## Note re: Day of year analysis

DRJ 110303

1. Considering the restricted wards data, initially for $\mathbf{1 9 9 8}$ only: Dr 2 had total of 102 events, Dr 150 . However, there is no 'time at risk' information. The analyses below are based on comparing rates for Dr1 during apparent periods of absence of Dr2 with rates for Dr2 when (apparently) present. Absence of Dr2 is inferred from prolonged gaps between consecutive Dr2 events. However, such gaps could of course occur by chance even when Dr2 is present.
2. Such errors of inference are less likely for longer than shorter gaps (as the chance of event-free periods decreases as the length of period increases). Other errors in assumptions made in the calculation are also less severe in the case of longer periods. The principal results below are based on event-free periods of at least 14 consecutive days for Dr2; those for shorter periods are more prone to error.
3. There are periods of apparent absence of Dr2 in 1998, each of 17 days. The number observed from Drl in these periods is 12. Assuming the 102 events are thus truly ascribable to the remaining 366-17-17=332 days, the number expected (E) from Dr2 during the 332 day period, if Dr1's rate applies is $(12 / 34) \times 332=117.2$, to be compared with the observed (O) number of events 102. There is thus no evidence from this analysis of an excess in Dr2 events $\{\mathbf{O} / \mathbf{E}$ $=0.87$ ( $\mathbf{9 5 \% c i} \mathbf{0 . 4 8 - 1 . 7 4 \}}$
4. As noted in 2. above, use of shorter periods may introduce errors of inference about absence, and more important 'end errors' in calculations. However, proceeding in the same way but with a 10 -day (cf 14-day) definition of absence yields: O for $\mathrm{Dr} 2=102 ; \mathrm{E}=(26 / 104) \times 262=65.5 .\{\mathrm{O} / \mathrm{E}=1.56(1.00-2.50)\}$
5. For a 7 -day period (where errors may be a major problem) O for $\mathrm{Dr} 2=102, \mathrm{E}=$ $(32 / 155) \times 211=43.6 .\{\mathrm{O} / \mathrm{E}=2.34(1.56-3.60)\}$
6. Thus overall (for 1998) there is no clear evidence of a raised rate in Dr2, since the a priori more reliable calculations do not support this conclusion, and the nominally significant less secure calculations need to be treated with considerable caution.
7. Extending the 14-day analyses for the later period of 1999 and first half of $\mathbf{2 0 0 0}, \mathrm{O}$ for $\mathrm{Dr} 2=125, \mathrm{E}=(16 / 85) \times 456=85.8\{\mathbf{O} / \mathbf{E}=\mathbf{1 . 4 5}(\mathbf{0 . 8 6 - 2 . 6 2})\}$
8. Combining the $14+$ day analyses for 1998,1999 and part of 2000 (ie those in paras 3 and 7 above) yields $\mathrm{O}=227, \mathrm{E}=185.4\{\mathbf{O} / \mathbf{E}=\mathbf{1 . 2 2}(\mathbf{0 . 8 2 - 1 . 8 8})\}$
9. Overall, although the most reliable (14-day + ) analyses may suggest that Dr2's rate is somewhat raised, perhaps especially from 1999 onwards, this is not a statistically clearcut result.
