5. The rejection region: Reject $H_0: \rho = 0$, if $t$ is not between $\pm t_{0.975}(320) \approx 2.00$

6. The result:

\[ t = 0.83 \sqrt{\frac{320}{1 - 0.83^2}} = 0.83 \sqrt{\frac{320}{1 - 0.69}} \]

\[ t = 0.83 \sqrt{\frac{320}{0.31}} = 0.83 \sqrt{1032.26} \]

\[ t = 0.83(32.13) = 26.67 \]

7. The conclusion:

Reject $H_0: \rho = 0$

$t = 26.67$ is not between
Module 20: Correlation

This module focuses on the calculating, interpreting and testing hypotheses about the Pearson Product Moment Correlation Coefficient.
Correlation

In module 19, we examined how two variables, x and y, relate to each other by using the simple linear regression tool. In that context, x is the independent variable and y is the dependent variable. Typical examples for the independent variable include measures of time, including age; whereas, typical examples for the dependent variable are continuous measurements such as blood cholesterol level. The general assumption is that there are separate normal distributions of the dependent variable y for each value of the independent variable x. Further, we need to assume that these separate normal distributions for the dependent variable all have the same population variance.
Clearly these assumptions are quite restrictive in that we are often interested in the relationship between two variables, x and y, where it is not at all clear which should be labeled the independent variable and which the dependent one. An example is the relationship between blood cholesterol level and blood pressure level.
For this situation, we have another tool we can use to measure and test hypotheses about the relationship between these two variables. The tool is called correlation and we focus here only on what is usually called the Pearson Product Moment Correlation Coefficient. There are other measures of correlation which we will not discuss here. There are restrictions for the use of this correlation tool as well, which include the basic assumption that the $x$ and $y$ variables together have a joint frequency distribution which is called the bivariate normal distribution. This distribution looks like a three-dimensional bell in a manner similar to the way a normal distribution for one variable looks like a cross section of a bell.
The degree of association or correlation between two variables is measured by the correlation coefficient. This is done in a manner similar to that for other population parameters and estimates of these parameters obtained by calculating statistics from samples. That is, there is a value for the population parameter for this coefficient which is estimated by selecting a random sample and calculating the appropriate coefficient using the data from this sample. We can also use the information from the sample to test hypotheses about the population.
The population parameter for the Pearson Product Moment Correlation Coefficient is defined as

\[ \rho = \frac{\sum (x-u_x)(y-u_y)}{\sqrt{\sum (x-u_x)^2 \sum (y-u_y)^2}} \]

which is typically called rho, for the Greek letter it represents.

The estimate of \( \rho \) calculated from the sample data is the statistic

\[ r_w = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sqrt{\sum (x-\bar{x})^2 \sum (y-\bar{y})^2}} \]

\[ r_w = \frac{\sum x y - (\sum x)(\sum y)/n}{\sqrt{[\sum x^2 - (\sum x)^2/n][\sum y^2 - (\sum y)^2/n]}} \]

\[ r_w = \frac{SS(xy)}{\sqrt{SS(x)SS(y)}} \]
Fecal Fat (g/24 hr) and Urinary Oxalate (mg/24 hr) secreted by a random sample of $n = 11$ persons

<table>
<thead>
<tr>
<th>Patient</th>
<th>Fecal Fat (g/24hr)</th>
<th>Urinary Oxalate (mg/24hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16</td>
<td>31</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>40</td>
</tr>
<tr>
<td>3</td>
<td>38</td>
<td>41</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>70</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>75</td>
</tr>
<tr>
<td>6</td>
<td>22</td>
<td>65</td>
</tr>
<tr>
<td>7</td>
<td>28</td>
<td>85</td>
</tr>
<tr>
<td>8</td>
<td>27</td>
<td>115</td>
</tr>
<tr>
<td>9</td>
<td>14</td>
<td>128</td>
</tr>
<tr>
<td>10</td>
<td>45</td>
<td>145</td>
</tr>
<tr>
<td>11</td>
<td>46</td>
<td>140</td>
</tr>
<tr>
<td>Sum</td>
<td>273</td>
<td>935</td>
</tr>
<tr>
<td>Mean</td>
<td>24.8</td>
<td>85.0</td>
</tr>
<tr>
<td>Person</td>
<td>Fecal Fat</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-----------</td>
<td>---</td>
</tr>
<tr>
<td>1</td>
<td>16</td>
<td>256</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>196</td>
</tr>
<tr>
<td>3</td>
<td>38</td>
<td>1,444</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>64</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>225</td>
</tr>
<tr>
<td>6</td>
<td>22</td>
<td>484</td>
</tr>
<tr>
<td>7</td>
<td>28</td>
<td>784</td>
</tr>
<tr>
<td>8</td>
<td>27</td>
<td>729</td>
</tr>
<tr>
<td>9</td>
<td>14</td>
<td>196</td>
</tr>
<tr>
<td>10</td>
<td>45</td>
<td>2,025</td>
</tr>
<tr>
<td>11</td>
<td>46</td>
<td>2,116</td>
</tr>
<tr>
<td>Sum</td>
<td>273</td>
<td>8,519</td>
</tr>
<tr>
<td>Mean</td>
<td>24.8</td>
<td>85.0</td>
</tr>
<tr>
<td>( \frac{\text{Sum}}{n} )</td>
<td>6,775.36</td>
<td>79,475.00</td>
</tr>
<tr>
<td>SS</td>
<td>1,743.64</td>
<td>16,976.00</td>
</tr>
<tr>
<td>Variance</td>
<td>174.36</td>
<td>1,697.60</td>
</tr>
<tr>
<td>SD</td>
<td>13.2</td>
<td></td>
</tr>
</tbody>
</table>
Regression Tools

\[ \bar{y} = 85.0 \]
\[ \bar{x} = 24.8 \]
\[ SS(x) = 1,743.64 \]
\[ SS(y) = 16,976.00 \]
\[ SS(xy) = 2,766.00 \]
So we can calculate

\[
Slope = b = \frac{SS(xy)}{SS(x)} = \frac{2,766.00}{1,743.64} = 1.59
\]

\[
Intercept = a = \bar{y} - b\bar{x}
\]

\[
= 85.0 + 1.59(24.8)
\]

\[
= 45.54
\]

The straight line depicting the regression relationship of \(y\) on \(x\) is

\[
\hat{y} = a + bx
\]

\[
\hat{y} = 45.54 + 1.59x
\]

At \(x = 40\), the regression estimate for \(y\) is:

\[
\hat{y} = 45.54 + 1.59(40) = 109.14
\]
With this information, we can add the regression line 
\[ \hat{y} = 45.54 + 1.59x \]
to the scatter plot, as shown below.

\[ \hat{y} = 45.54 + 1.59(x = 45) = 117.09 \]
Hypothesis test for regression of urinary oxalate on fecal fat

1. The Hypothesis: \( H_0: \beta = 0 \) vs \( H_1: \beta \neq 0 \)

2. The \( \alpha \) level: \( \alpha = 0.05 \)

3. The assumptions: Random normal samples for \( y \)-variable from populations defined by \( x \)-variable

4. The test statistic: ANOVA as specified by

<table>
<thead>
<tr>
<th>ANOVA</th>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>1</td>
<td></td>
<td>( bSS(xy) )</td>
<td>( \frac{SS(Reg)}{1} )</td>
<td>( \frac{MS(Reg)}{MS(Res)} )</td>
</tr>
<tr>
<td>Residual</td>
<td>( n-2 )</td>
<td></td>
<td>( SS(Res) )</td>
<td>( \frac{SS(Res)}{(n-2)} )</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>( n-1 )</td>
<td></td>
<td>( SS(y) )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( SS(Residual) = SS(y) - SS(Regression) \)
5. The critical region: Reject $H_0: \beta = 0$ if the value calculated for $F$ is greater than $F_{0.05}(1, 9) = 5.12$

6. The result:

\[
\begin{align*}
\text{SS(Reg)} &= b\text{SS(xy)} = 1.59(2,766.00) = 4,397.94 \\
\text{SS(Total)} &= \text{SS(y)} = 16,976.00 \\
\text{SS(Res)} &= 16,976.00 - 4,397.94 = 12,578.06
\end{align*}
\]

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>1</td>
<td>4,397.94</td>
<td>4,397.94</td>
<td>3.15</td>
</tr>
<tr>
<td>Residual</td>
<td>9</td>
<td>12,578.06</td>
<td>1,397.56</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>16,976.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7. The conclusion: Accept $H_0: \beta = 0$ since $F < 5.12$
\( R^2 = 0.2985 \)

\( \hat{y} = \text{regression estimate} = 45.54 + 1.59x = 117 \)

\[
C = A + B = y - \bar{y} \quad \text{Total deviation} \quad \Rightarrow SS(\bar{y})
\]

\[
A = \hat{y} - \bar{y} \quad \text{Explained by line} \quad \Rightarrow SS(\text{Reg})
\]

\[
B = y - \hat{y} \quad \text{Left over} \quad \Rightarrow SS(\text{Residual})
\]
Correlation Tools

\[ SS(x) = 1,743.64 \]
\[ SS(y) = 16,976.00 \]
\[ SS(xy) = 2,766.00 \]

The estimate of the correlation coefficient is:

\[ r_{xy} = \frac{SS(xy)}{\sqrt{SS(x)SS(y)}} \]

\[ r_{xy} = \frac{2,766.00}{\sqrt{1,743.64 \cdot 16,976.00}} = 0.5084 \]
The Correlation Coefficient

\( r \) measures linear association

\( r \) has values between \(-1 \leq r \leq +1\)

\( r \approx +1 \) implies strong positive linear association

\( r \approx -1 \) implies strong negative linear association

\( r \approx 0 \) implies no linear association
Correlation Hypothesis Testing

The hypothesis of interest deals with whether there is linear association between $x$ and $y$. If there is no such association, we would have $\rho = 0$. Hence, the hypotheses of interest are:

$$H_0: \rho = 0 \quad \text{vs} \quad H_1: \rho \neq 0$$

which we can test by using the test statistic:

$$t = r \left( \frac{n-2}{1-r^2} \right)^{\frac{1}{2}} \approx t_{(n-2)}$$

Note that this calculation requires only the sample estimate $r$ of the correlation coefficient $\rho$ and the sample size $n$ and that we need to use the $t$ distribution with $n - 2$ degrees of freedom.
Test of correlation hypothesis for urinary oxalate and fecal fat, \( n = 11, r = 0.5084 \)

1. The Hypothesis: \( H_0: \rho = 0 \ vs \ H_1: \rho \neq 0 \)

2. The \( \alpha \) level: \( \alpha = 0.05 \)

3. The assumptions: Random sample from bivariate normal distribution

4. The test statistic:

\[
t = r \left[ \frac{n-2}{1-r^2} \right]^{1/2} \approx t_{(n-2)}
\]
5. The critical region: \[ \text{Reject } H_0: \rho = 0 \text{ if the value calculated for } t \text{ is not between } \pm t_{0.975}(9) = 2.262 \]

6. The result: \[ r = 0.5084, \quad n = 11 \]
\[ t = 0.5084 \left[ \frac{9}{1 - (0.5084)^2} \right]^{\frac{1}{2}} \]
\[ = 0.5084 \left[ \frac{9}{1 - 0.2585} \right]^{\frac{1}{2}} \]
\[ = 1.77 \]

7. The conclusion: \[ \text{Accept } H_0: \rho = 0 \text{ since } t = 1.77 \text{ is between } \pm t_{0.975}(9) = 2.262 \]
Test of correlation hypothesis for Tono-Pen vs Goldman intraocular pressure, $n = 40$, $r = 0.6574$

1. The hypothesis: $H_0: \rho = 0$ vs $H_1: \rho \neq 0$

2. The assumptions: Random sample
   bivariate normal distribution

3. The $\alpha$-level: $\alpha = 0.05$

4. The test statistic: $t = r \sqrt{\frac{n-2}{1-r^2}}$
5. The rejection region: Reject $H_0: \rho = 0$, if $t$ is not between $\pm t_{0.975}(38) = 2.02$

6. The result:

\[ n = 40, \quad r = 0.6574, \quad r^2 = 0.44, \]
\[ t = 0.6574 \sqrt{\frac{38}{1 - 0.6574^2}} = 0.66 \sqrt{\frac{38}{0.56}} = 5.44 \]

7. The conclusion:

Reject $H_0: \rho = 0$

Since $t = 5.44$ is not between $\pm 2.02$
Health Indicators and the Organization of Health Care Systems in Western Europe

ABSTRACT

Objective. This study investigated the association between health care systems and health indicators in 17 Western European countries. Cross-national comparisons were conducted with regression analysis between 17 Western European countries with respect to health care systems, national health services, and social security systems.

Methods. Cross-national comparisons were conducted with regression analysis between 17 Western European countries with respect to health care systems, national health services, and social security systems.

Results. Health care expenditures were unevenly distributed in potential years of life, how to improve and to infant mortality rates they were positively correlated to the expectancy for females. Regression models predicted that countries with national health services systems would have lower infant mortality rates at similar levels of gross domestic product (GDP) and health care expenditures. Finally, increases in health care expenditures would decrease the risk of observed in potential years of life, mortality rates decreasing to GDPs. This decrease would be greater in countries with national health services than in those with social security systems. The model predicted this difference to be about 15% at average levels of health expenditures.


Introduction

The health of a population is influenced by many factors, including lifestyle and the health care system. Environmental and lifestyle factors are strongly dependent on each country's level of development and position within the worldwide division of labor.

The socioeconomic level of development, as measured by per capita gross domestic product (GDP), has been shown to have an impact on the health of a population. However, many national health care systems have been shown to have a negative impact on the health of their populations. In addition, there is a significant relationship between the level of social security and the health care expenditures of a country.

Relational factors influence health care systems. The structure of the labor market and the power of social security systems have an impact on the organization of the health care system. In Western European countries, social security systems are involved in a process of political and economic integration. The value of the guarantee of access to health care services and social security systems is their social security systems in Western European countries. The aim of this study is to analyze the efficiency of two health care systems in terms of their efficiency—which is the relationship between a country's health care expenditures and health indicators.

The study was conducted in Western European countries. The data were collected using a questionnaire administered to health care professionals in each country. The questionnaire included questions on the organization of health care systems, the level of social security, and the health care expenditures of each country.

At the onset of the study, the authors were concerned with the organization of health care systems and the health care services in Western European countries. Differences in the level of social security and the health care system are evident in countries with lower social security systems and higher health care expenditures. However, the relationship between health care expenditures and health indicators is not always clear.

The study was carried out by Peter England and colleagues. The results were published in the Journal of Pain & Palliative Care Pharmacotherapy, 1985; 2: 33-48.

The study was supported by the European Commission.
TABLE 3—Correlation Matrix of Selected Health and Socioeconomic Indicators of Western European Countries

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Health Care Expenditures (per Capita)</th>
<th>% Population Covered</th>
<th>Public Health Expenditures (per Capita)</th>
<th>Infant Mortality Rate</th>
<th>Maternal Mortality Rate</th>
<th>PYLL Maternal</th>
<th>PYLL Paternal</th>
<th>Life Expectancy Males</th>
<th>Life Expectancy Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP</td>
<td>0.97</td>
<td>-0.60</td>
<td>0.50</td>
<td>-0.35</td>
<td>-0.40</td>
<td>-0.09</td>
<td>-0.59</td>
<td>-0.46</td>
<td>-0.31</td>
</tr>
<tr>
<td>Health care expenditures (per Capita)</td>
<td>-0.12</td>
<td>-0.14</td>
<td>0.50</td>
<td>0.35</td>
<td>0.22</td>
<td>0.16</td>
<td>-0.07</td>
<td>0.25</td>
<td>0.23</td>
</tr>
<tr>
<td>% population covered by health care</td>
<td>-0.20</td>
<td>0.31</td>
<td>0.80</td>
<td>-0.05</td>
<td>-0.15</td>
<td>0.07</td>
<td>-0.07</td>
<td>0.18</td>
<td>-0.07</td>
</tr>
<tr>
<td>Public health expenditure (% of GPD)</td>
<td>-0.10</td>
<td>0.31</td>
<td>0.80</td>
<td>-0.05</td>
<td>-0.15</td>
<td>0.07</td>
<td>-0.07</td>
<td>0.18</td>
<td>-0.07</td>
</tr>
<tr>
<td>Infant mortality rate</td>
<td>0.80</td>
<td>0.31</td>
<td>0.80</td>
<td>-0.05</td>
<td>-0.15</td>
<td>0.07</td>
<td>-0.07</td>
<td>0.18</td>
<td>-0.07</td>
</tr>
</tbody>
</table>

Note: P-values = 0.01. Maternal mortality rate = 0.05.

**Results**

Mean values of all variables for the year 1970 for each country system are shown in Table 1. When we computed the matrix of the socioeconomic indicators between the two variables, only the two variables GDP and health care expenditures correlated significantly.

Pearson's correlation coefficients were statistically significant for health care expenditures and GDP and for both variables and infant mortality (Table 2).

In simple regression models, health care expenditures and GDP were associated with life expectancy for males and females associated with polio deaths and tobacco deaths. However, these were not associated with the same health indicators for males. Health care expenditures and tobacco deaths were correlated 12% of potential men of life lost to females (F = 0.05) and 37% of life expectancy for females (F = 0.25).

GDP explained 39% of the variability in potential life years lost to females (F = 0.34, P < 0.01) and 22% of life expectancy for females (F = 0.34, P < 0.01). When other socioeconomic indicators explained significantly, they explained variability in these two health indicators.

In the overall health care expenditures proved to be the most explanatory variable of infant mortality (W = 0.1). Type of health care system was included in models III and IV (Table 3) according to. Infant mortality rates were lower for national health services compared with those in other systems in a similar level of GDP model. Model III and health care expenditures (model IV) the magnitude of the coefficient of R² would be of approximately 23% (model III) or 31% (model IV).

Table 4 shows the best model that estimates the ratio of observed to predicted infant mortality rates for countries with similar health care expenditures. Model III, the infant mortality rate would be about 15% lower in national health services. Figure 1 illustrates the regression effect on this model in the ratio for the countries included in the study, as health care expenditures increase, the decreasing variability of infant mortality.

**TABLE 4—Regression Models for Infant Mortality Rate as a Function of Health Care Expenditures and Type of Health Care System, Western Europe**

<table>
<thead>
<tr>
<th>Model</th>
<th>Model II</th>
<th>Model III</th>
<th>Model IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gross domestic product</td>
<td>-0.20 (0.12)</td>
<td>-0.40 (0.11)</td>
<td>-0.20 (0.11)</td>
</tr>
<tr>
<td>Health care expenditure</td>
<td>-0.20 (0.079)</td>
<td>0.31 (0.01)</td>
<td>0.11 (0.005)</td>
</tr>
<tr>
<td>Social security</td>
<td>0.10 (0.01)</td>
<td>0.10 (0.01)</td>
<td>0.01 (0.01)</td>
</tr>
<tr>
<td>Income per capita</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Note: As variables were negatively correlated.

20 - 25
Test of correlation hypothesis for infant mortality rate and gross domestic product, $n = 17$, $r = -0.64$

**Example**

1. The hypothesis: $H_0: \rho = 0$ vs $H_1: \rho \neq 0$

2. The assumptions: Random sample, bivariate normal distribution

3. The $\alpha$-level: $\alpha = 0.05$

4. The test statistic: $t = r \sqrt{\frac{n-2}{1-r^2}}$
5. The rejection region: Reject $H_0$: $\rho = 0$, if $t$ is not between $\pm t_{0.975}(15) = 2.13$

6. The result: $n = 17$, $r = -0.64$,

$$t = -0.64 \sqrt{\frac{15}{1 - (-0.64)^2}} = -0.64 \sqrt{\frac{15}{0.59}} = -3.23$$

7. The conclusion: Reject $H_0$: $\rho = 0$
Since $t = -3.23$ is not between $\pm 2.1315$
Test of correlation hypothesis for life expectancy for males and females, \( n = 17, r = 0.67 \)

1. **The hypothesis:**  
   \[ H_0: \rho = 0 \quad \text{vs} \quad H_1: \rho \neq 0 \]

2. **The assumptions:**  
   Random sample  
   Bivariate normal distribution

3. **The \( \alpha \)-level:**  
   \( \alpha = 0.05 \)

4. **The test statistic:**  
   \[ t = r \sqrt{\frac{n-2}{1-r^2}} \]
5. The rejection region: Reject $H_0: \rho = 0$, if $t$ is not between $\pm t_{0.975}(15) = 2.1315$

6. The result:

\[ n = 17, r = 0.67, \]
\[ t = 0.67 \sqrt{\frac{15}{1 - 0.67^2}} = 0.67 \sqrt{\frac{15}{1 - 0.45}} \]
\[ t = 0.67 \sqrt{\frac{15}{0.55}} = 0.67 \sqrt{27.27} \]
\[ t = 0.67(5.22) \quad t = 3.49 \]

7. The conclusion:
Reject $H_0: \rho = 0$
Since $t = 3.49$ is not between $\pm 2.1315$
Social Capital, Income Inequality, and Mortality

ABSTRACT

Objective: Recent studies have demonstrated that income inequality is related to mortality rates. It was hypothesized in this study, that income inequality is related to reduction in social cohesion and that improvement in social capital is in turn associated with increased mortality.

Methods: In this cross-sectional ecological study, based on data from 30 states, social capital was measured by weighted responses to two items from the General Social Survey: proportion of respondents in each state who agree to the statement that people can be trusted. Age-standardized total and cause-specific mortality rates in 1990 were obtained for each state.

Results: Inequality was strongly correlated with both pre-exposure group membership (r = -.49) and lack of social trust (r = .76). In turn, both social trust and group membership were associated with lower mortality, as well as deaths from coronary heart disease, malignant neoplasms, and infant mortality.

Conclusions: These data support the notion that income inequality leads to increased mortality via disinvestment in social capital. Am J Public Health. 1997;87:1591-1596.

Introduction

A number of cross-sectional studies have indicated that the degree of income inequality in a given society is strongly related to the society's level of mortality. In one investigation of nine nations included in the Luxembourg Income Study, a correlation of .60 was reported between inequality in income distribution and the proportion of income received by the top 10% of the population at the lowest income levels. Two recent US studies independently demonstrated an association between income inequality and mortality. 1-2

Kawachi and Kennedy observed that using income inequality results in increased levels of isolation, which may have deleterious behavioral and health consequences. In another study, income inequality was associated with increased mortality, after adjustment for income in all-cause mortality in 1990 was 34.4 per 100,000 persons, confidence interval 32.6-36.2. The Robson Index was also associated with deaths from specific causes, including coronary heart disease, cancer, and infant mortality.

In an independent study, Kaplan et al. 3 examined the association between income inequality—a measured by the share of aggregate income earned by the bottom 50% of households—and state-level variations in total mortality. A strong association was found between their measure of income inequality and age-adjusted death rates in 1990 (r = .35, p < .001). Moreover, the degree of income inequality in each state in 1980 was a powerful predictor of levels of total mortality 10 years later.

The present study attempts to identify the mechanisms underlying the association between income inequality and mortality levels remain is the established. One hypothesis is that rising income inequality results in increased levels of isolation, which may have deleterious behavioral and health consequences. In another study, income inequality was associated with increased mortality, after adjustment for income in all-cause mortality in 1990 was 34.4 per 100,000 persons, confidence interval 32.6-36.2. The Robson Index was also associated with deaths from specific causes, including coronary heart disease, cancer, and infant mortality.

In an independent study, Kaplan et al. 3 examined the association between income inequality—a measured by the share of aggregate income earned by the bottom 50% of households—and state-level variations in total mortality. A strong association was found between their measure of income inequality and age-adjusted death rates in 1990 (r = .35, p < .001). Moreover, the degree of income inequality in each state in 1980 was a powerful predictor of levels of total mortality 10 years later.

The present study attempts to identify the mechanisms underlying the association between income inequality and mortality levels remain is the established. One hypothesis is that rising income inequality results in increased levels of isolation, which may have deleterious behavioral and health consequences. In another study, income inequality was associated with increased mortality, after adjustment for income in all-cause mortality in 1990 was 34.4 per 100,000 persons, confidence interval 32.6-36.2. The Robson Index was also associated with deaths from specific causes, including coronary heart disease, cancer, and infant mortality.
The relationship between income inequality, as measured by the Gini index, and lack of social trust.

**Figure 1** - The relationship between income inequality, as measured by the Gini index, and lack of social trust.

**Figure 2** - The relationship between age-adjusted mortality rates and lack of social trust.

**Table 1** - Correlations among indicators of social capital, poverty, income inequality, and mortality

<table>
<thead>
<tr>
<th>Poverty</th>
<th>Mortality</th>
<th>Group Membership</th>
<th>Social Trust</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immunity</td>
<td>0.70</td>
<td>0.65</td>
<td>0.53</td>
</tr>
<tr>
<td>Health</td>
<td>0.74</td>
<td>0.65</td>
<td>0.53</td>
</tr>
<tr>
<td>Membership</td>
<td>0.65</td>
<td>0.74</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Data Analysis

Statistical models were used to examine the relationships of social capital indicators to mortality rates. Two sets of models were examined in each outcome of interest. In the first set of models, we regressed the weighted social capital measures (i.e., weighted average group membership and weighted average social trust) against all cause and cause-specific mortality rates. In the second set of models, we included the regression models for each outcome in prevalence of poverty. To examine the effects of social trust, we measured the relationship between social trust, lack of social trust, and social capital. The results of this analysis were used to draw conclusions about the impact of social capital on mortality rates. We conducted a path analysis based on a causal model in which social capital affects mortality through its impact on social trust.

Results

**Relationships among Social Capital, Measures of Income, and Mortality**

The four indicators of social capital—number of participants in civic associations and the weighted proportions of respondents who agree that “most people would not take advantage of you if they could” “people usually look out for themselves” “people don’t care what happens to other people”—were highly correlated with each other (Table 1). Since these variables may not represent an exhaustive list of “social capital” indices,
For Table 1, Correlation between Mortality and Social Mistrust, $n = 39$, $r = 0.79$

1. The hypothesis: $H_0: \rho = 0$ vs $H_1: \rho \neq 0$

2. The assumptions: Random sample, bivariate normal distribution

3. The $\alpha$ level: $\alpha = 0.05$

4. The test statistic: $t = r\sqrt{\frac{n-2}{1-r^2}}$
5. The rejection region:  
Reject $H_0: \rho = 0$, if $t$ is not between $\pm t_{0.975}(37) = 2.02$

6. The result:  
$n = 39, r = 0.79, r^2 = 0.6241,$  
$t = 0.79 \sqrt{\frac{37}{1-0.79^2}} = 0.79 \sqrt{\frac{37}{0.3759}} = 7.8$

7. The conclusion:  
Reject $H_0: \rho = 0$  
Since $t = 7.8$ is not between $\pm 2.02$
The Evolving Epidemiology of Chlamydial and Gonococcal Infections in Response to Control Programs in Winnipeg, Canada

ABSTRACT

Objectives. The purpose of this study was to describe and compare the transmission dynamics of chlamydial and gonorrheal infections in Winnipeg, Manitoba, Canada, and to assess implications for control programs.

Methods. Chlamydial and gonorrheal surveillance case reports (1988 through 1993) and contact-tracing reports (1993 through 1995) were examined.

Results. High incidence rates of both chlamydial and gonorrheal infections occurred in geographic core areas characterized by low socioeconomic status. A decline in the number of reported cases of chlamydial (51%) and gonorrheal (64%) infections occurred between 1992 and 1993. For chlamydial, the decline was most pronounced in non-core areas and, while the proportion was similar in core and non-core areas.

Conclusions. Chlamydial and gonorrheal infections are evolving through different epidemiologic, clinical, and transmission patterns, in response to newly introduced control programs. Emerging core group behavior and changing transmission patterns are likely to be the basis of any sustained control efforts.

Lysén E. Bouchard, MB, PhD, Stephen Meehan, MD, MPH
Christine Garncarski, MD, FRCP, G. Peace OZ, MB, FRCPC
Gary W. Macleod, MD, FRCPC, and Robert C. Krampl, MD, FRCP

Bacterial sexually transmitted diseases (STDs) remain an important public health problem worldwide. Most prevalent STDs are carried by relatively small numbers of antimicrobial-resistant strains and therefore become the focus of major disease control efforts. The approach to control of STDs has been guided by our understanding of the transmission and behavior dynamics of STDs. New epidemiological and behavioral findings have provided valuable insights into the transmission dynamics of STDs. Several findings have related the proportion of core groups among a selected group of individuals to the spread of STDs within populations. The mathematical modeling of the core group concept has provided insights into the behavior of these groups.
Methods

Descriptive Epidemiology

Manitoba is a province in Canada with a population of approximately 1.2 million. Winnipeg, with a population of close to 600,000, is the only major city in the province. All cases of chlamydia and gonorrhea are reportable to the provincial health department by both physicians and laboratories under the Public Health Act. Since laboratories in the province routinely report all positive diagnostic tests for chlamydia and gonorrhea, few laboratory-confirmed cases were missed. In Winnipeg, a specially trained public health nurse performed interview and conducted contact tracing for most persons infected with chlamydia or gonorrhea.11 Sexual contacts of those with chlamydia and gonorrhea are also available to the provincial health department. As a measure of comparing the incidence of chlamydia and gonorrhea by geographic location, the province was divided into 222 geographic areas defined by provincial administrative boundaries.28 Rural municipalities for non-Winnipeg residents and the 24 postal areas for Winnipeg residents. The average annual incidence of chlamydia and gonorrhea per 100,000 population was computed on the basis of all reported cases. Incidence rates were directly standardized for age and gender to the total 1991 Manitoba population. Pearson moment correlation coefficients were used to assess the geographic relationship between chlamydia and gonorrhea rates. Two-sample t-tests were used to assess the statistical significance of differences in proportions. SAS (version 6.12, SAS Institute Inc., Cary, NC) was used in performing all statistical analyses.

Winnipeg study areas

Winnipeg was divided into 3 geographic risk areas ("core," "adjacent," and "peripheral") separately for chlamydia and gonorrhea. This was done empirically by aggregating postal areas based on their average annual incidence rates between 1991 and 1995 and visually examining the break points from the distribution curve (see Figure 1c). The 1991 Census of Canada was used to estimate unemployment rates, mean household incomes, language usually spoken at home, mobility status, and population density for the Winnipeg postal areas (Statistics Canada, Methodology Release 1991). The incidence of other selected communicable diseases was derived from the provincial notifiable disease registry.

Results

STD Incidence Rates by Geographic Site

Figure 1 showed that the incidence of chlamydia and gonorrhea was highest in the Winnipeg core. The provinces of Manitoba were divided into 222 geographic areas defined by provincial administrative boundaries. The 1991 Census of Canada was used to estimate unemployment rates, mean household income, language usually spoken at home, mobility status, and population density for the Winnipeg postal areas (Statistics Canada, Methodology Release 1991). The incidence of other selected communicable diseases was derived from the provincial notifiable disease registry.
Hypothesis test for correlation between gonorrhea rate and chlamydia rate

1. The hypothesis: \( H_0: \rho = 0 \) vs \( H_1: \rho \neq 0 \)

2. The assumptions:
   - random sample
   - bivariate normal distribution

3. The \( \alpha \)-level: \( \alpha = 0.05 \)

4. The test statistic:
   \[
   t = r \sqrt{\frac{n-2}{1-r^2}}
   \]