MEASURING THE RISK OF INFECTION AT WORK

BY

DAVID HEWITT and ALICE STEWART

From the Institute of Social Medicine, Oxford

The present investigation was prompted by the Northamptonshire Miniature Mass Radiography Survey of 1945-46. One of the findings of this survey, as presented by Smith (1947), was a significantly higher proportion of newly-discovered cases of active pulmonary tuberculosis* among boot and shoe factory workers than among workers in other trades. In July, 1947, the manufacturers asked for an investigation to be carried out, with the object of discovering whether this higher incidence of tuberculosis was the result of working conditions, or whether it was due to factors outside the control of the industry. Since then five related papers have been published. Stewart and Hughes (1949, 1951) describe how an intensive scrutiny of the mass radiography records broke down the total of cases found in boot and shoe factories into a number of homogeneous groups, and, notably, how it was possible to discern an association between cases of chronic tuberculosis and newly-infected cases. The theory that selective recruitment is solely responsible for the high tuberculosis morbidity rate is criticized, and attention drawn to the much higher rates prevalent in large than in small factories. An experiment carried out in Northampton and Leicester by Hirch (1951) shows the level of bacterial contamination of the air in boot and shoe factory workshops to be largely determined by the number of persons present, however close together or far apart they may be. Cairns and Stewart (1951) present a comparative history of the printing and shoe-making trades, to show that, while the living standards enjoyed by shoemakers are not such as would be expected to cause an excess of tuberculosis, there is some plausibility in the theory of a long-wave epidemic in the factory section of the shoe-making community. Webb, Stewart, and Sutherland (1951) discuss the tendency, revealed in a street survey of Northampton, for cases of tuberculosis to occur in adjacent houses. These papers all tend to strengthen the view that infection outside the home is an important factor in the tuberculosis morbidity of boot and shoe workers.

Since the original mass radiography survey, fresh material has been gathered which forms the basis of this paper. A further attempt is now made to trace the path of infection in boot and shoe factories, to suggest appropriate measures for

* Throughout this paper "tuberculosis" is to be understood as "pulmonary tuberculosis".
reducing the disease incidence, and to draw some conclusions which may be of general epidemiological interest.

(1) MATERIAL USED IN THE PRESENT STUDY

The new material comprises:

(a) Records of all boot and shoe factory operatives notified as cases of tuberculosis in the Borough and County of Northampton during the decade 1940-49. There were nearly 600 such records, each including the occupation and last place of employment before notification. The placing of cases in individual workshops (see Table I, opposite) is based on this information.

(b) Board of Trade returns from 194 boot and shoe factories in the same area, showing the number of workers employed on production in each of the years 1941-1949. These include all the factories in continuous production during the period of study.

(c) The results of a survey of the same factories made in 1948 to determine the number of men and women working in each room or compartment, distinguishing the six main occupational groups.

(d) For a smaller number of factories*, approximately two-thirds of the whole, the floor area of each room or compartment was also obtained. From these measurements the floor space per worker was calculated (see Table II, p. 212).

Compared with the mass radiography survey, the present body of data is more comprehensive but less clear cut. It is, for instance, less certain that the last place of work before notification was the place at which the worker first began to develop tuberculosis. With the mass radiography data there was no difficulty in establishing comparable morbidity rates for different sections of the population examined. In the present instance, so-called notification rates have been calculated from the total number of cases and the number of persons at work in 1948, but these rates are only comparable one with another if the relationship between the 1948 working population and the total exposure is the same for each group examined. Fortunately, such a relationship may be assumed, since the processes involved and the machinery used prescribe very nearly standard working conditions for all the factories. The history of most of the factories was also very similar throughout the decade.† While, therefore, the lack of precision in the data may have blurred the outlines of the main pattern, there is no reason to suppose that it has caused distortion.

(2) UNITS OF STUDY

Three basic groupings of workers have been studied:

(a) factories; (b) walled rooms; (c) compartments‡.

* This sample had the same age, sex and occupational composition as the whole of the material and included factories of all sizes. There was, however, a tendency for the measurements to be obtained more frequently from the larger factories.

† Of the 194 factories covered by the Board of Trade returns, 17 were "concentrated" during the war. These have been excluded from the calculations. A further 20 subsidiary factories which were included in the mass radiography survey but which did not carry out the complete manufacturing processes were also excluded. The remaining 177 factories accounted for 515 persons notified during the decade 1940-49.

‡ A compartment is co-extensive with a walled room where there is no subdivision of the room. Wherever a permanent partition was found which reached eye level but did not touch the ceiling, it was held that the workers on either side belonged to separate compartments.
MEASURING THE RISK OF INFECTION AT WORK

The obvious advantage of the factories as units of study was that their notification rates, being based on relatively large numbers, were much less erratic than the corresponding rates for the other two groupings. There are, however, in factory populations, significant variations associated with age, sex, and occupation, which influence the liability to develop tuberculosis. Since all the factories included in the survey had approximately the same proportion of workers of each sex and occupation, these variations would inevitably have been overlooked, had factories been made the basis of study. Compared with the walled rooms, the compartments more often contained members of only one sex or only one occupational group. It was decided, therefore, that from the point of view of homogeneity they were better than either of the other groupings.

(3) Effect of Partitions

Since partitions have some of the characteristics of walls, it is of interest to compare the notification rates in partitioned and unpartitioned rooms. This point is examined in Table I, where, for each of three size groups, the notification rates for the two types of room are compared. In each group the difference between the frequency of tuberculosis in rooms with and without partitions is statistically significant, the lower rates occurring always in the partitioned rooms. This finding suggests that the presence of partitions has had an appreciable effect on the risk of cross-infection at work and is a further reason for preferring compartments as the basic units of study.

### TABLE I

**Effect of Partitions on Frequency of Tuberculosis**

<table>
<thead>
<tr>
<th>Type of Room</th>
<th>No. of Workers</th>
<th>Partitions</th>
<th>No. of such Rooms</th>
<th>No. of Workers, 1948</th>
<th>Cases of Pulmonary Tuberculosis, 1940-49</th>
<th>Notification Rate, 1940-49</th>
<th>Average No. of Workers per Room, 1948</th>
<th>Rate in Partitioned Rooms lower by (per cent.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 to 75</td>
<td>1 to 75</td>
<td>None</td>
<td>502</td>
<td>10,590</td>
<td>183</td>
<td>1.73</td>
<td>21.1</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>One or more</td>
<td>54</td>
<td>1,854</td>
<td>22</td>
<td>1.19</td>
<td>34.3</td>
<td></td>
</tr>
<tr>
<td>76 to 150</td>
<td>76 to 150</td>
<td>None</td>
<td>35</td>
<td>3,681</td>
<td>101</td>
<td>2.74</td>
<td>105.2</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>One or more</td>
<td>26</td>
<td>2,722</td>
<td>46</td>
<td>1.69</td>
<td>104.7</td>
<td></td>
</tr>
<tr>
<td>151 or more</td>
<td>151 or more</td>
<td>None</td>
<td>7</td>
<td>1,555</td>
<td>48</td>
<td>3.09</td>
<td>222.1</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>One or more</td>
<td>18</td>
<td>4,026</td>
<td>115</td>
<td>2.86</td>
<td>223.7</td>
<td></td>
</tr>
<tr>
<td>All Rooms</td>
<td></td>
<td></td>
<td>642</td>
<td>24,428</td>
<td>515</td>
<td>2.11</td>
<td>38.1</td>
<td></td>
</tr>
</tbody>
</table>

The rooms were divided into three groups according to the number of workers they contained. For each group a $\chi^2$ test of the hypothesis that risk of tuberculosis is not affected by partitioning was performed. The combined results of these tests give a value of $\chi^2$ equal to 10.54 with three degrees of freedom, the corresponding value for $P$ being less than 0.02. In the first size group it so happened that the rooms with partitions were
of greater average size than those without. Hence the 31 per cent. reduction in the notification rate is likely to be an underestimate of the effect of partitions. On the other hand, the reduction of 7 per cent. in rooms with over 150 workers could not by itself be held to have proved a real difference between the two types of room.

(4) Effect of Spacing

It is a familiar notion that overcrowding, both by multiplying opportunities for infection and by lowering hygienic standards, increases the risk of contracting infectious diseases. Tuberculosis in particular, since it is thought to require close and sustained contact for its passage from one person to another, is spoken of as a disease of overcrowding. We expected, therefore, to find the highest notification rates in the most congested working units, but this expectation was not fulfilled (see Table II).

**TABLE II**

**APPARENT ASSOCIATION BETWEEN SPACING OF WORKERS AND FREQUENCY OF TUBERCULOSIS**

<table>
<thead>
<tr>
<th>Range of Area per Worker (sq. ft)</th>
<th>No. of Workers</th>
<th>No. of Notifications</th>
<th>Notification Rate (%)</th>
<th>Average No. of Workers per Compartment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-55</td>
<td>6,279</td>
<td>140</td>
<td>2.23</td>
<td>54.6</td>
</tr>
<tr>
<td>56-85</td>
<td>4,344</td>
<td>102</td>
<td>2.35</td>
<td>33.7</td>
</tr>
<tr>
<td>86-120</td>
<td>4,681</td>
<td>128</td>
<td>2.73</td>
<td>36.9</td>
</tr>
<tr>
<td>121-175</td>
<td>3,126</td>
<td>76</td>
<td>2.43</td>
<td>25.4</td>
</tr>
<tr>
<td>Over 175</td>
<td>1,460</td>
<td>15</td>
<td>1.03</td>
<td>11.9</td>
</tr>
</tbody>
</table>

When the compartments were first classified by area-per-worker it seemed that the maximum risk of tuberculosis occurred in the middle of the range; i.e. units with more than 120 sq. ft per worker had notification rates below the maximum, *but so did those with less than 55 sq. ft.* This obvious paradox was resolved when it was realized that there was considerably more variation of area-per-worker from one occupational group to another than within each occupational group. The notification rate for units with less than 70 sq. ft per worker was in effect a rate for "closers"; in units with more than 150 sq. ft per worker there was a preponderance of "rough stuff" and "shoe room" workers, while the other occupational groups also fell within fairly well-defined ranges of area-per-worker. Since the occupational groups differ widely in sex composition, they also differ in their liability to develop tuberculosis. It was decided, therefore, that the association between crowding and risk of infection should be studied separately in each occupational group (see Table III).

All units containing workers from more than one occupation were excluded, and the remainder divided into six occupational groups. Each group was in turn subdivided into as many area-per-worker groups as the material allowed. For each of the six occupational groups a $\chi^2$ test was then performed for the hypothesis that, within each occupation, the distribution of cases of tuberculosis was independent of area-per-worker. The results of the tests were then pooled to obtain a value of $\chi^2$ equal to 23.48 with 17 degrees of freedom, for which $P$ is greater than 0.10. For the sake of contrast six more tests were
MEASURING THE RISK OF INFECTION AT WORK

TABLE III

LACK OF GENUINE ASSOCIATION BETWEEN SPACING OF WORKERS AND FREQUENCY OF TUBERCULOSIS

<table>
<thead>
<tr>
<th>Occupational Groups</th>
<th>Degrees of Freedom in $\chi^2$ Test</th>
<th>Result of Testing whether Tuberculosis is Independent of</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(a) Area per Worker</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\chi^2$</td>
</tr>
<tr>
<td>Clickers</td>
<td>5</td>
<td>5·28</td>
</tr>
<tr>
<td>Closers</td>
<td>4</td>
<td>3·70</td>
</tr>
<tr>
<td>Rough Stuff</td>
<td>1</td>
<td>9·17</td>
</tr>
<tr>
<td>Makers</td>
<td>3</td>
<td>2·78</td>
</tr>
<tr>
<td>Finishers</td>
<td>2</td>
<td>1·10</td>
</tr>
<tr>
<td>Shoe Room</td>
<td>2</td>
<td>1·45</td>
</tr>
<tr>
<td>All Occupations</td>
<td>17</td>
<td>23·48</td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Occupations</td>
<td>16</td>
<td>14·21</td>
</tr>
</tbody>
</table>

made for the hypothesis that, within the same restricted material, the distribution of cases was independent of the number of workers per unit. In this case the sum of $\chi^2$ equals 47·25, for which $P$ is less than 0·001.

These tests still show that there is no significant evidence of higher morbidity rates in the more congested working units. Since, in general, the compartments with most workers are also those in which area-per-worker is least (Table II, column 4), it is all the more remarkable that no significant area effect appears. It has to be concluded that spacing does not affect the risk of infection within the range of variation observed. Area-per-worker has therefore not been taken into consideration in the following analysis.

Some explanation should, however, be attempted for the special case of the "rough stuff" workers. For this one occupational group the $\chi^2$ tests suggest that risk of infection may depend on area-per-worker, rather than on number of workers per unit. A possible explanation is that there may be some critical distance between workers, beyond which the benefit of wider spacing begins to be felt, and that only in the rough stuff rooms is this critical distance exceeded. On this view a crowding effect fails to appear, not because area-per-worker has no bearing on risk of infection, but because boot and shoe factories as a whole are "overcrowded"*, in the sense that the distance between workers is nearly always less than some critical amount. The possibility that such a critical distance may exist will be discussed again in Section 5. Whether it exists or not, there are three possible advantages enjoyed by rough stuff workers which are worth mentioning:

(a) Available evidence suggests that known tuberculous subjects tend to avoid rough stuff work. If this is so, sources of infection will be much fewer than in the industry as a whole.

* This use of the term overcrowding should not be confused with "statutory overcrowding". In fact, the air-space per worker in most of the workshops far exceeded the minimum prescribed by law.
DAVID HEWITT AND ALICE STEWART

(b) The bulky machines and tall racks found in rough stuff work may act like the partitions already discussed.

(c) Bacterial counts per cu. ft of air recorded by Hirch (1951) were lower in the rough stuff departments than elsewhere.

It should also be mentioned that although rough stuff is the smallest of the six occupational groups examined, and too small for further subdivision, it is not homogeneous either in sex or in operation performed. The difference in risk which appears to be due to difference in spacing might disappear if it were possible to give separate treatment to the individual occupations grouped together as "rough stuff".

(5) FIRST MEASUREMENT OF RISK GRADIENT

Having decided to discard the variable "area-per-worker", the next step was to attempt to measure the rate at which risk of infection increased with compartment size. The notification rates were therefore correlated with workers-per-compartment to obtain estimating equations of the form

\[ y = a + bW, \]

where \( y \) is the estimated notification rate, and \( W \) the number of workers in the unit.

Since the units were of widely different sizes, their notification rates were subject to different sampling errors. To allow for this a weighting system had to be introduced. The exact sampling distribution of the notification rates was not investigated, but it was assumed that they could be treated as binomials. The correct weight for each rate would then be

\[ \frac{n}{p (1-p)}, \]

where \( n \) is the number of workers on which the sample rate is based, and \( p \) the corresponding "true" proportion of workers notified. Values of \( p \) would normally have been estimated from the sample rates themselves, but the presence of many sample proportions of zero—implying infinite weight—made this impossible. The method adopted was to proceed in two stages. In the first stage all sample rates were given equal weight, and, from the equation thus calculated, weights were obtained for use in the second stage.

The second stage estimate of regression based on 807 compartment sizes and rates was

\[ y = 1.40 + 0.0096W. \]

According to this regression equation, the expected notification rate varies from 1.41 in the smallest unit encountered (one worker), to 4.40 in the largest (312 workers). That is to say, the extra hazards of factory employment can more than treble the risk of tuberculosis in the extreme case.

Is this the best estimating equation that can be obtained? In Table IV are set out the actual distribution of notifications between units of various sizes, the distribution expected if risk of tuberculosis were independent of compartment size (Model "A"), and that expected on the basis of the equation above (Model "B").

It is clear that Model "B", although a considerable improvement on Model "A", is still far from being a satisfactory explanation of the way in which cases of tuberculosis were distributed. Compared with the expectation based on Model
**MEASURING THE RISK OF INFECTION AT WORK**

**TABLE IV**

**TEST OF MODELS "A", "B", AND "C"**

Model "A": Risk independent of number of workers per compartment.
Model "B": Linear increase of risk with number of workers per compartment.
Model "C": Curvilinear increase of risk with number of workers per compartment (see Figure).

<table>
<thead>
<tr>
<th>Compartment Size Group</th>
<th>Number of Workers in such Compartments, 1948</th>
<th>Number of Cases of Tuberculosis, 1940-49</th>
<th>Number of Cases Expected on Basis of Model &quot;A&quot;, &quot;B&quot;, or &quot;C&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>&quot;A&quot;</td>
</tr>
<tr>
<td>1 to 25</td>
<td>5,698</td>
<td>72</td>
<td>120.13</td>
</tr>
<tr>
<td>26 to 50</td>
<td>6,124</td>
<td>115</td>
<td>129.11</td>
</tr>
<tr>
<td>51 to 75</td>
<td>4,235</td>
<td>102</td>
<td>89.28</td>
</tr>
<tr>
<td>76 to 100</td>
<td>2,837</td>
<td>71</td>
<td>59.81</td>
</tr>
<tr>
<td>100 to 150</td>
<td>2,974</td>
<td>80</td>
<td>62.7</td>
</tr>
<tr>
<td>151 to 250</td>
<td>1,692</td>
<td>50</td>
<td>35.67</td>
</tr>
<tr>
<td>Over 250</td>
<td>868</td>
<td>25</td>
<td>18.3</td>
</tr>
<tr>
<td>Total</td>
<td>24,428</td>
<td>515</td>
<td>515.00</td>
</tr>
</tbody>
</table>

Degrees of Freedom left for $\chi^2$ Test ........................................ 6

Probability that Discrepancy between Expected and Actual Numbers of Cases could have arisen by Chance ..................................................... $P<0.001$

<table>
<thead>
<tr>
<th>Probability that Discrepancy between Expected and Actual Numbers of Cases could have arisen by Chance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0.05&lt;P&gt;0.02$</td>
</tr>
<tr>
<td>$0.45&gt;P&gt;0.35$</td>
</tr>
</tbody>
</table>

"B", there is a deficiency of cases at each end of the range of compartment sizes and an excess in the middle. From a graph showing notification rates in these unit-size groups it appeared that a better fit to the data would be obtained by a curved line, possibly one tending to an upper limit. A systematic departure from linearity would justify a more complicated equation relating risk to numbers, but, before complicating the hypothesis, it is prudent to consider how else the appearance of curvilinearity could have arisen. It might occur, for instance, if either (a) the susceptible occupational groups were over-represented in the medium-sized compartments, or (b) the larger compartments contained groups of workers too remote to transmit infection to one another, causing the effective number of persons to fall below the total number present.

As far as (a) is concerned, it is impossible to disprove any such over-representation without becoming involved in a circular argument, since the "susceptibility" of an occupational group is inferred from its notification rate and this, in turn, is known to be largely affected by the size of the working units of which that occupational group is composed. But at least there is nothing amounting to evidence that occupational differences have caused an appearance of curvilinearity.

Possibility (b)—that large compartments may contain numbers of workers out of "infecting range" of each other—is that mentioned above in connexion with rough stuff workers. It has been rejected for a number of reasons. First, there is the evidence of an experimental demonstration by Williams (1950) of the rapidity with which exhaled droplets can reach all corners of a working unit. Secondly, droplets are not the only medium of infection to be considered; the notorious
hardihood of the tubercle bacillus enables it to survive in dust for long periods, and infected dust can persist in a room despite repeated and thorough sweepings (Lidwell, 1950; Lidwell and Lowbury, 1950). Thirdly, since operatives are not rooted in their place of work, "infecting range" may extend beyond the actual compartment.

(6) SECOND MEASUREMENT OF RISK GRADIENT

It was therefore considered proper to try a non-linear form of estimating equation. There are precedents (deriving from Farr's "Law of Density") for describing cognate material in terms of an equation of the form

\[ y = a x^n, \]

where \( y \) is a mortality or morbidity rate, \( x \) an independent variable, and \( n \) a positive constant smaller than unity (Brownlee, 1920; Cheeseman and others, 1939). There is no doubt that to use a curve of this type here would result in a better fit to the data, as would also the addition of a term \(-cW^2\) to the equation already used. But both these types of curves have the disadvantage that they do not tend to an upper limit, and that they cannot be interpreted in terms of any known mechanism of infection. Instead of using them, therefore, an exponential form has been adopted, of which the rationale is as follows:*

Suppose that in the course of daily life a person is in touch with \( F \) "effective contacts"—effective in the sense that, if any one or more of them is a carrier of a certain disease, then he or she is in contact with the disease. (The number \( F \) will not, of course, be the same for, say, airborne as for contagious diseases.) The total of "effective contacts" may be thought of in two parts, of which \( U \) represents the contacts made at work, and \( D \) all other contacts. For large groups of workers it is reasonable to treat \( D \) as a constant, and \( U \) as a variable, whose value depends upon the specification of the particular group. In the present example, \( U \) will be set to equal the number of persons assumed to be contacted at work, i.e. one less than the total number of workers in the same compartment. Now let \( \pi \) be the proportion of carriers in a community and \( R \) the risk of making contact with the infection. It follows that

\[ R = 1 - (1-\pi)^D + U \]

We may also suppose that \( R = y/M \), where \( y \), as before, is the expected notification rate and \( M \) the percentage of the community which, once in "effective contact" with tuberculosis, will develop the disease within a specified period. As the number of "effective contacts" increases, the risk \( R \) approaches unity, and \( y \) approaches its upper limit \( M \). Substituting \( y/M \) for \( R \) in Equation (i) and expressing it in the more convenient logarithmic form we have

\[ \log (1 - y/M) = D \log (1 - \pi) + U \log (1 - \pi) \]

The percentage of any group observed to develop tuberculosis obviously depends in part on the length of time for which the group is under observation. By expressing each notification rate as a fraction of the ten-year maximum rate, we can make an

* Since writing this paper we have found the same argument applied by Hogben (1947) to the risks of jaundice following transfusion with pooled plasma or serum.
adequate allowance for this time factor. But there is a more subtle way in which the passage of time complicates the assessment of risk. A person continuously in contact with $F$ others will not always be in contact with the same $F$ individuals. Therefore in a long period he will have a total number of contacts which exceeds $F$. No correction has been attempted for this second factor and, as a consequence, the estimates of $\pi$ and $D$ are each subject to an upward bias.

The compartments were arranged in twelve size groups. The notification rate $y$ was calculated for each group and also a weighted average value of $U$ for all units in the group. The weighted average used was $\Sigma W (W - 1) / \Sigma W$. The twelve values of $y$ were plotted against values of $U$ and a freehand curve drawn to obtain an estimate of 3.50 per cent. for the upper limit $M$. Each value of $y$ was divided by $M$ to express the risk of infection in the unit size group. The least-squares method, with appropriate weighting, was used to determine the risk equation in its logarithmic form. It was then expressed in exponential form and the associated correlation coefficient found to be $0.215$ based on 804 degrees of freedom, compared with $0.16$ based on 805 degrees of freedom for the linear equation above.

The equation finally obtained in this way was

$$y = 3.50 \times (1 - (1 - 0.0085)^{(U - 2)}) \quad \text{(iii)}$$

That is to say, the carrier rate was estimated at 0.85 per cent., and the number of effective contacts from domestic and other non-working sources as 51. When $U$ is set equal to zero, $y$ has the value 1.23 compared with an average notification rate of 2.11. This implies that more than 40 per cent. of the cases of tuberculosis occurring in industry as a whole were the result of infection at work. In the larger compartments the proportion of cases so caused is, of course, higher, rising to about 60 per cent. in the largest units of all. The distribution of cases expected on the basis of the new equation is shown in Table IV (Model "C"), which illustrates the progress made by adopting a non-linear estimate of risk. In order to maintain the earlier Model "B" it would have been necessary to accept that a chance of about 1 in 25 had occurred. But if Model "C" is correct, then the differences between actual and expected cases are no larger than would occur in more than 1 in 3 of repeated samples.

(7) Effect of Factory Size

Even Model "C" does not exhaust the possibilities of accounting for variations in the notification rates in terms of contacts made at work. In Table V (overleaf) the actual notification rates in a number of factory size groups are compared with the rates which would have occurred if cross-infection within the factory had been confined to infection within compartments. It is evident that over and above the increase of risk with compartment size there is a further increase associated with factory size.

For each room in a factory the number of cases or fractions of case expected on the basis of Model "C" were summed, and the factory totals were summed in ten factory-size groups. A chi-squared test confirmed that the departure of actual from expected cases was statistically significant.

The close simultaneous relation between compartment size and factory size made
it impracticable to include them both as "independent" variables in the same estimating equation. Otherwise we could have obtained a still better estimator of risk of the form

\[ y = M \left\{ 1 - (1 - \pi)^{(T-U)/(T-U)} \right\}, \]

in which \( T \) is the total number of workers in the factory.

One result of taking factory size into account would have been a substantial increase in the estimated proportion of cases resulting from infection at work. Another result would have been to reduce the estimate of \( D \).

**DISCUSSION**

Is 40 per cent. an unreasonably high estimate of the proportion of tuberculosis cases originating in infection at work? A preoccupation with the contacts who can most easily be identified (i.e. those in the home and close friends) has given rise to the impression that tuberculosis is only passed between persons whose relationship is more intimate than that of fellow workers. But since it is virtually impossible to develop human type tuberculosis without meeting an infectious person, the figure 40 per cent. is actually quite small. Only 1 in 4 of Northamptonshire boot and shoe factory operatives who developed tuberculosis between 1920 and 1948 had any history of family contact, and among 2,829 pulmonary cases notified in Glasgow during 1949 only 509 (18 per cent.) had ever lived in the same house with a tuberculosis case of any kind (Laidlaw, 1949). In all, therefore, about three-quarters of tuberculosis cases remain to be accounted for in terms of non-household contacts. Since this is so, the remarkably good fit obtained with the second form of estimating equation justifies a discussion of the theoretical model on which it is based.

To illustrate the discussion, a diagram has been drawn in which risk of infection \( R \) is shown as a function of number of effective contacts \( F \), on four different assumptions about the carrier rate \( \pi \). These are the curves \( G, H, J, \) and \( K \). For \( G \) the value of \( \pi \) is set equal to the percentage of newly-discovered active cases of tuberculosis found by mass radiography (1945-46) amongst Northamptonshire workers not in the boot and shoe trade (i.e. 0.33 per cent.); for \( H \) to the corresponding percentage amongst boot and shoe workers (i.e. 0.59 per cent.); for \( J \) the value
of \( \pi \) is that estimated from notification data in the present analysis (i.e. 0.85 per cent.); for \( K \) the value of \( \pi \) has been set at the arbitrary figure of 5 per cent.

The Figure shows clearly the special relevance of the theoretical model to diseases with low carrier rates. Where there are large numbers of carriers (5 per cent. is a very modest estimate of the carrier rate for the common cold at certain times of the year) then only a few contacts are needed to subject a person to his maximum risk of infection. But when the carrier rate is below 1 per cent., meeting an extra one or two hundred persons in the course of the day's work can easily double or treble the risk of infection.

According to the diagram we should expect the incidence of tuberculosis to rise whenever two or more groups of persons are thrown together for the first time. This phenomenon is well known when civilized and primitive societies meet, but is one which the medical profession is apt to overlook when attempting to explain the prompt rise in the incidence of tuberculosis (and also of infective hepatitis) on the outbreak of war.*

The same argument applied to individuals throws light on morbidity at two

* During the first half of the 19th century there was an unprecedented migration of country folk into towns accompanied by explosive outbreaks of many infectious diseases. Since real wages were increasing and food supplies improving during the relevant period (Clapham, 1926), it is impossible to ascribe the outbreaks to a decline in the standard of living. One has, therefore, to conclude that the sudden collision of large numbers of susceptible persons was alone sufficient to cause giant epidemics of diseases such as cholera, typhoid, scarlet fever, and diphtheria, which in their endemic phase have low carrier rates. In this connexion it is of interest to note that during the long spell of unemployment in the cotton industry brought about by the American Civil War, the death rate in several Lancashire towns temporarily declined (Simon, 1862). Towards the end of the century when boot and shoe factories began to replace the old cottage workshops there was undoubtedly an improvement in living and working conditions, yet the excess of tuberculosis in the boot and shoe industry dates from this period (Cairns and Stewart, 1951).
particular ages. The first is that of the child moving from a home to a school environment at the age of five. At this age the peak incidence of measles and of whooping cough has already been passed, but the rarer exanthemata, such as scarlet fever, occur more frequently than before. The second is that of young adults, moving from a school to a working community, when a rise in the incidence of tuberculosis and lobar pneumonia occurs. So long as these two diseases are thought of solely in terms of resistance it is difficult to explain this phenomenon, but if they are once thought of in terms of infection the paradox is resolved.

In the case of tuberculosis, the incidence of the disease increases even more than would be expected from the curves in the diagram. This is probably due to the fact that the school-leaver, besides increasing his number of effective contacts, is moving into an adult community in which the carrier rate is relatively high. That is to say, besides moving to the right along one of the risk curves, he also shifts onto a higher curve.*

One of the assumptions which had to be made in the present calculations was that carriers were distributed at random. This may have been true originally in the boot and shoe trade, and may still be true in more strenuous trades, where the nature of the work tends to eliminate new cases of tuberculosis as they arise. But in comparatively light work, such as shoemaking, provided no deliberate measures are taken to eliminate cases, there will always be an excess of carriers in those parts of the industry where the most infection has been taking place, that is, in the largest working units. These are precisely the places in which the most chronic cases have been found and where, man for man, they produce the most victims (Stewart and Hughes, 1951). It follows that any given percentage of carriers in the boot and shoe industry as a whole will do more damage than the same percentage in an industry where carriers are evenly distributed.

The non-random distribution of carriers will have given an upward bias to the ex post estimate of the carrier rate (in addition to the bias already mentioned). Without this bias curve J would lie nearer to curve H. Curve L illustrates what would happen if there were no elimination of new cases: an entrant to the industry who added only fifty to his number of effective contacts would move from the point $Q_1$ to $Q_2$, a fivefold increase of risk.

Multiplication of risk on this scale wholly outweighs any effect due to age, sex, occupation, spacing of workers, and differences between "good" and "bad" factories as judged by factory inspectors. The first three of those variables are in any case not amenable to control, and neither re-spacing workers nor reconditioning factories, however desirable, is likely to have an appreciable effect on tuberculosis morbidity. There remain, therefore, only two courses open: to reduce the carrier rate, or to reduce the number of "effective contacts".

* The "control" group of young persons used in the Prophit Survey (Daniels and others, 1948) were said to be "at no greater risk of encountering infection than any town-dweller". Of controls under 17 years old, 76 per cent. were tuberculin positive, and of those between 18 and 20, 81 per cent. But in the course of B.C.G. trials in the industrial Midlands it has been found that out of 2,700 children in their last year at school only 969 (36 per cent.) were found to be positive (Hughes, 1951). If these findings are to be reconciled, we must assume that the amount of primary infection taking place in the first years after leaving school is far greater than in any other period of life.
MEASURING THE RISK OF INFECTION AT WORK

The first of these courses is the more promising. In mass radiography there exists an ideal weapon for the detection and removal of sources of infection. Active cases of tuberculosis discovered amongst male shoemakers in the 1945-46 survey were, on the average, 8 years older than the male cases notified other than by mass radiography during the same period.* This illustrates how such surveys can locate the chronic male cases which otherwise tend to escape diagnosis, and which constitute the hard core of infection. It is too early to assess how far this first survey has succeeded in reducing the rate of current infection, but if the theoretical model is correct, an initial halving of the carrier rate should have reduced the average risk of infection by more than half. In this connexion it is of interest to note that in the repeat mass radiography survey of Northamptonshire (1948-49) active tuberculosis was found no more frequently among shoemakers than among other workers, even in those individuals x-rayed for the first time (Fisher, 1951).

The pattern of infection revealed by recent research shows how to make continued control by mass radiography at once more efficient and more economical. The correct strategy will be to concentrate on the largest factories, where sources of infection are most frequent, most dangerous, and, for administrative reasons, easiest and cheapest to find. If, however, case-finding by mass radiography is pursued without providing alternative employment for tuberculous subjects after treatment, it may result in a falling off of volunteers in just those sections of the working population which it is most important to examine. This difficulty could be avoided if special factories, similar to those already provided under the Ministry of Labour Remploy Scheme, were available for all tuberculous subjects who are fit for work. Even though known tuberculous subjects are less dangerous sources of infection than untreated cases, it is important from the point of view of policy not to discriminate between these two types of case. For if it became generally known that suitable working conditions and economic security were guaranteed for all tuberculous persons, reluctance to submit to periodic examination might be reduced to vanishing point. At present a great deal of attention is paid to protecting members of infected households, but persons equally in need of protection are fellow-workers from uninfected houses.

The second possible line of action, namely reduction of the number of effective contacts, would in practice be difficult, if not impossible. The present tendency in the boot and shoe industry, as in others, is for factories to grow larger. This tendency is accelerated by the modern preference for vast, multi-process workshops instead of separate rooms for each process. More partitions, or at least the preservation of existing partitions, would be of some protective value against tuberculosis, and perhaps also against that troublesome cause of absenteeism, the common cold.

SUMMARY

At least 40 per cent. of the cases of tuberculosis occurring amongst workers in

* The average ages of these groups were (numbers of cases in brackets): mass radiography notifications, 45·3 years (56); other notifications, 37·5 (66). A similar age relationship was found for male workers in other trades.
Northamptonshire boot and shoe factories during 1940-49 may be attributed to infection at work. For large factories the figure may be as high as 60 per cent.

There is a definite association between risk of tuberculosis and numbers of persons working together, but no perceptible association between risk of tuberculosis and "overcrowding".

The hypothesis that risk of tuberculosis varies in direct proportion to the number of contacts at work, though not entirely worthless, is shown to be unsatisfactory. A second hypothesis is developed according to which risk bears a curvilinear relation to number of contacts. This provides an adequate explanation of the observed facts. From this hypothesis it is argued that:

1. Any given percentage of active cases of tuberculosis in a light industry constitutes a greater danger to fellow-workers than the same percentage in a heavy industry.
2. An excess of tuberculosis amongst boot and shoe workers implies neither sub-average powers of resistance, nor the existence of a specific occupational hazard, nor bad working or living conditions.
3. There is nothing paradoxical in the rising incidence of tuberculosis at young adult ages, although acquired specific resistance is higher at this age than in childhood.

Mass radiography has a vital role to play in the prevention of tuberculosis. Suggestions are made for the most efficient use of mass radiography resources.

The hypothesis advanced may have relevance outside the industrial field, and may apply to communicable diseases other than tuberculosis.

Our thanks are due to members of the boot and shoe industry who allowed us access to their factories, and to the Medical Officers of Health for Northampton County Borough and for Northamptonshire who put their records at our disposal. In the writing of the paper we have benefited from the advice of our colleagues, Dr. W. T. Russell and Dr. I. Sutherland, and of Mr. P. A. P. Moran of the Oxford University Institute of Statistics. The Medical Research Council contributed generously towards the expenses of the investigation.

REFERENCES