.53
8	Any other ethnic group	1,046	19	196	64	67	36	87	44	**1,511**	3,070	0.30	0.96	0.49
	Total	26,789	857	2,890	631	698	237	3,193	371	5,048	**40,714**			

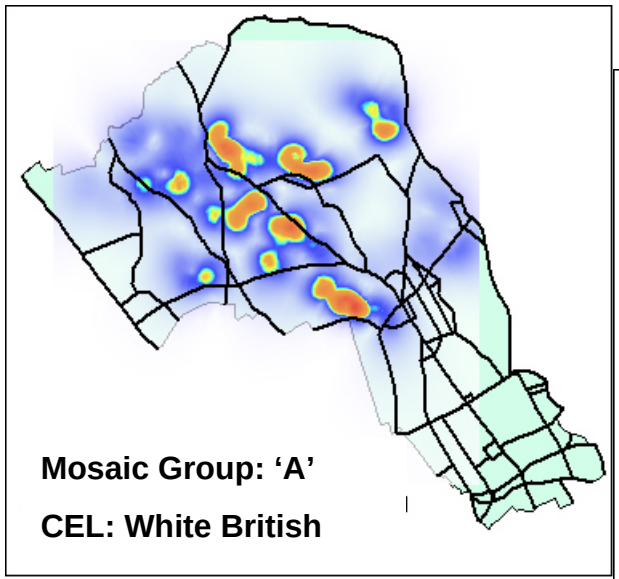
Evaluating the Model (II)

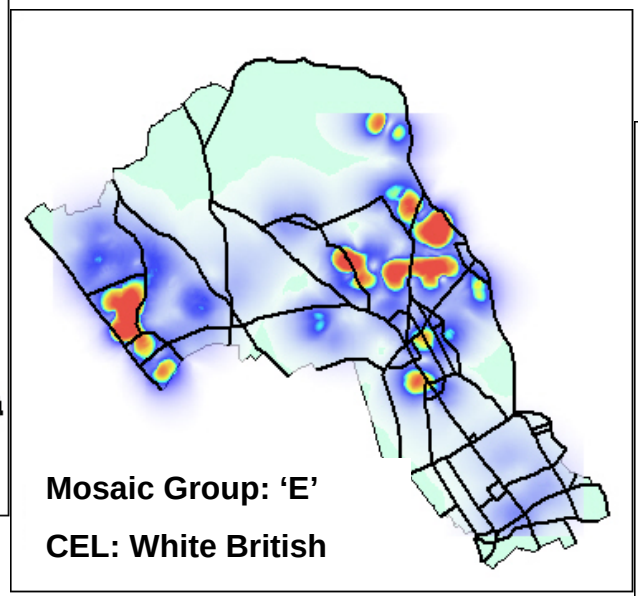
- National Evaluation against Census data

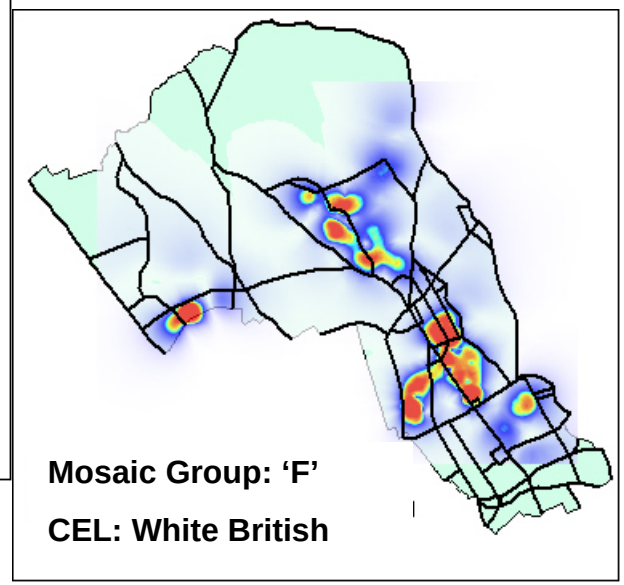
Census: Religion	Census: Ethnicity	Census: Country of Birth	% recognised family names	% all names (a)	% census (b)	Ratio (b/a)	Coverage
Hindu	Any	Any	1.01	1.07	0.98	**109.39**	Great Britain
Jewish	Any	Any	0.22	0.23	0.47	**49.79**	Great Britain
Muslim	Asian or Black African	Any	2.11	2.15	2.22	**96.63**	Great Britain
Sikh	Any	Any	0.72	0.71	0.59	**120.68**	Great Britain
Any	Black British or Black Afric	C, S, W Africa	0.28	0.33	0.32	**101.23**	UK
Any	White	Europe or Latin America	1.73	1.92	1.81	**105.91**	UK
Any	Chinese	Any	0.38	0.38	0.43	**88.84**	Great Britain
Any	Any	Bangladesh	0.43	0.43	0.50	**86.64**	Great Britain
Any	Any	India	1.66	1.76	1.84	**95.73**	Great Britain
Any	Any	Pakistan	1.33	1.35	1.31	**103.20**	Great Britain
Any	White, Mixed or Black Carib	Exc Europe or Latin Americ	93.74	93.08	92.39	**100.75**	UK
Any	Other	Other	6.26	6.92	7.61	**90.90**	UK
						(F/D)	
Any	Any	Scandinavia	0.08		0.09	**112.58**	Great Britain
Any	Any	Greece and Cyprus	0.19		0.19	**99.33**	Great Britain
Any	Any	Iberia and South America	0.29		0.40	**135.44**	Great Britain
		Rest of Europe	1.16		1.13	**97.10**	Great Britain

Non-responders to Breast Screening

White British Non-respondents by Mosaic type







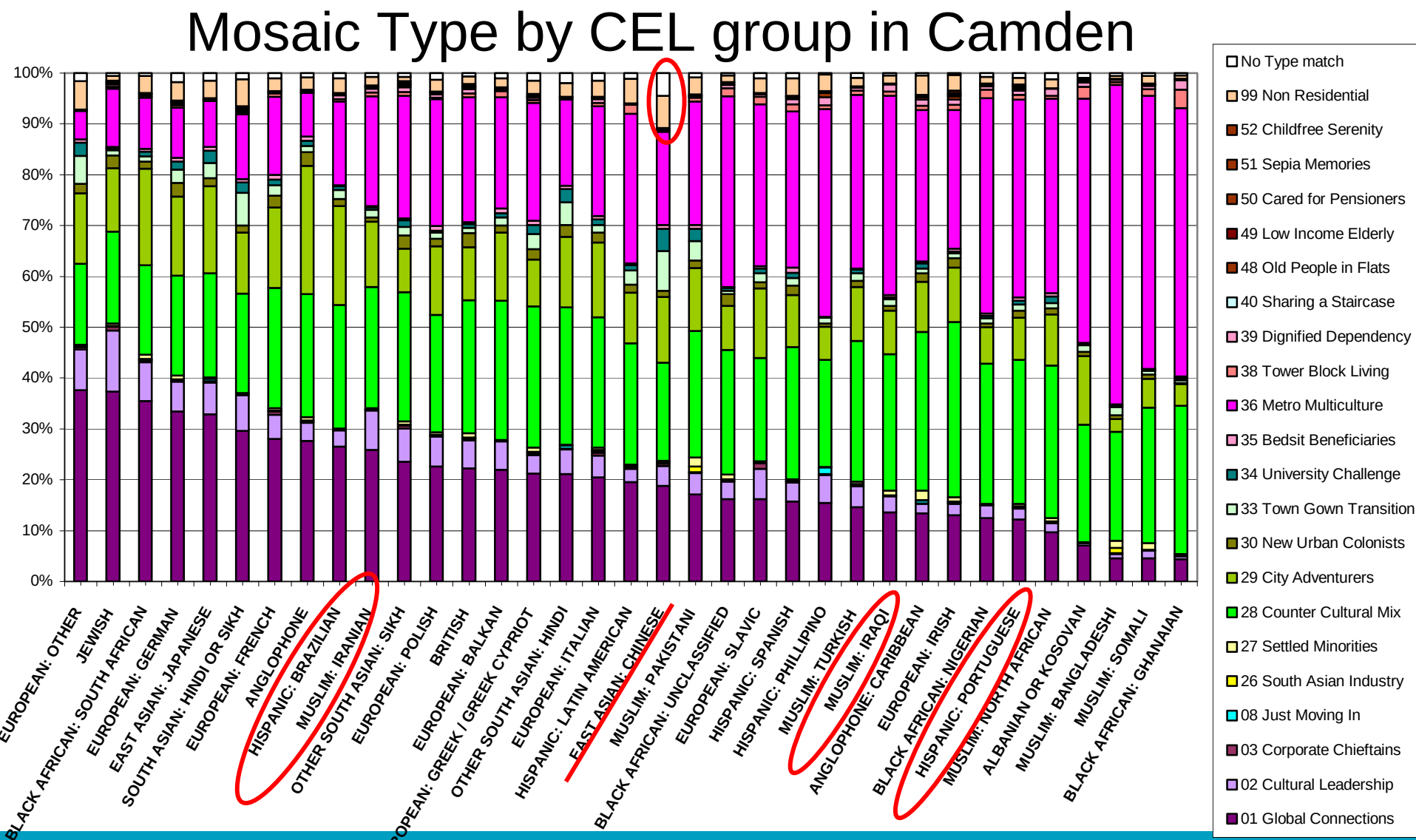
Most Affluent → **Least Affluent**

Camden NHS Primary Care Trust

Jones *et al* (2005)

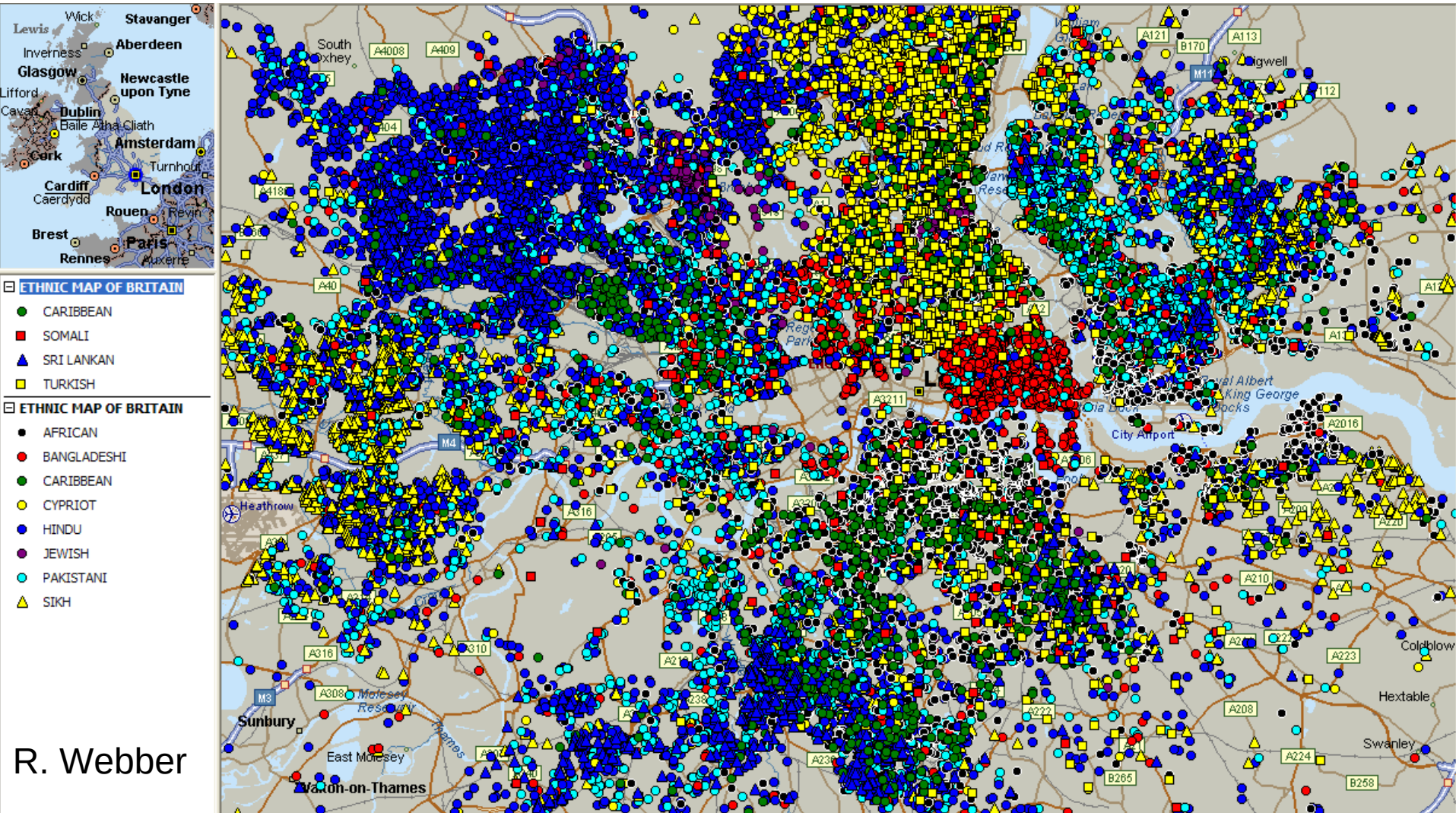
UCL

Ethnic Inequalities



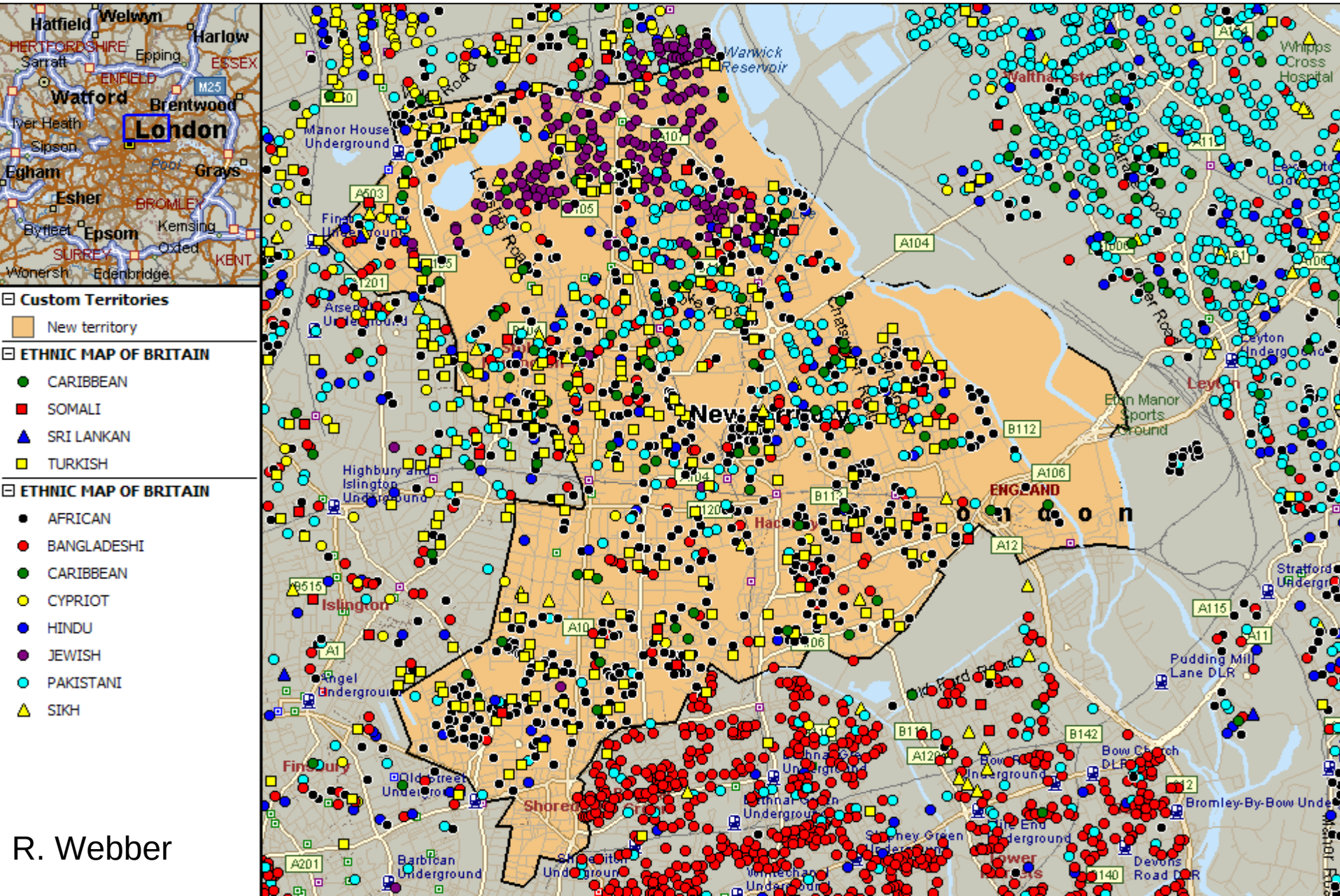
The ethnic map of London

Only postcodes with 8 or more foreign names are shown. Colours and symbols indicate the most frequently occurring minority based on name



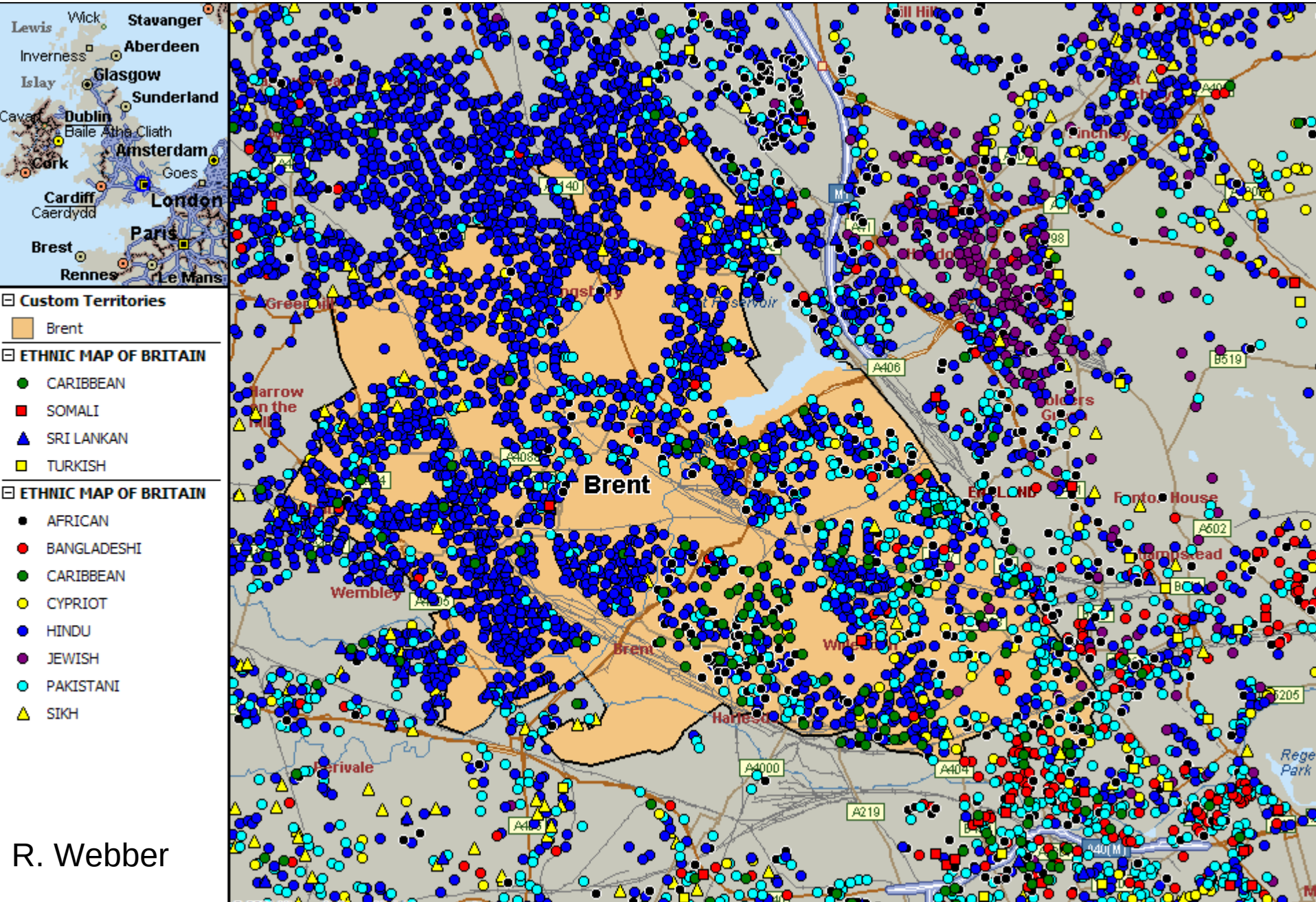
R. Webber

Hackney



R. Webber

Brent



R. Webber

Future enhancements to the ethnic classification model

- Improve household structure and overall model algorithms
- Expand name analysis
  - Introduce language and religion at subnational geographies
  - Introduce probabilistic and fuzzy CEL allocations
- Involve other users (currently working with several London PCTs)
  - Broaden the placename alias tables
  - Disseminate the methods & tools

    www.casa.ucl.ac.uk/geonom

6- Conclusion

Conclusion

- The geographies of ethnic inequalities in health cannot be understood with current:
  - Broad classifications of ethnicity
  - Coarse aggregated geographical units
- Spatial segregation processes are most likely hidden under those coarse units and closely linked to socieconomic factors
- This PhD will propose a new ontology of ethnicity based on determinining its different dimensions (CEL+)
- Methods will be developed and applied to ethnic health inequalities in Camden and London, to search for explanations at the individual level through CEL allocations and address level analysis

Thank you!
Any Questions?

www.casa.ucl.ac.uk/geonom