A mathematical model of a heroin epidemic: implications for control policies

DOUGLAS R. MACKINTOSH
From the Department of Management and Marketing, College of Business Administration, Louisiana State University in New Orleans

GORDON T. STEWART
From the Department of Community Medicine, University of Glasgow

SUMMARY An exponential model based on the infectious disease model of Kermack and McKendrick has been simplified to illustrate how the use of heroin spreads in epidemic fashion. A numerical simulation is arranged to show how the dynamics of spread are influenced by the original number of users, rates of conversion, and time of removal from the drug scene of those secondarily affected. The spread is significantly increased by small increases of those originally affected, in which case reduction of spread requires a large increase in rate of removal. The model indicates a strategy for intervention which is discussed in relation to policies for control of drug abuse.

The term epidemic is usually applied to the spread of infectious disease, but there are situations in which disorder can spread in similar fashion without infectious agents being involved. The spread of heroin habituation and addiction presents many of the well-known phenomena of epidemics, including rapid diffusion and clear geographic boundaries (Bejerot, 1969; Dupont, 1971; Hughes et al., 1972; Hughes and Crawford, 1972). Heroin is the agent, man is the host, the pusher or the heroin user is the vector, and the social milieu, usually but not always a lower socioeconomic neighbourhood, is the environment (Jonas, 1972). Heroin epidemics, as well as other epidemics involving communicable diseases, are characterised by three common factors—a state of susceptibility in the given community, introduction of a provocative stimulus, and some degree of sensory contact between those primarily and those secondarily affected. The dynamics of spread of such states of mind are poorly understood, and it is therefore of interest to explore the problem mathematically, reducing the factors to the essential minima even though the situation, in real life, is much more complex because of imponderable and often intervening biological, psychological, and social conditions.

In general, any process of social diffusion results in diverse forms of distribution of acts in time and among people. Previous mathematical models of social diffusion have dealt with the phenomena of airborne leafleting (Dodd, 1955; Dodd, 1958–59), rumour spread (Rapoport and Rehbn, 1953), and other novel behaviour including pest control (Chatterjee, 1972), reading a news item, buying a new product, following a new fashion, adapting imitatively to any new situation, educating children, and acculturating immigrants (Dodd and Winthrop, 1953). In these models an S-shaped logistic growth curve will occur in so far as an attribute spreads solely by random steady pairings; this principle seems to hold for any type of behaviour and group or location of act.

Stewart (1969) suggested the possibility that an exponential model might describe the development of a riot resulting from the conversion of riot-prone individuals. This model did not provide explicit representation of the effect of the rioters themselves on the riot-prone. In order to take this factor into account, a simple form of deterministic model of Kermack and McKendrick (1927) was developed (Mackintosh and Stewart, 1972). Like the Reed-Frost epidemic model (Abbey, 1952), this model is based on the assumption that at the beginning of each successive interval (incubation period) of an epidemic there is a constant probability of infectious contact between a case and a susceptible, that an infected susceptible at the end of the interval will become infectious, and that those
who were infectious at the beginning of the interval will become immune. The model allows for the occurrence of more than one infectious contact for a given susceptible (Serfling, 1952) so that if \( p \) is taken as the probability of an infectious contact, then \( q = 1 - p \) is the probability that an exposed susceptible will escape infection. If the susceptible has contact at time \( t \) with \( C_t \) infectious persons, the probability that the susceptible does not become infected is \( q^{C_t} \) and the probability of at least one infectious contact is \( 1 - q^{C_t} \). Since this probability is the same for all susceptibles, the expected number of new cases at the end of the interval is

\[
C_{t+1} = S_t (1 - q^{C_t})
\]

where \( S_t \) is the number of susceptibles at the time \( t \) and \( C_{t+1} \) is the number of new cases at time \( t + 1 \).

In the present model, a susceptible is defined as an individual who is exposed to heroin because he lives in a particular social milieu but has not used the narcotic recently although he or she may have used other drugs or be an ex-user of heroin. A susceptible may also be an individual who is motivated to sell or distribute heroin but has not actually done so within the neighbourhood (Prebel and Casey, 1969). The affected vector is defined as an individual who lives in the same environment and either (1) experiments with, is habituated to, or is addicted to, heroin or (2) is a non-user of heroin, but sells or distributes the drug. The term ‘affected’ is employed to denote the difference between the heroin user and the ‘infected’ person in a microbial infectious process. Since heroin is powerfully addictive it is also assumed that during the heroin epidemic there is no sudden reversion to the susceptible phase. Consequently,

\[
C_{t+1} = S_t \left( 1 - q \sum_{i=1}^{C_t} i \right)
\]

In order to examine the effect of physical removal of affected vectors, the assumption was made that at the end of period \( k \) and at each subsequent period, a fixed proportion, \( I_t \), of affected were removed. Removal is defined as death, relocation in another environment, imprisonment, and closed ward treatment, plus any rehabilitation, detoxification, immunisation, maturing out, or change in occupational or psychological disposition to the extent that the previously affected individual will no longer convert susceptibles.

If the total population \( n = s + c + r \), where \( r \) = removals, the basic differential equation is

\[
\frac{dr}{dt} = Y (n - s - r) \]

where \( Y \) = rate of removal. Let the number thus removed be designated by \( R_k (I) \).

The difference equation for \( C_{t+1} \) then becomes

\[
C_{t+1} = S_t \left[ 1 - q^{C_t} - R_t (I) \right], \ t > k \tag{3}
\]

in which \( C_t \) represents the number of affected at time \( t \) after allowance for removal of \( R_j (I) \) affected at times \( j = k, k + 1, \ldots, t - 1 \)

\[
R_t (I) = I : C_t \tag{4}
\]

represents the number of removals at the current time \( t \) and

\[
C_t = \sum_{j=t-1}^{j=0} C_j \sum_{j=k}^{j=t-1} R_j (I)
\]

The length of the interval \((t, t + 1)\) is not specified and it is acknowledged in models of this kind that there are uncertainties due to variation in the time interval between exposure of a ‘susceptible’ subject to an infectious one, as well as by variation in the duration of the incubation period of the ‘infection itself (Bailey, 1975). These uncertainties may be assumed to be of a lesser order in the present model of the drug scene because there is no latent period and the only contact which we are considering is the effective one, that is, the acceptance and continued use of heroin by the ‘susceptible’, in which case the ‘incubation period’ between exposure and use of the drug is usually a matter of hours rather than days. We are not considering here, any more than in conventional models of infection, the behavioural variables which influence conversion of the ‘susceptible’ to a receptive state because, until this conversion has occurred, he need not be regarded as being susceptible. If allowance has to be made for variations in the chance of infection, this could be done by the substitution \( p = 1 - e^{-\lambda t} \), where \( \lambda \) = the effective contact rate and \( \alpha \) = the infectious period. It is obvious that, under the conditions described here, the alteration in \( p \) by this substitution will be small.

**Numerical Illustration**

Reports of recent heroin epidemics are beginning to accumulate in the literature (Bewley, 1968; de Alarcon, 1969; Cherubin et al., 1972; Hughes et al., 1972; Hughes and Crawford, 1972), but there is a scarcity of numerical data relating to incidence, intervention, and conversion rates. In epidemics of drug abuse and sniffing observed in schools in New Orleans, New York, and Glasgow, however, we found, like Oliver (1977), that it was not uncommon for 20 to 50 persons to become affected within a few
A mathematical model of a heroin epidemic: implications for control policies

weeks. For this example, therefore, the initial number of involved persons, susceptibles, and affected vectors was set at \( n = 50 \). When \( p = 0.08 \), the conversion of susceptibles to affected was so rapid that all susceptibles were converted to the affected stage at or before \( t = 4 \). With \( p = 0.02 \) and only one affected against 49 susceptibles at \( t = 0 \), all 50 persons would convert to affected vectors by 10 time-intervals. The conversion rate \( p \) is not necessarily constant for each time interval, and this variable is represented by \( p_t \) for \( t = 0, 1, \ldots, 10 \). The initial conversion rate \( p_0 \) equals 0.02 in all subsequent variations of the parameters.

As the number of affected, \( C_0 \), increases, the RT50 (or time at which 50% of the population is affected) decreases (Table 1). The number of person-time units of involvement after 10 time-periods (\( D_{10} \)) is a possible indicator of the amount of crime, both personal injury and property damage, which increases proportionally as \( C_0 \) increases and RT50 decreases. Small variations in the number initially affected clearly affect the development of the epidemic.

The effect of removal policies was examined on the assumption that the initial number of affected is \( C_0 = 4 \) (Table 2). The effect of removal is shown for a constant rate of conversion (\( p_t = 0.02 \) for \( t = 0, 1, \ldots, 10 \)), an increase in the rate of conversion after initiation of removal,

\[ \Delta p_t = p_{t+1} - p_t = 0.01, \ t = k, k + 1, \ldots, 10 \]

and a decrease in the rate of conversion after initiation of removal,

\[ \Delta p_t = p_{t+1} - p_t = 0.001, \ t = k, k + 1, \ldots, 10 \]

The consequences of no removal of affected, removal of 20% and 80% of affected, \( R_1 (0.02) \) and \( R_1 (0.8) \), respectively, at times \( k = 1 \) and each subsequent period, are indicated in Table 2.

Obviously the conversion rate might change according to whether removal is voluntary or involuntary, how the community perceives the justice involved in the removal, and whether certain exogenous variables change. There are significant differences in RT50 and \( D_{10} \) with different removal rates (Table 2). Delay in removal until the sixth time-interval, \( k = 6 \), has a striking effect (Table 3), but \( D_{10} \) varies little with either increases or decreases in the conversion rate. One time unit could be any period from a day to a month but our impression is that usually it would be about three to seven days. In social studies in general and in studies of drug addicts in particular, reliable data and hard end-points are elusive. In the present study it is assumed that census information is not available and that data are no more accessible than usual. Hence the need for simulations as pointers to logistic possibilities or impossibilities with given sets of numbers.

**APPLICATION TO CONTROL POLICIES**

The main use of a mathematical model in this field is to assist in conceptualisation and in definition of primary parameters. The model described also has a possible application to policy-making in demonstrating the extent of change as a product of the period of time over which each person is affected, that is, in person-time units. The entire decision tree, with change at \( t = 10 \) (Figure), then shows more clearly the results of intervention. Postponing the removal decision results in rapid spread and more damage.
Figure  Summary of results of intervention at different times and rates of removal (n = 50, P₀ = 0·02, C₀ = 4).

We are studying here only the general rate of development of a heroin epidemic, with and without intervention at various stages. Whenever the propensity to use or ‘push’ heroin is high (significant peer pressure or personal problems; large profits from sales) early intervention could be decisive; if intervention is delayed its effect becomes smaller until a point is reached when, to be effective, intervention has to be on a massive scale. Early action is the most effective action, but there is a point beyond which action is too small to affect the outcome appreciably, especially since a high removal rate might in many situations be tactically impossible. A simulation model has reaffirmed this difficulty (Levin et al., 1972).

Intervention at the commencement of the sixth time-period has to be on a major scale to reduce the number affected to less than 50% of the total population in less than three subsequent time intervals. When the decision to intervene is taken, there is very little difference from that point onward in the effect of modifying the conversion rate of ‘susceptible persons’ to ‘affected’ since the residual proportion of ‘susceptibles’ is small.

The implications of removing one source of supply, that is, the drug ‘pusher’, as well as the addict himself, are obvious and suggest, numerically at least, how drug dependence might be reduced. Intervention would necessarily have to be selective among users, since some users do not sell heroin to support their habit nor do they attempt to influence others to use the narcotic. Those individuals who have affected others could be removed by such techniques as compulsory closed-ward treatment, immunisation with long-acting antagonists like cyclazocine, or, as a last resort, imprisonment. One problem with these techniques is the high incidence of relapse.

Alternatively, the drug or a suitable substitute (methadone) could be made available at relatively low cost in a supervised out-patient programme, as is done in many American cities. Legalised distribution of a substitute drug lessens the motivation of the affected vector to convert the susceptible, but some removal of addicts on programmes may still be necessary to ensure that peer pressure, particularly from affected relatives and friendship groups, plus brokering of the legal supply, do not convert susceptibles. There is evidence of some discrimination among individuals about the distribution of drugs (Moore, 1973). Heroin would be more difficult for the susceptible novice to obtain than for the experienced non-infectious addict to procure.

Alternatively, or perhaps concurrently, the number of susceptibles could be reduced. Non-users of dangerous drugs tend to avoid susceptibility because of satisfaction gained in exploring positive alternatives, rather than fear of consequent harm (Cohen, 1972). If there was a legal, low cost supply, the number of individuals who would be motivated to sell heroin would be reduced. The longer the time that a vulnerable individual stays in the susceptible state, that is, the longer he or she is exposed to drug use and peer pressure, the more likely it is that that
individual will be converted (Bowden, 1971). The problem with the differential-association theory and this mathematical model is that not all susceptibles or those who are exposed to drug-related associations become drug users or ‘pushers’. However, contact between a susceptible and an affected vector is an important predisposing factor in the developmental pathology of heroin addiction (Hughes et al., 1972).

Although sophisticated sampling techniques have been developed to determine the number of narcotic addicts within an area (Greenwood, 1971; 1977), it is equally important that methods be developed to determine the total number of affected vectors, the number of susceptibles (the number ‘at risk’), and the conversion rates for both micro- and macro-epidemics in homogeneous and heterogeneous communities. In a homogeneous community, peer pressure is a dominant factor (Smith and Gay, 1972) whereas, in a heterogeneous community, the influence of pushers offering inducements in an otherwise hostile environment is often decisive (Stewart, 1972). In this connection it may be noted that psychopathic deviance and high levels of neuroticism may predisperse not only to drug-seeking behaviour but also to relapse during remedial activities (Waddell and Stewart, 1972).

Whatever the intermediate mechanism, the outcome in real life depends primarily upon the number of users or sources of a drug, the possibilities for spreading the habit, and the numbers and state of mind of the non-drug-using population. It is therefore logical as a first step toward rational control policies to have a means of analysing the dynamics of spread according to the numbers of users and susceptibles and the rate of spread, which is what this model does in deterministic terms.

We thank the National Institute of Mental Health (research grant No. MH 18229-01) and the Graduate Research Council of Louisiana State University in New Orleans for partial support.

Reprints from Professor Gordon T. Stewart, Department of Community Medicine, University of Glasgow, Ruchill Hospital, Glasgow G20 9NB.

**SUMMARY OF SYMBOLS**

- \( \alpha \) = infectious period
- \( \lambda \) = contact rate
- \( N \) = number of susceptibles + number of affected persons removed
- \( N_0 \) = number of affected at the beginning of the outbreak
- \( C_t \) = number of new affected produced during the time period \( t - 1 \) to \( t \)

\[
S_t = \text{number of susceptibles available at the commencement of time period } t
\]

\[
PO = \text{conversion rate of susceptibles to affected at the beginning of the outbreak}
\]

\[
p_t = \text{conversion rate of susceptibles to affected at time periods during the outbreak}
\]

\[
\Delta p_t = \text{increment of change in conversion rate of susceptibles to affected at time period } t + 1 \text{ and thereafter following initiation of removal at time } t = k
\]

\[
R(t) = \text{number of removals if the proportion of affected } I \text{ are removed at time } t = k \text{ and each subsequent time period}
\]

\[
RT_{50} = \text{time at which } 50\% \text{ of the population is affected}
\]

\[
D_{10} = \text{person-time units of involvement after 10 time-units of the outbreak}
\]

**References**


Douglas R. Mackintosh and Gordon T. Stewart