Taylor & Francis
Taylor & Francis Group

A classification system for hospital-based infection outbreaks

Paul S. Ganney[a,b]*, Maurice Madeo[c1] and Roger Phillips[b2]

[a]*Department of Medical Physics, Hull and East Yorkshire Hospitals NHS Trust, Hull, England;*
[b]*Department of Computer Science, University of Hull, Hull, England;* [c]*Infection Control, Hull and East Yorkshire Hospitals NHS Trust, Hull, England*

Outbreaks of infection within semi-closed environments such as hospitals, whether inherent in the environment (such as Clostridium difficile (C.Diff) or Methicillin-resistant Staphylococcus aureus (MRSA) or imported from the wider community (such as Norwalk-like viruses (NLVs)), are difficult to manage. As part of our work on modelling such outbreaks, we have developed a classification system to describe the impact of a particular outbreak upon an organization. This classification system may then be used in comparing appropriate computer models to real outbreaks, as well as in comparing different real outbreaks in, for example, the comparison of differing management and containment techniques and strategies. Data from NLV outbreaks in the Hull and East Yorkshire Hospitals NHS Trust (the Trust) over several previous years are analysed and classified, both for infection within staff (where the end of infection date may not be known) and within patients (where it generally is known). A classification system consisting of seven elements is described, along with a goodness-of-fit method for comparing a new classification to previously known ones, for use in evaluating a simulation against history and thereby determining how 'realistic' (or otherwise) it is.

1. Introduction

A semi-closed system such as a hospital environment is susceptible to certain types of infection outbreak, especially those transmitted by contact or proximity. Given that hospitals are generally warm environments populated primarily by humans with reduced immune systems, they require careful management in order to avoid infections such as Noroviruses (Norwalk-like viruses or NLVs, generally regarded as non-life threatening) [10], seriously reducing the capability of the organization to perform its primary function – the treatment of the ill.

In describing an outbreak, there appears to be no simple method of doing so – certainly nothing as simple and useful as the Richter scale for earthquakes [13].

Several computer models have been constructed in recent years ([3,11,12,14] are examples) in order to try and mimic the impact of such outbreaks. One difficulty experienced by such investigators is the comparison of the modelled results with those from real outbreaks.

*Corresponding author. Email: paul.ganney@hey.nhs.uk

This work describes a classification method for comparing outbreaks within a semi-closed environment (*e.g.* where differing management and containment techniques have been applied), simply describing them (for news and executive reporting purposes) and to aid in the evaluation of simulated results. The NHS Trust from which the data were drawn consisted of three sites at the time the data were taken, with 800 beds spread across 58 wards (34, 21 and 3 per site). It employs approximately 8000 staff and sees 120,000 inpatients and 130,000 outpatients per year drawn from a population of 600,000 in the immediate area and 1.2 million in the wider area.

2.   Description

In producing a classified description of an outbreak, the primary metrics are the length of the outbreak (number of days) and its prevalence (how widely the infection permeated the organization). There are, however, several ways of reporting these and each will show a different impact.

For the purpose of the classification system to be described, an outbreak is taken as being a continuous infected state, where an infected state exists when the infection is present in the entity/entities under investigation. For example, when classifying patient data, an outbreak commences when a patient is recorded as having the appropriate symptoms and concludes when all patients are non-infectious (which may be the cessation of symptoms or be a defined period beyond it – in the case of Norovirus, this is 48 h post asymptomatic [5]). This continuous infected state is not restricted to a single ward but is measured across the whole organization (Figure 1, in which Wards 30 and 31 would constitute a single outbreak even if all other wards were discounted).

The classification system described here uses a seven-digit number where each represents a different element of the classification:

$$< \text{Extent} >< \text{Amp} >< \text{LSpan} >< \text{AvIncidence} >< \text{Vol} >< \text{PkSev} >< \text{Recurr} >,$$

where Extent is the infection extent as a proportion of the organization, Amp is the amplitude of infected wards, LSpan is the infection lifespan, AvIncidence is the average incidence per ward, Vol is the infection volume (see below for a description of this metric), PkSev is the peak severity index of the infection (see below for a description of this metric) and Recurr is the infection recurrence (does it recur soon after removal from the organization?).

The elements were refined over several iterations and therefore the thresholds reported here may be descriptive of the Trust only. The element thresholds are not linear, as Figure 2 illustrates. These were selected so that the metrics exhibit sufficient granularity without becoming overly cumbersome.

2.1   *Infection extent*

0: The infection does not spread beyond the initial individual.
1: The infection spreads to 1–1.9% of the wards (containing the case where the outbreak is restricted to a single ward).
2: The infection spreads to 2–2.9% of the wards.
3: The infection spreads to 3–3.9% of the wards.
4: The infection spreads to 4–9% of the wards.
5: The infection spreads to 10–24% of the wards.
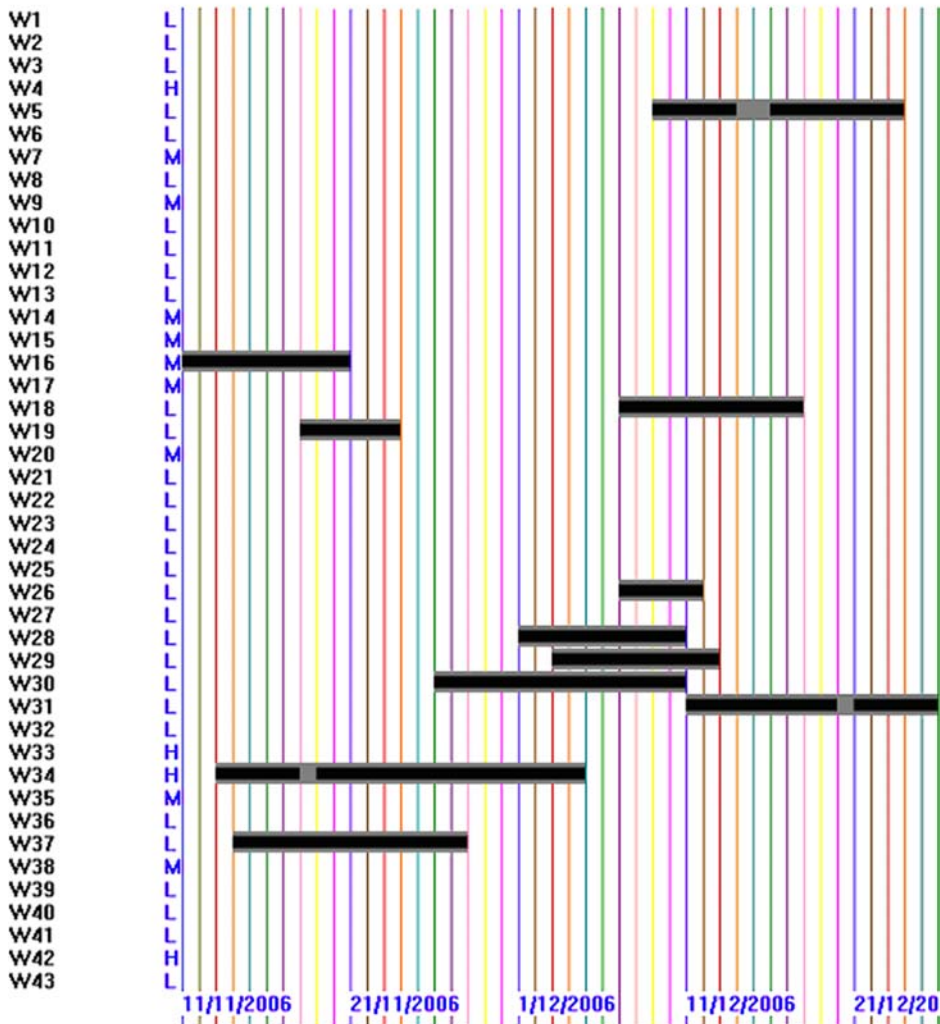6: The infection spreads to 25–49% of the wards.

Figure 1. Pictorial representation of an outbreak. The black bars show the time that an infection is present on a ward, with the grey ones show the time span between the start of the first infection and end of the final one. The ward names have been anonymized. The letters (L, M and H) alongside the ward names are the assigned severity values. Dates are attached to the left-hand grid line.

   7: The infection spreads to 50–74% of the wards.
   8: The infection spreads to 75–99% of the wards.
   9: The infection spreads to 100% of the wards.

2.2 Amplitude of infected wards

This is the maximum number of wards where the infection exists within a single time.

   0: The infection is present in <1% of the wards at one time.
   1: The infection is present in 1–1.9% of the wards at one time.
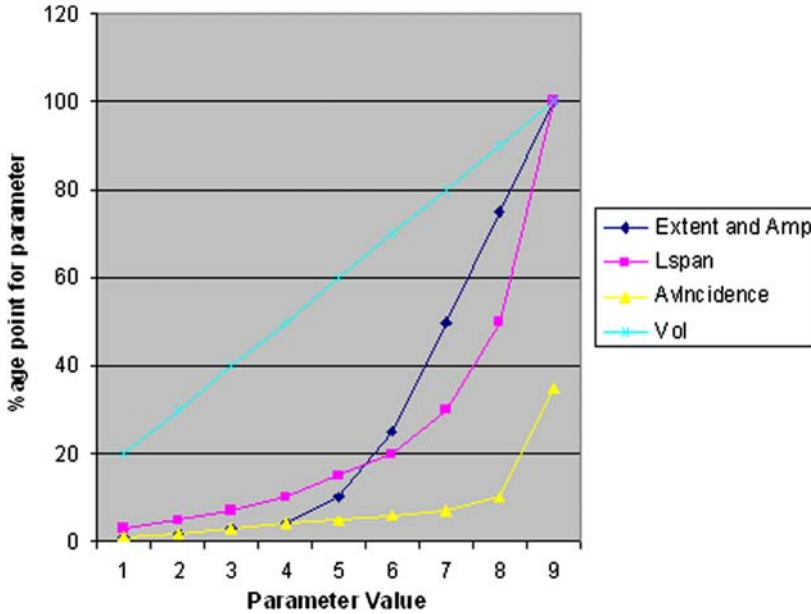   2: The infection is present in 2–2.9% of the wards at one time.

Figure 2.   The threshold points for each metric showing the non-linearity of them. Recurr is omitted from this diagram, as its values are flags rather than thresholds. PkSev is omitted, as its values are non-continuous.

   3: The infection is present in 3–3.9% of the wards at one time.
   4: The infection is present in 4–9% of the wards at one time.
   5: The infection is present in 10–24% of the wards at one time.
   6: The infection is present in 25–49% of the wards at one time.
   7: The infection is present in 50–74% of the wards at one time.
   8: The infection is present in 75–99% of the wards at one time.
   9: The infection reaches all the wards at one time.

2.3   Infection lifespan

('Maximum' is set according to the investigation – see 'Results')

   0: The infection lasts <3% of maximum.
   1: The infection lasts 3–4.9% of maximum.
   2: The infection lasts 5–6.9% of maximum.
   3: The infection lasts 7–9.9% of maximum.
   4: The infection lasts 10–14.9% of maximum.
   5: The infection lasts 15–19.9% of maximum.
   6: The infection lasts 20–29.9% of maximum.
   7: The infection lasts 30–49.9% of maximum.
   8: The infection lasts 50–99% of maximum.
   9: The infection reaches maximum.

2.4   Average incidence per infected ward

This is the average number of incidences on infected wards, that is, the total number of incidences divided by the number of wards they occurred in. This is not the same as the

incidence rate, as it does not relate to the size of the pool of susceptibles.

   0: Average of 0–1 infection per ward.
   1: Average of 1.01–2 infections per ward.
   2: Average of 2.01–3 infections per ward.
   3: Average of 3.01–4 infections per ward.
   4: Average of 4.01–5 infections per ward.
   5: Average of 5.01–6 infections per ward.
   6: Average of 6.01–7 infections per ward.
   7: Average of 7.01–10 infections per ward.
   8: Average of 10.01–35 infections per ward.
   9: Average of more than 35 infections per ward.

2.5 Infection volume

This is a measure of the overall prevalence of the outbreak and is the sum of distinct days/wards for which the infection exists. For example, a 10-day infection which only exists in a single ward has a volume measure of 10. Similarly, a 10-day infection which lasts for 5 days in one ward and then 5 days in another has a volume measure of $5 + 5 = 10$. Finally, a 10-day infection that lasts for 10 days in each of the two wards has a volume measure of $10 + 10 = 20$, showing that it has pervaded most of the organization. This measure is analogous to that of fractal dimension [9], in that it describes the extent to which an infection has filled the available space. In order to make this measure useful, the 'available space' is limited to the number of days for which the outbreak lasts and to the number of affected wards (as opposed to the maximum for each). This measure is therefore a reflection of how non-sparse the infection extent is. It has similarities with cumulative incidence [6] but does not investigate the number of individual incidences, only the number of areas in which it has occurred. It is also a measure of where the infection currently exists, rather than new cases. If the time period under consideration is that measured from the onset of the outbreak, and wards are considered as individual units, with the population being only wards that are infected, then they are the same, although Vol is probably closely linked to cumulative prevalence. As this is a specialized case, the term 'Vol' is used to avoid confusion. Figure 1 is worth referring to for a visualization of this element.

   ('Maximum' is set according to the investigation – see 'Results')

   0: The infection volume is <20% of maximum.
   1: The infection volume is 20–29% of maximum.
   2: The infection volume is 30–39% of maximum.
   3: The infection volume is 40–49% of maximum.
   4: The infection volume is 50–59% of maximum.
   5: The infection volume is 60–69% of maximum.
   6: The infection volume is 70–79% of maximum.
   7: The infection volume is 80–89% of maximum.
   8: The infection volume is 90–99% of maximum.
   9: The infection volume is 100% of maximum.

2.6 Peak severity index

This is a measure of how severe (in terms of impact on the Trust) the outbreak is. Each ward is given a severity measure: high, medium or low. High is given to wards such as

Table 1.   The thresholding classification system for the PkSev element.

Criteria	PkSev value	Low	Medium	High
	0	<5	0	0
	1	5–9	0	0
	2	<5	1	0
	3	5–9	1	0
	4	<5	2	0
	5	5–9	2	0
1 of	6	10 OR	3 OR	1
2 of	7	10 OR	3 OR	1
1 of	8	>10 OR	>3 OR	>1
	9	Any	>3	>1

Notes: For example, a severity of 3 is attained if 5–9 wards classified as 'low severity' are infected at the same time as one ward is classified as 'medium' and no wards classified 'high' are infected.

intensive care units, where only one or two exist (so patients may not be relocated), and wards where patients are already in very poor health and where relocating patients may not be possible. Medium is given to wards where only a few exist in that specialty so that relocating patients (or diverting the intake) is highly problematic. Low is given to general medical wards, of which there are many (so relocation is very feasible) and where outbreaks are common so that procedures for tackling outbreaks are well known and simply executed. The assumption is that the Trust reaches a critical state when 10 low wards are infected simultaneously or 3 medium wards or 1 high ward. To reflect this, the thresholding classification in Table 1 was developed.

2.7   Infection recurrence

  0: Recurrence is not searched for.
  1: Once the hospital is clear of infection, it stays uninfected.
  2: The hospital is clear of infection at some point but is reinfected within 2 weeks.

Element Recurr is a flag, whereas the others are measures within a continuum (PkSev could be regarded as being either but is more closely allied to a continuum).

3.   Results

Two sets of data for NLV outbreaks were examined initially: those for staff (September 2004–November 2007) and those for patients (August 2004–December 2007). As no end date was available for staff data, an infection length of 5 days was used (derived from [10,5] and Trust data). This was also used for patient data where no end date was recorded. For patient data with a recorded date, when the patient was asymptomatic, the end date for the infection was taken as the asymptomatic date plus 2 (as NLV is infectious for 48 h beyond the cessation of symptoms [5]).

   The Trust has 58 wards and the maximum extent of an outbreak was set at 100 days, thus giving a maximum volume of 5800. Although the Trust is mainly based on two sites, it is treated as one organization for the purpose of this analysis.

   Staff and patient data were analysed separately and in two ways: recurrence not checked for (so each outbreak is treated as independent) and where it is checked for (so outbreaks previously identified may combine).

3.1 Staff data

The data showed 38 distinct outbreaks, which reduced to 17 when recurrence was checked for. These were classified and these classifications were then statistically analysed for mean, median, mode and standard deviation of each element. The modal value for the full classification was also determined – the results are presented in Tables 2 and 3.

It should be noted that the mean and standard deviation cannot be taken literally, as the classification system is not linear and certainly does not occupy a continuous space. The values are included in order to give an indication of spread. The most common classifications (taking each element individually) are therefore 1120900 and 1173902, that is, the infection spreads to between 1 and 2% of the wards (and interestingly between 1 and 2% at any one time), with a lifespan of 2–7 days (increasing, as might be expected, to 30–49 days when recurrence is checked for) and an average incidence of 0–1 (increasing to 3–4 for recurrence, which again might be expected). The infection volume is 100% of the maximum (*i.e.* all infected wards are infected for the full life of the outbreak) and the PkSev index is 0 (less than five low risk wards infected). Vol remaining at 100% when recurrence is checked for is counter-intuitive. However, a comparison of Figures 3 and 4 shows that, while the values are still skewed towards 0, the range has been extended from 4 to 5 (a maximum possible extension from 9 to 24%). Vol has polarized having either high or low values.

When the classification is taken as a whole, the most common classification is 0120900. The only difference is in the Extent element, showing that the most common overall classification comes from a low spread, whereas it is higher spread when the elements are taken individually. The use of the other statistics can then be seen in explaining the preservation of the modal value for LSpan (which would be expected to be higher for a multi-ward outbreak than a single-ward one). While the modal value is 2 in each case, the mean and median are closer to 4 with a standard deviation of 1.55, implying a spread of 0–8, which can be clearly seen in Figure 4. Equally, the final statistical line

Table 2. Classification results for staff data, where recurrence is not checked for.

	Classification	Extent	Amp	LSpan	AvIncidence	Vol	PkSev	Recurr
Mode	0120900	1	1	2	0	9	0	0
Mean		2.13	2.29	3.74	2.92	6.87	1.16	0.00
Median		1.00	1.00	4.00	3.00	9.00	0.00	0.00
SD		1.79	1.49	1.55	2.21	2.64	2.11	0.00
Md–Mn		1.13	1.29	1.74	2.92	2.13	1.16	0.00

Notes: Each element is analysed separately for each statistical measure, whereas only the mode is determined for the full classification.

Table 3. Classification results for staff data, where recurrence is checked for.

	Classification	Extent	Amp	LSpan	AvIncidence	Vol	PkSev	Recurr
Mode	1144901	1	1	7	3	9	0	2
Mean		3.06	2.71	5.29	3.65	4.94	1.65	1.53
Median		4.00	3.00	6.00	3.00	3.00	0.00	2.00
SD		2.08	1.65	2.23	2.03	3.90	2.67	0.51
Md–Mn		2.06	1.71	1.71	0.65	4.06	1.65	0.47

Notes: Each element is analysed separately for each statistical measure, whereas only the mode is determined for full classification.
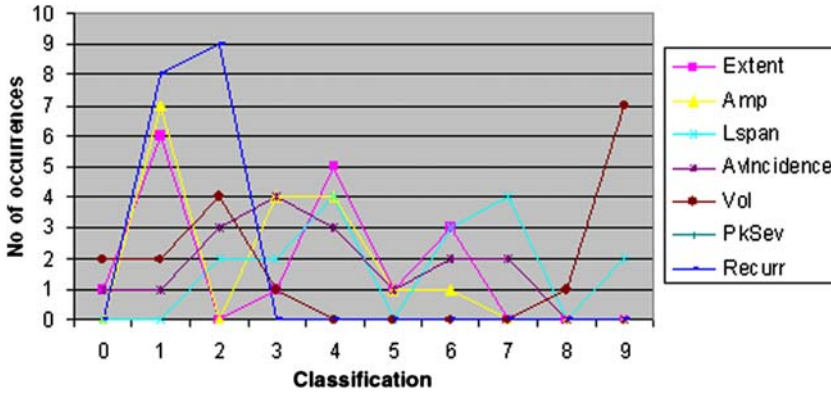
Figure 3. A graphical representation of the spread of classification values for each element, showing how often a particular value appears (staff data, recurrence checked for).
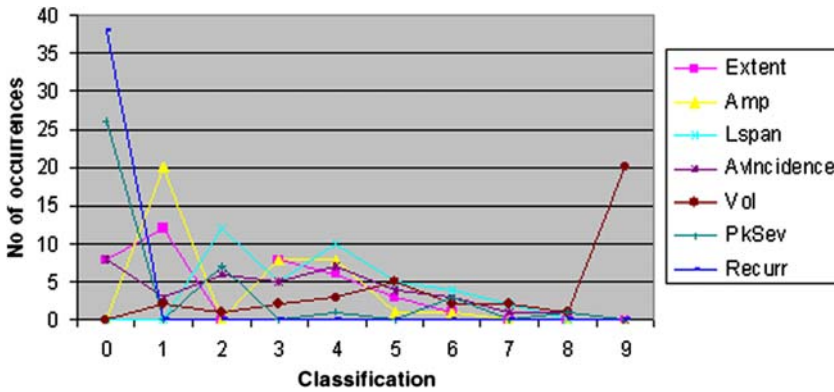

Figure 4. A graphical representation of the spread of classification values for each element, showing how often a particular value appears (staff data, recurrence not checked for). Recurr only has a limited range (0, 1 or 2), but the other elements are well distributed, tailing off in frequency at the higher levels, as would be hoped. The high incidence of Vol = 9 is due to single-ward outbreaks and is discussed below.

(the absolute value of the Mode–Mean, Md–Mn) shows how closely the mode represents the data, with low values denoting a closer (and therefore 'better') fit. When recurrence is checked, the modal classification is 1144901, showing a difference in the Extent, LSpan, AvIncidence and Recurr elements. All of these have increased, as might be expected.

As PkSev index is a measure of how severe the outbreak has been, it is a useful measure on a daily basis also, showing how severe an outbreak is at any point in time (Figure 5).

3.2 Patient data

The data showed 31 distinct outbreaks (reducing to 13 when recurrence is checked) which were analysed as before and are presented in Tables 4 and 5.

The most common (modal) classifications (taking each element individually) are therefore 1148900 and 1147901, that is, the infection spreads to between 1 and 2% of the wards (another analysis shows that this is single-ward incidence when recurrence is not

Figure 5. Graph of the daily severity index for the outbreak in Figure 1. The letters beside the ward names in Figure 1 are the severity measures for those wards, so it can be seen that the index reaches 6 when a ward rated high (in this case W34) is infected.

checked; interestingly between 1 and 2% at any one time), with a lifespan of 10–15 days and an average incidence of 10–35 (or 7–10 for recurrence). The infection volume is 100% of the maximum and the PkSev index is 0 (less than five low-risk wards infected).

The reduction in AvIncidence when recurrence is checked is counter-intuitive, but an examination of Figures 6 and 7 shows that an element value of 8 has remained constant, whereas the occurrences of value 7 have increased (as we would expect) sufficiently so to become the modal value.

Table 4. Classification results for patient data, where recurrence is not checked for.

	Classification	Extent	Amp	LSpan	AvIncidence	Vol	PkSev	Recurr
Mode	120900	1	1	4	8	9	0	0
Mean		2.23	2.29	4.32	5.39	6.65	1.16	0.00
Median		1.00	1.00	4.00	7.00	9.00	0.00	0.00
SD		1.86	1.64	2.01	2.94	3.09	2.28	0.00
Md–Mn		1.23	1.29	0.32	2.61	2.35	1.16	0.00

Notes: Each element is analysed separately for each statistical measure, whereas only the mode is determined for the full classification.

Table 5. Classification results for patient data, where recurrence is checked for.

	Classification	Extent	Amp	LSpan	AvIncidence	Vol	PkSev	Recurr
Mode	#N/A	1	1	4	7	9	0	1
Mean		3.38	2.85	5.62	5.92	4.00	1.85	1.46
Median		3.00	3.00	6.00	7.00	2.00	0.00	1.00
SD		1.98	1.91	2.57	2.06	3.76	2.97	0.52
Md–Mn		2.38	1.85	1.62	1.08	5.00	1.85	0.46

Notes: Each element is analysed separately for each statistical measure. The mode could not be determined for the full classification.
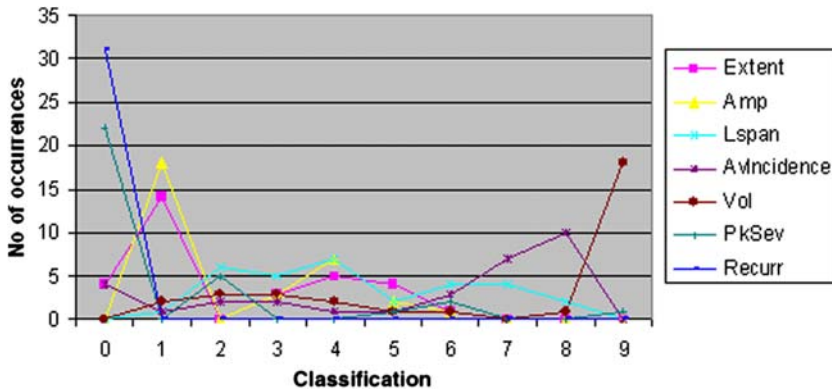
Figure 6. A graphical representation of the spread of classification values for each element, showing how often a particular value appears (patient data, recurrence not checked for).
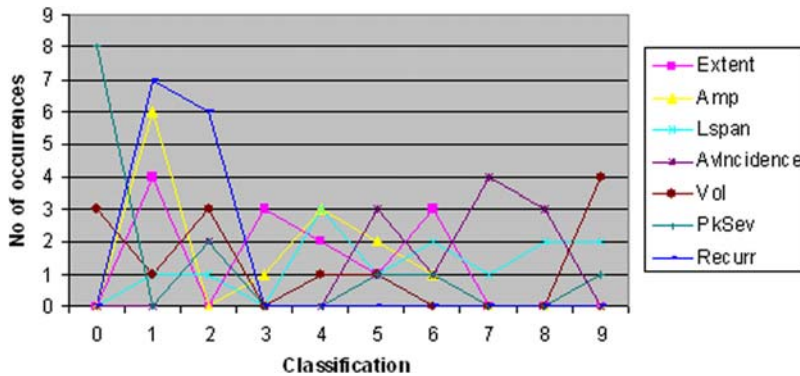

Figure 7. A graphical representation of the spread of classification values for each element, showing how often a particular value appears (patient data, recurrence checked for).

Interestingly, the modal classification (0120900) is the same as for staff. There is no modal classification when recurrence is checked as no value occurs more than once.

4. Discussion

It should be noted that this classification system is a descriptive (in the style of Richter) rather than being a predictive one (in the style of the modified Duke system for diagnosing infective endocarditis [8]). The use of this description will now be examined.

For staff, the most common classifications are 1120900 and 1173902. For patients, these are 1148900 and 1147901. Although these look similar, there are clear differences. The SD, although not a statistically useful value (due to the non-continuous nature of the variable space), assists here in giving a picture of the spread of values and therefore how 'tightly' the modal value can be considered to fit the data. For example, LSpan for staff is 2 or 7, with SDs of 1.55 and 2.23 (a loose fit), giving ranges of −2.65 to 6.65 and 0.31–13.69 within which 2 fits and 7 is close to fitting – therefore, they are not as wildly different as might at first appear.

Some elements are interlinked, and this can aid in validation. For example, an Extent value of 0 (single individual) must have the lowest possible Amp value (determined by the number of wards in the Trust) and a Vol value of 9. For Extent = 0, an AvIncidence value

of 0 and a Recurr value of 0 or 1 (due to immunity, a single re-infection should not happen) are the only possible outcomes. For Extent = 0, LSpan should be in a range commensurate with the lifespan of the infection. Extent and Amp are linked, in that Amp can never be greater than Extent. Examining the data, we see that this is borne out in that when Extent = 0, Amp = 0 and Vol = 9. When Extent = 0, AvIncidence = 0, Recurr = 0 and LSpan = 2 mostly (5–7 days) and 4 on one occasion (10–15 days). Also, Amp is never greater than Extent.

Likewise, some elements can be predicted for certain common scenarios. For example, an outbreak of NLV that does not spread beyond a single individual will have a classification as described above. Thus, for the Trust examined here, a value of 01(0/1/2/3/4)09(0/2/6)* would be expected, with 01209(0/2/6)(0/1/2) for staff data, as the lifespan is fixed at 5 days. In the non-recurrent data, we see 12 outbreaks with Extent values of 0 (8 for staff and 4 for patients). Nine of these have classifications of 012090* with two of 0120920 and one of 0140900, which is consistent as the last classification here is for patient data. This result has an LSpan value within that predicted but examining the data (single patient: symptoms commenced 7 February 2005, asymptomatic 15 February 2005, infectious until 16 February 2005, giving a lifespan of 10 days and thus the LSpan value of 4) shows a lifespan within a statistically possible range (3–15 days) but, as evidenced by the single occurrence of this value, at a less likely point. For validation purposes, it is thus better to describe the expected value as 01(0/1/2)09(0/2/6)* and investigate any that do not comply.

Likewise, an outbreak that infects more than one individual but does not spread beyond the single ward would have an expected classification of 11**9(0/2/6)* and all outbreaks with Extent values of 1 have classifications that are consistent with this expectation.

At the extreme, an outbreak that infects all wards for the maximum possible time would have a classification of 999*99(0/1). (This does not occur in the data.)

Having determined that the data examined using the classification system are consistent with predicted outcomes, we now examine its usefulness in describing an outbreak. While the system is of such use, possibly the most useful element in this respect is PkSev. It is reassuring to see that the modal value for this is zero in every case, with a low mean but with an SD of 2–3, meaning that while severe outbreaks have occurred, these are in the minority. Unlike other elements, PkSev may also be used on a daily basis to describe the progress of the infection and the current status within the organization (an example graph is shown in Figure 5). The other elements which may be plotted in this way (and therefore calculated during the outbreak rather than solely at its termination) are Amp (Figure 8) and possibly AvIncidence, although as a daily 'barometer', these are less useful than PkSev, which may be used to trigger ramping up/down of interventional procedures during an outbreak.

One use of the classification system (and the major one to which we have put it) is in the evaluation of stochastic computer models. (Deterministic models, such as that in [15], can reach steady state. Both are valid modelling techniques, but as this classification system requires a completed infection cycle, only stochastic models will be considered hereon.) As the system has several elements, the results from a computer model can be compared for 'goodness-of-fit' (GoF). A model producing results outside of the range seen in the real data is not necessarily incorrect but should be treated with caution. It should be noted that there are restrictions on the type of model that can be compared in this way: the classic susceptible–infected–removed model is not well suited to this classification.
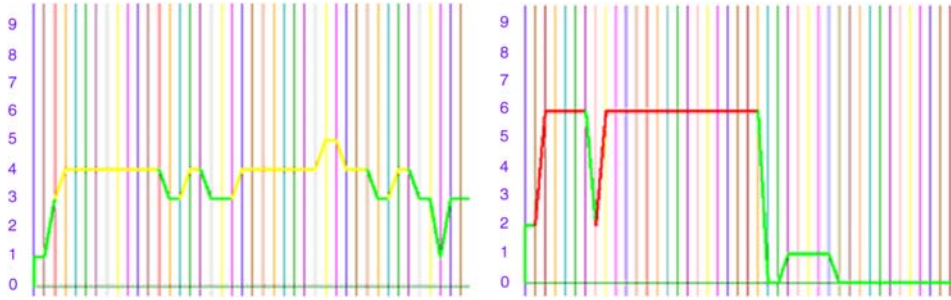
Figure 8.   The number of wards infected at one time (Amp element). The PkSev graph from Figure 5 is reproduced for comparison.

However, a network model [1,2,11] is, and it is with such a model that we have used this system.

Before embarking on this description, it is worth noting that neither the Hamming distance [4] nor the Levenshtein distance [7] is a sufficient tool for this purpose. Both are used by spell checkers for evaluating the 'closeness' or similarity between words in order to suggest replacements. Levenshtein is the more powerful in this respect yet is the less useful here, as it places no importance on the position of the individual characters and thereby handles transposition errors (possibly the most common in typing) admirably – thus 'chioce' can be easily rendered 'choice', yet a comparison of '12344321' and '12434321' would give a low Levenshtein distance, but is actually further, due to the different meaning of the elements. The Hamming distance compares two strings of equal length (Levenshtein has no such restriction) and counts the number of characters that are the same. Having been designed for comparing and validating binary transmissions, it places no weight on closeness of the error – only on whether there is one. Thus '1235' is as far from '1234' as is '1239'.

Four new GoF measures were investigated, the one presented here being an amalgam of two of these. Development used the classifications from the real outbreaks, the realistic classifications derived from modal values, plus 14 additional realistic classifications, drawn from realistic scenarios and 14 additional unrealistic classifications, drawn from unrealistic scenarios.

The measure presented is as follows: the elements are taken as triplets (order matters, so there are 21 comparisons to be made) and the range of possible values of the third for each value of the first two is determined. This then gives three values for each possible value of each element: a minimum, a maximum and a mode. In order to compare a classification for GoF, the scoring shown in Table 6 is applied.

Table 6.   GoF scoring (see text for description).

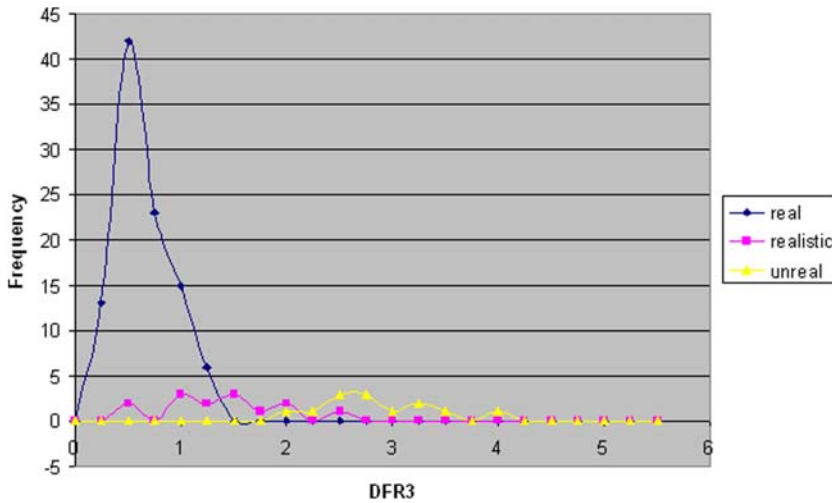|	
Elements 1 and 2 do not match	Ignored
Elements 1 and 2 match and Element 3 matches the mode (NB there may be several modal values)	0
Elements 1 and 2 match and Element 3 is within the min:max range	|value-mode| [in the case of several modal values, the mean of these is used (rounded integer)]
Elements 1 and 2 match and Element 3 is outside the min:max range	10

Figure 9. The frequency plots for the GoF values computed using DFR3 for all classifications (real and simulated), showing an overlap between the ranges of DFR3 values for realistic and unrealistic values but with none for real and unrealistic and thereby a region of very confident 'unrealistic' classifications.

These scores are then summed and normalized by dividing by the number of matches attempted (*i.e.* the number of times elements 1 and 2 match). As there is the possibility of a division by zero (giving a theoretical value of infinity), INT_MAX is assigned in this case (often taken to be 2147483647). The lower the value computed, the better the GoF using this measure. This measure is therefore a 'difference from realistic' measure (DFR3). The results for the data (real and simulated) can be seen in Figure 9.

These results were examined to give three regions as shown in Table 7.

In order to test the use of this demarcation, real classifications that had not been used in the development were tested. Since the project had commenced, a further year of data had been collected and a fresh copy of the database has been obtained. This showed a few differences in the data previously obtained, due to date corrections and the inclusion of data not in the original copy. This enabled six more test cases to be drawn from these data than had been anticipated. The full set of new real classifications was

Staff data: 3 revised and 13 new distinct outbreaks and
Patient data: 4 revised and 13 new distinct outbreaks.

These new results were then analysed using DFR3 and compared to the previous analysis. The results are shown in Figure 10.

Table 7. Demarcation regions for DRF3 (see text for description).

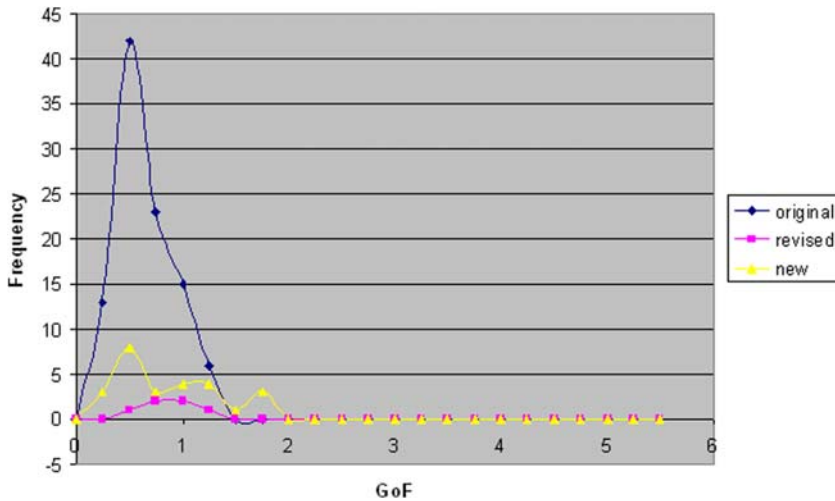| DFR3	Analysis	% Real values in this band	% Realistic values in this band	% Unrealistic values in this band
0–1.8	Realistic	100	79	0
1.8–2.3	Further investigation required	0	21	14
2.3–INT_MAX	Unrealistic	0	0	86

Figure 10.   The frequency plots for the GoF values computed using DFR3 for the original real outbreaks along with the new and revised classifications, showing all revised and new values within the 'real' range.

Table 8.   DFR3 results for new real classifications (see text for description).

DFR3	Analysis	% Original values in this band	% Revised values in this band	% New values in this band
0–1.8	Realistic	100	100	100
1.8–2.3	Further investigation required	0	0	0
2.3–INT_MAX	Unrealistic	0	0	0

The three regions described in Table 7 were applied to give the results in Table 8.

We therefore conclude that DFR3 can be used to determine how well a classification matches to those known to be real. The most practical use of this is likely to be in examining the output from suitable randomly generated computer models, allowing the results to be sifted so that unrealistic results can be quickly discarded and the model thus honed or proved.

The system is, however, of limited use in Trust-to-Trust comparisons for the following reasons:

1. The number of wards will be different. A Trust with, for example, 50 wards will see a single ward infection as affecting 2%, whereas a Trust with 100 wards will see this as 1% – thereby giving different values for Amp.
2. The severity element has been devised using circumstances from the Trust investigated. Other Trusts may be able to cope better or worse with the scenarios used as demarcations and should therefore be adjusted for local conditions.

5.   Conclusion

We have created a classification system that is useful in describing an outbreak of an infectious disease within a semi-closed environment such as a hospital, one element in particular PkSev being of use during the outbreak. Additionally, we have developed a GoF measure to assist in the analysis of the classification produced by computer models.

There are three main uses of this classification system:

1. Outbreaks, once complete, may be compared to previous outbreaks in order to determine how effectively lessons have been learned from previous outbreaks or in the comparison of differing management and containment techniques applied.
2. During the course of an outbreak, two elements can be used to track the outbreak and trigger action: Amp (number of wards the infection is present in) and PkSev.
3. Simulated outbreaks produced by computer models may be compared to known real values via the DFR3 GoF measure in order to determine how realistic (or otherwise) the simulation is.

Additionally, there may be a further use of the DFR3 value in that it may give a simple metric for describing the overall outbreak, analogous to the Richter scale, but this has not been investigated. It is likely, though, that the PkSev value alone would suffice locally.

Notes

1. Email: maurice.madeo@hey.nhs.uk
2. Email: r.phillips@hull.ac.uk

References

[1] L. Ancel Meyers, M.E.J. Newman, and B. Pourbohloul, *Predicting epidemics on directed contact networks*, J. Theor. Biol. 240(3) (2006), pp. 400–418.
[2] K.T.D. Eames, *Modelling disease spread through random and regular contacts in clustered populations*, Theor. Popul. Biol. 73(1) (2008), pp. 104–111.
[3] P.S. Ganney, *A framework for investigating the spread of infection in a hospital environment utilising a stochastic and small-world model*, Proceedings of the Annual Scientific Meeting of the Institute of Physics and Engineering in Medicine, 2003, p. 105.
[4] R.W. Hamming, *Error detecting and error correcting codes*, Bell Syst. Tech. J. 29(2) (1955), pp. 147–160.
[5] J. Hawker, N. Begg, I. Blair, R. Reintjes, and J. Weinberg, *Communicable Disease Control Handbook*, Blackwell, Oxford, 2005.
[6] J.M. Last, *A Dictionary of Epidemiology*, Oxford University Press, New York, 2008.
[7] V.I. Levenshtein, *Binary codes capable of correcting deletions, insertions, and reversals*, Soviet. Phys. Dokl. 10 (1966), pp. 707–710.
[8] J.S. Li, D.J. Sexton, N. Mick, R. Nettles, V.G. Fowler, Jr., T. Ryan, T. Bashore, and G.R. Corey, *Proposed modifications to the Duke criteria for the diagnosis of infective endocarditis*, Clin. Infect. Dis. 30 (2000), pp. 633–638.
[9] B.B. Mandelbrot, *Self-affine fractals and fractal dimension*, Phys. Scr. 32 (1985), pp. 257–260.
[10] C.G. Mayhall, *Hospital Epidemiology and Infection Control*, 3rd ed., Lippincott Williams and Wilkins, New York, 2004.
[11] M.E.J. Newman, *Models of the small world a review*, J. Stat. Phys. 101 (2000), pp. 819–841, Available at http://arxiv.org/abs/cond-mat/0001118.
[12] C.J. Noakes, C.B. Beggs, P.A. Sleigh, and K.G. Kerr, *Modelling the transmission of airborne infections in enclosed spaces*, Epidemiol. Infect. 134(5) (2006), pp. 1082–1091.
[13] C.F. Richter, *An instrumental earthquake magnitude scale*, Bull. Seismol. Soc. Am. 25(1) (1935), pp. 1–32.
[14] Z. Toroczkaia and H. Guclu, *Proximity networks and epidemics*, Phys. A: Stat. Mech. Appl. 378(1) (2007), pp. 68–75.
[15] J. Vanderpas, J. Louis, M. Reynders, G. Mascart, and O. Vandenberg, *Mathematical model for the control of nosocomial norovirus*, J. Hosp. Infect. 71 (2009), pp. 214–222.