Positional cloning: statistical approaches to gene mapping, i.e. locating genes on the genome

- Linkage analysis
- Association studies (Linkage disequilibrium)
Linkage analysis

- Uses a genetic marker map (a map of polymorphic loci)

- Looks for co-segregation with a marker (polymorphic locus)

- **Simple Idea:**
  - To determine if marker allele at a known location travels with the disease in a family
Linkage Analysis

To identify the chromosomal region in which the disease-gene is located without knowing its function

- If in a family a disease D is transmitted associated with specific markers M, then the disease-gene mapped near these markers and D and M segregate together

Main aim of linkage analysis:

To evaluate the distance between D and M
Chromosomal region **non in-linkage** with the disease

Meiotic recombination is exploited to define the **small region in-linkage** with the disease
Linkage Analysis is based on **RECOMBINATION**

**Genetic mapping**
the aim is to discover **how often two loci are separated by meiotic recombination**

If two loci are **on different chromosome** they will segregate **independently**

Children will have 50% chance to receive each of these loci.
**Recombination fraction is** $\theta = 0.5$

If loci are on the **same chromosome** they are expected to **segregate together** ($\theta = 0$) but due to meiotic recombination this **does not always happen** ($0 < \theta < 0.5$)

The further two loci are on the chromosome the more they recombine

The $\theta$ is a **measure of the distance** between two loci
Recombination will rarely separate loci which lie very close ($\theta=0$)

Alleles on the same small chromosome segment tend to be transmitted as a block through a pedigree called **haplotype** -that can be tracked in families and populations

Recombination fraction defines **genetic distance**

The proportion of children who are recombinant is the recombination fraction between the two loci A and B during meiosis

How can we calculate the $\theta$?
Two loci which show 1% of recombination are defined as 1 centimorgan distance (cM)

The mathematical relationship between recombination fraction and genetic map distance is described by the mapping function

Haldane function \( d = 0.5 \ln (1 - 2\theta) \)

However there is interference during crossing-over since one chiasma can inhibit another...

Kosambi function \( d = 0.25 \ln \left(\frac{1 + 2\theta}{1 - 2\theta}\right) \)

Recombination map

Physical map

1 male cM = 0.9 Mb
1 female cM = 0.7 Mb
To perform linkage analysis you need **genetic markers**... 

**Characteristics:**

- highly polymorphic
- the rarest allele with a frequency of at least 1%
- feasible and stable in the pedigree
- well-known position in the genome
- genetic map of markers

**Genetic map of markers in 1980...**

**Human complete genetic map of microsatellite markers in 1992 by Cohen and colleagues**
Genetic linkage and disease

• Suitable large families are collected and segregation of the disease is compared with the segregation of the markers

• By using statistics it is tested the probability that the two loci (markers) are not in linkage (null hypothesis; threshold is p=0.05) in one family (LOD SCORE)

• Data from different families are collected and combined

**LOD SCORE** = is the logarithm of the odds that the loci are linked rather than unlinked

\[
LOD = Z = \log 10 \frac{\text{probability of birth sequence with a given linkage value}}{\text{probability of birth sequence with no linkage}}
= \log 10 \frac{(1 - \theta)^N \times \theta^R}{0.5(NR + R)}
\]

• It is a function of recombination fraction and is the product of the probabilities in each individual family

• When \( \theta \) is 0.5, lod score is 0

• \( Z = 3 \) is the threshold to accept linkage
The overall probability of linkage in a set of families is the addition of LOD scores in each family.

Linkage analysis can be more efficient if data for more than two loci are analyzed simultaneously.

Multipoint mapping
Linkage analysis and positional cloning: BRCA1

“The existence of BRCA1 was proven in 1990 by mapping predisposition to young-onset breast cancer in families to chromosome 17q21. Knowing that such a gene existed and approximately where it lays triggered efforts by public and private groups to clone and sequence it....BRCA1 was positionally cloned in September 1994....”

Mary Claire King

- identification of the more informative families
- identification of the locus for the susceptibility
Linkage analysis and positional cloning: BRCA1

- Families were analyzed with microsatellites spanning the region

- Lod scores were calculated for each family and add up for all the families

- Characterization of an open reading frame
**Parametric Linkage Analysis**

Parametric linkage analysis can be applied when there is a probability that a gene important for a disease is linked to a genetic marker.

It is studied using the LOD score, which assesses the probability that the disease and the marker are cosegregating.

It can be used when we have a pedigree with a clear type of inheritance and genotype-phenotype correlation.

**Non- Parametric Linkage Analysis**

Non-parametric linkage analysis studies the probability of an allele being identical by descent with itself.

It is used when the type of inheritance is not known.

Less powerful, but you can apply it to a lot of families.
Linkage analysis

• Great success in identifying genes for simple Mendelian diseases

• Few success in identifying genes contributing to complex disease

• Unsuccessful in identifying genes contributing to common complex disease
Linkage disequilibrium (LD)

• The nonrandom association of alleles in the population

• Alleles at neighboring loci tend to cosegregate

• Linkage disequilibrium implies population allelic association
Linkage Disequilibrium Mapping

- Population based
- Look for variant allele in LD with disease
- If most affected individuals in a population share the same mutant allele, then LD can be used to locate the chromosomal region harboring the disease
Case-Control Studies

Association studies: which allele of which gene is associated with the disease?

Case-Control Studies

- Common method in epidemiology
- Cases
- Controls from the same population
  - this implies that cases and controls should have similar genetic backgrounds
Linkage and association are different phenomena

- Association is a statistical statement about the co-occurrence of alleles or phenotypes (e.i. allele A is associated with disease D if people who have D also have more A). The association can have many possible causes (not all genetics).

- Linkage is a relationship between loci and does not of itself produce any association in the general population. Linkage creates association within families, but not among unrelated people.

However, if two supposedly unrelated people with disease D have inherited it from a distant common ancestor, they may well also tend to share particular ancestral alleles at loci closely linked to D (Linkage disequilibrium)

Statistical association can develop for different reason:
• direct cause-effect
• natural selection
• Stratification of the population
• linkage disequilibrium

Association studies are based on the use of haplotypes
Odds ratio = \frac{\text{odds allele } a \text{ in cases}}{\text{odds allele } a \text{ in controls}}

\text{Odds ratio}= \frac{A/B}{B/C} = A/B \times D/C = AD/BC

OR= 1 (no association); OR>1 the allele contributes to the disease
http://www.hapmap.org/index.html
About 10 million SNPs exist in human populations, where the rarer SNP allele has a frequency of at least 1%.

Researchers trying to discover the genes that affect a disease, such as diabetes, will compare a group of people with the disease to a group of people without the disease.

Chromosome regions where the two groups differ in their haplotype frequencies might contain genes affecting the disease.

Theoretically, researchers could look for these regions by genotyping 10 million SNPs. However, the methods to do this are currently too expensive.

The HapMap identifies which 200,000 to 1 million tag SNPs provide almost as much mapping information as the 10 million SNPs.

This substantial cost reduction makes such studies feasible to do.
“Sporadic” CRC

• is a multifactorial (complex) condition
  – environmental factors
  – genetic factors

• study of twins: 35% of all CRC cases have a genetic component
• first-degree relatives of CRC patients are well-recognized to have a 2- to 4-fold increased risk of developing the disease
  – recessive genes?
  – pathogenic mutations of low penetrance
  – complex gene-gene and gene-environment interactions
Tenesa et al. *Genome-wide association scan identifies colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21*

Tomlinson et al. *A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3*
Nature Genetics 40, 623 - 630 (2008)
In a genome-wide association study to identify loci associated with colorectal cancer (CRC) risk, we genotyped 555,510 SNPs in 1,012 early-onset Scottish CRC cases and 1,012 controls (phase 1). In phase 2, we genotyped the 15,008 highest-ranked SNPs in 2,057 Scottish cases and 2,111 controls. We then genotyped the five highest-ranked SNPs from the joint phase 1 and 2 analysis in 14,500 cases and 13,294 controls from seven populations, and identified a previously unreported association, rs3802842 on 11q23 (OR = 1.1; P = 5.8 times 10-10), showing population differences in risk. We also replicated and fine-mapped associations at 8q24 (rs7014346; OR = 1.19; P = 8.6 times 10^-26) and 18q21 (rs4939827; OR = 1.2; P = 7.8 times 10^-28).

Carrying all six possible risk alleles yielded OR = 2.6 (95% CI = 1.75-3.89) for CRC. These findings extend our understanding of the role of common genetic variation in CRC etiology.

Tenesa et al. Nat. Genet., 2008
Genetic factors play a role in cancer

- Epidemiology studies show that relatives of cancer patients have a higher risk of developing the disease

<table>
<thead>
<tr>
<th>Cancer site</th>
<th>1(^{st}) relatives</th>
<th>2(^{nd}) relatives</th>
<th>3(^{rd}) relatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thyroid</td>
<td>3.02</td>
<td>1.64</td>
<td>1.13</td>
</tr>
<tr>
<td>Kidney</td>
<td>2.3</td>
<td>1.31</td>
<td>1.32</td>
</tr>
<tr>
<td>Breast</td>
<td>2.02</td>
<td>1.36</td>
<td>1.21</td>
</tr>
<tr>
<td>Lung</td>
<td>2</td>
<td>1.39</td>
<td>1.1</td>
</tr>
<tr>
<td>Prostate</td>
<td>1.89</td>
<td>1.36</td>
<td>1.19</td>
</tr>
</tbody>
</table>
Only a fraction of cancer cases have family history of the disease

- **Cancer syndromes (1-2%)**
  - Rare, highly penetrant mutations e.g. p53 in Li-Fraumeni syndrome

- **Familial cases (10-15%)**
  - Mutations of intermediate penetrance, e.g. HNPCC, BRCA1/2

- **Sporadic cases (>80%)**
  - Family history not notable
  - Genetic factors still play a role
How much of cancer risk is explained by genetics?

- Thyroid: 53%
- Breast: 27%
- Ovary: 22%
- Colorectal: 35%
- Lung: 8%
- Prostate: 42%

Lichtenstein et al., NEJM 2000
Czen et al., Int J Cancer 2002
Cross-risk of cancer in relatives

Icelandic Cancer Registry

Genealogy of all Icelanders

Can we find genes that explain this picture? “Multi-cancer” genes

Amundadottir et al PLOS Med 2004
Icelandic history

Iceland founded in 9th century by settlers of mixed Northern European descent

Current Pop. 300,000

- N-European descent
- Isolation for 11 centuries
- Same genetic background
- Smaller number of mutations
First major cancer project - prostate cancer

- A major public health problem
  - The most common cancer in males in the US
  - Lifetime risk 10% in EU - 16% USA
  - The second leading cause of cancer related deaths in men

- A genetic enigma
  - Genetic component one of the largest of all cancers
  - No highly-penetrant cancer genes isolated that can explain the familiality
  - Common polymorphisms in numerous genes reported to be associated with risk - hard to replicate
### Results on 8q24 replicated

<table>
<thead>
<tr>
<th>Study population</th>
<th>Marker</th>
<th>Allele</th>
<th>Cases</th>
<th>Controls</th>
<th>OR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Iceland</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1291/997)</td>
<td>DG8S737</td>
<td>-8</td>
<td>0.131</td>
<td>0.078</td>
<td>1.77</td>
<td>2.3x10^{-8}</td>
</tr>
<tr>
<td></td>
<td>rs1447295</td>
<td>A</td>
<td>0.169</td>
<td>0.106</td>
<td>1.72</td>
<td>1.7x10^{-9}</td>
</tr>
<tr>
<td><strong>Sweden</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1435/779)</td>
<td>DG8S737</td>
<td>-8</td>
<td>0.101</td>
<td>0.079</td>
<td>1.38</td>
<td>4.3x10^{-3}</td>
</tr>
<tr>
<td></td>
<td>rs1447295</td>
<td>A</td>
<td>0.164</td>
<td>0.133</td>
<td>1.29</td>
<td>4.5x10^{-3}</td>
</tr>
<tr>
<td><strong>European Americans</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Chicago</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(458/247)</td>
<td>DG8S737</td>
<td>-8</td>
<td>0.082</td>
<td>0.041</td>
<td>2.10</td>
<td>2.9x10^{-3}</td>
</tr>
<tr>
<td></td>
<td>rs1447295</td>
<td>A</td>
<td>0.127</td>
<td>0.081</td>
<td>1.66</td>
<td>6.7x10^{-3}</td>
</tr>
<tr>
<td><strong>African Americans</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Michigan</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(246/352)</td>
<td>DG8S737</td>
<td>-8</td>
<td>0.234</td>
<td>0.161</td>
<td>1.60</td>
<td>2.2x10^{-3}</td>
</tr>
<tr>
<td></td>
<td>rs1447295</td>
<td>A</td>
<td>0.344</td>
<td>0.313</td>
<td>1.15</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Alleles for the markers DG8S737 and rs1447295 at 8q24.21 are shown and the corresponding numbers of cases and controls (N), allelic frequencies of variants in affected and control individuals, the odds-ratio (OR) and two-sided P values.
GWA identifies a second signal on 8q24

Gudmundsson et al., Nat.Genetics, 2007
Several independent prostate cancer risk loci on chr8q24

A common variant associated with prostate cancer in European and African populations

Genome-wide association study of prostate cancer identifies a second risk locus at 8q24

Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24

Multiple newly identified loci associated with prostate cancer susceptibility
GWA studies have identified >16 loci involved in prostate cancer

Eeles et al. 2008
Gudmundsson et al 2008
Haiman et al 2007
# 16 prostate cancer variants...

| Prostate cancer | 2p15 | 3p12 | 6q25 | 7q21 | JAZF1 | 8q24 | 8q24 | 8q24 | HNF1B | HNF1B | 17q | MSMB | CTBP2 | 11q13 | KLK2/KLK3 | Xp11 | rs721048 | rs2660753 | rs9364554 | rs6465657 | rs10486567 | rs1447295, DG8S737 | rs6983267 | rs16901979, hapC | rs4430796 | rs11649743 | rs1859962 | rs10993994 | rs4962416 | rs931342 | rs2735839 | rs5945619 | $8 \times 10^{-9}$ | $3 \times 10^{-8}$ | $6 \times 10^{-10}$ | $10^{-9}$ | $10^{-7}$ | $3 \times 10^{-11}$ | $9 \times 10^{-13}$ | $3 \times 10^{-15}$ | $10^{-11}$ | $2 \times 10^{-9}$ | $3 \times 10^{-10}$ | $9 \times 10^{-29}$ | $3 \times 10^{-8}$ | $2 \times 10^{-12}$ | $2 \times 10^{-18}$ | $2 \times 10^{-9}$ |
Cancer risk variants on chr8q24

Most variants specific for a particular cancer
Does not explain cross-risk of cancers
Genetic risk assessment model for prostate cancer

- Genotype 13 variants
  - Multiplicative model
  - Results presented as
    - Relative risk of developing the disease
    - Lifetime risk (average 10% in the EU)
- Does not include family history
Ways to find cancer risk variants

• Perform a GWA study on a large number of cancer cases and controls
  – Power depends on frequency of variant and OR
  – Need thousands of cancer cases to find variants with OR<1.2

• Risk factors for cancer may also be genetic
  – e.g. pigmentation and risk of skin cancer
  – Sample sizes often very large
Results from prostate cancer risk test

Fraction of the Population

Relative genetic risk

0.76

Top 5%

RRR 3%

OR > 2

Lifetime risk >20%lifetime risk >30%

deCODE genetics
Smoking behavior and lung cancer

- Smoking is the major risk factor for lung cancer
  - Over 90% of cases in males and 80% of cases in females attributed to smoking
- Evidence for genetic influence on smoking behaviour and nicotine addiction
- Genetic studies on smoking....
Studies on smoking behaviour

- 11,000 smokers; divided into 4 groups based on the number of cigarettes per day (cpd)
  1-10
  11-20
  21-30
  31 or more
- Genotyped for 370,000 SNPs
  - Search for variants that are more common in heavier smokers
Variants in nicotine receptor cluster associate with more smoking

6 SNPs in the nicotinic acetylcholine receptor gene cluster on chromosome 15q
Associate with more smoking and nicotine addiction

rs1051730 (T)
Genetics of smoking

- Results confirmed in an independent group of smokers from
  - Iceland (2950)
  - The Netherlands (1375)
  - Spain (523)
- For all groups combined, regression analysis adjusted for gender and age, $P=6 \times 10^{-20}$.
  - Each copy of the “risk” variant increases smoking by 1 cigarette per day
- Also associated with nicotine addiction

Thorgeirsson et al Nature 2008
Association between rs1051730(T) and number of cigarettes per day

Thorgeirsson et al Nature 2008
rs1051730(T) associates with risk of lung cancer

Table 4: Association of rs1051730 allele T with Lung Cancer and PAD

<table>
<thead>
<tr>
<th>Study Group</th>
<th>Controls</th>
<th>Cases</th>
<th>OR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>freq</td>
<td>n</td>
<td>freq</td>
<td></td>
</tr>
<tr>
<td>Lung Cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iceland</td>
<td>28,752</td>
<td>0.342</td>
<td>565</td>
<td>0.398</td>
<td>1.27</td>
</tr>
<tr>
<td>Spain</td>
<td>1,474</td>
<td>0.390</td>
<td>269</td>
<td>0.483</td>
<td>1.46</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>2,018</td>
<td>0.314</td>
<td>90</td>
<td>0.350</td>
<td>1.18</td>
</tr>
<tr>
<td>Foreign combined</td>
<td>3,492</td>
<td>-</td>
<td>359</td>
<td>-</td>
<td>1.38</td>
</tr>
<tr>
<td>All combined</td>
<td>32,244</td>
<td>-</td>
<td>1,024</td>
<td>-</td>
<td>1.31</td>
</tr>
</tbody>
</table>

Thorgeirsson et al Nature 2008
Is the increase in lung cancer risk only through effect on smoking?

- Increase in cpd too small to explain the full lung cancer risk
  - Direct effect of nicotine in the lung?
- Nicotin receptors are expressed in many tissues, including lung epithelium
  - stimulations causes proliferation and malignant transformation
- No increase in risk in non-smokers
- Variants not significantly associated with other cancer types
  - Not even bladder cancer......
Finally, a variant that affects risk of many types of cancer!

- GWA study on basal cell carcinoma (BCC) identified several regions that associate with increased risk of skin cancer (Stacey et al 2008)
- One on chr5p near two known “cancer genes”
  - *CLPTM1L* (cisplatin resistance related protein) gene
  - *hTERT* (human telomerase reverse transcriptase) gene
TERT plays a role in the progression of most forms of cancer

- Examine if variation in this region is associated with risk of other cancer types
- Test cancer at 17 cancer sites, using 30,000 cancer cases and 45,000 controls from Iceland, Europe and USA
### rs401681 (C) associates with risk of cancer at 5 sites

<table>
<thead>
<tr>
<th>Study population</th>
<th>Number</th>
<th>Frequency</th>
<th>OR</th>
<th>95% CI</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases</td>
<td>Controls</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basal cell carcinoma</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iceland all</td>
<td>2,040</td>
<td>28,890</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eastern Europe</td>
<td>525</td>
<td>515</td>
<td>0.616</td>
<td>0.575</td>
<td>1.16</td>
</tr>
<tr>
<td>All combined</td>
<td>2,565</td>
<td>515</td>
<td>0.610</td>
<td>0.650</td>
<td>1.25</td>
</tr>
<tr>
<td>Lung cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iceland all</td>
<td>1,449</td>
<td>28,890</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Netherlands</td>
<td>529</td>
<td>1,832</td>
<td>0.610</td>
<td>0.570</td>
<td>1.18</td>
</tr>
<tr>
<td>Spain</td>
<td>367</td>
<td>1,427</td>
<td>0.582</td>
<td>0.538</td>
<td>1.19</td>
</tr>
<tr>
<td>IARC</td>
<td>1,920</td>
<td>2,517</td>
<td>0.617</td>
<td>0.586</td>
<td>1.14</td>
</tr>
<tr>
<td>All combined</td>
<td>4,265</td>
<td>34,666</td>
<td>0.596</td>
<td>0.560</td>
<td>1.15</td>
</tr>
<tr>
<td>Bladder cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iceland all</td>
<td>780</td>
<td>28,890</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Netherlands</td>
<td>1,277</td>
<td>1,832</td>
<td>0.584</td>
<td>0.570</td>
<td>1.06</td>
</tr>
<tr>
<td>UK</td>
<td>707</td>
<td>506</td>
<td>0.564</td>
<td>0.514</td>
<td>1.23</td>
</tr>
<tr>
<td>Italy-Torino</td>
<td>329</td>
<td>379</td>
<td>0.550</td>
<td>0.545</td>
<td>1.02</td>
</tr>
<tr>
<td>Italy-Brescia</td>
<td>122</td>
<td>156</td>
<td>0.574</td>
<td>0.564</td>
<td>1.04</td>
</tr>
<tr>
<td>Belgium</td>
<td>199</td>
<td>378</td>
<td>0.603</td>
<td>0.554</td>
<td>1.22</td>
</tr>
<tr>
<td>Eastern Europe</td>
<td>214</td>
<td>515</td>
<td>0.619</td>
<td>0.575</td>
<td>1.20</td>
</tr>
<tr>
<td>Sweden</td>
<td>346</td>
<td>905</td>
<td>0.545</td>
<td>0.521</td>
<td>1.10</td>
</tr>
<tr>
<td>Spain</td>
<td>173</td>
<td>1,427</td>
<td>0.546</td>
<td>0.538</td>
<td>1.03</td>
</tr>
<tr>
<td>All combined</td>
<td>4,147</td>
<td>34,988</td>
<td>0.578</td>
<td>0.535</td>
<td>1.12</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iceland all</td>
<td>2,276</td>
<td>28,890</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Netherlands</td>
<td>994</td>
<td>1,832</td>
<td>0.576</td>
<td>0.570</td>
<td>1.02</td>
</tr>
<tr>
<td>Chicago, US</td>
<td>635</td>
<td>693</td>
<td>0.581</td>
<td>0.568</td>
<td>1.06</td>
</tr>
<tr>
<td>Spain</td>
<td>459</td>
<td>1,427</td>
<td>0.559</td>
<td>0.538</td>
<td>1.09</td>
</tr>
<tr>
<td>COGEMS</td>
<td>5,109</td>
<td>5,059</td>
<td>0.558</td>
<td>0.543</td>
<td>1.06</td>
</tr>
<tr>
<td>All combined</td>
<td>9,473</td>
<td>37,901</td>
<td>0.569</td>
<td>0.553</td>
<td>1.07</td>
</tr>
<tr>
<td>Cervical cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iceland all</td>
<td>369</td>
<td>28,890</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.611</td>
<td>0.545</td>
<td>1.31</td>
<td>1.13-1.51</td>
<td>2.6x10^-4</td>
</tr>
</tbody>
</table>