The new automated daily mortality surveillance system in Portugal

The experience reported in an earlier Eurosurveillance issue on a fast method to evaluate the impact of the 2003 heatwave on mortality in Portugal, generated a daily mortality surveillance system (VDM) that has been operating ever since jointly with the Portuguese Heat Health Watch Warning System. This work describes the VDM system and how it evolved to become an automated system operating year-round, and shows briefly its potential using mortality data from January 2006 to June 2009 collected by the system itself. The new system has important advantages such as: rapid information acquisition, completeness (the entire population is included), lightness (very little information is exchanged, date of death, age, sex, place of death registration). It allows rapid detection of impacts (within five days) and allows a quick preliminary quantification of impacts that usually took several years to be done. These characteristics make this system a powerful tool for public health action. The VDM system also represents an example of inter-institutional cooperation, bringing together organisations from two different ministries, Health and Justice, aiming at improving knowledge about the mortality in the population.

Introduction
The ongoing surveillance of mortality to detect and estimate the magnitude of deaths caused by epidemics, emergence of new diseases, or other relevant public health events or threats is envisaged as an important tool. Current efforts are in place to develop such a system at the European level [1].

In Portugal, heatwaves have been demonstrated to be an important health problem [2] and since 1999, a Portuguese Heat Health Watch Warning system (HHWW) has been operating, known as the ICARO Surveillance System (the acronym stands for Importância do Calor: Repercussão nos Óbitos — importance of heat and its repercussion on mortality) [3,4]. This ICARO system was based on models that forecast increases in mortality, and once this evidence was obtained, it was communicated to system partners in the form of a simple index called the ICARO-Index [3,5], where 0 (zero) means absence of risk and positive values added risk of mortality related with heat. Basically this index accounted for maximum temperatures above a fixed threshold of 32°C and the number of consecutive days on which such occurrences were observed. The Lisbon ICARO-index was updated later to integrate the experience of the 2003 heatwave [4], and it is currently referred to as ICARO-Index 2005.

In the time following the first ICARO alert in 2000, public health mitigation actions tended to wait for confirmation of an impact on the population (e.g. excess mortality), and once this evidence was obtained, it was too late to act. Given that there was no rapid method available to obtain that evidence, a tool that could provide it was needed.

It was within this framework that a first mortality monitoring system was tested in the hot summer of 2003. This early version (hereafter referred as ad-hoc VDM: Vigilância Diária da Mortalidade – Surveillance of Daily Mortality) was based on the daily number of death registrations in a set of Civil Registrar Offices (CROs) that accounted for about 40% of all Portuguese mortality [6]. In fact, this early version allowed the confirmation of the excess mortality predicted by the ICARO Surveillance System in the summer 2003 [3,7].

Generation of a routine daily mortality monitoring system – the original VDM system (2004-2007)
The ad-hoc system tested during the heatwave of 2003 used a sample of 31 CROs representing all capital cities of all districts of mainland Portugal [2]. In the summer of 2004, the routine daily surveillance was launched (original VDM system). This version differed from the ad-hoc system in the set of CROs, which consisted now


This article has been published on 1 April 2010
of a sample of 67 CROs, the first 31 CROs plus a random sample. This original VDM system was operational from 2004 to 2007, and the data collected consisted only of the numbers of registered deaths in each CRO by date of registration (total and for individuals aged 75 years or older). Data was sent to the Portuguese National Health Institute (INSA) at the end of every day or on the following morning. Data was mainly transmitted by fax and telephone with only a few offices being able to exchange data by e-mail. This non-automated version of the system required six or seven persons at INSA to collect the basic data. Given the original interdependence of this VDM system with the ICARO system, mortality information was disseminated within the daily ICARO bulletin to its partners which consisted of National Health and Civil Protection authorities (Figure 1).

Although logistically complex, this original VDM System was functional and fully executed on a daily basis (weekdays) during the summers, and at the end of the summer of 2005 it was decided to extend the data collection period to the full year. Figure 2 shows the evolution of the original VDM system. For the year 2004, the deaths registered by the 67 CROs are compared to weekday baselines of expected deaths based on information supplied by the 31 CROs involved in 2003, omitting weeks with national holidays and heatwaves. For better comparison with the baseline, the graph for the year 2004 also shows the deaths registered in 2004 by the 2003 sample of 31 CROs.

The baseline of expected deaths in the graphic for 2005 was based on information from the sample of 67 CROs in 2004, the same 67 CROs that also supplied the information on actual deaths in 2005. It is curious that the addition of 36 relatively less important CROs (with regards to the registrations of the number of deaths) led to a slightly different pattern, especially on Fridays. It is also remarkable that holidays, either national or local (in important centres like Lisbon and Porto, e.g. St. John on 24 June) caused delays in the acquisition of mortality information. Noteworthy is that a strike of the justice services in 2005 also affected the system information.

In 2006, an added feature was tested by including a mathematical model that would consider delayed information due to a holiday. This allowed an assessment of whether the observed number of death registrations

---

**Figure 1**

Original daily mortality surveillance system (VDM), Portugal, 2004-2007

---

CRO: Civil Registrar Offices; DGS: General Directorate of Health; INSA: National Health Institute Dr. Ricardo Jorge; IRN: Institute for Registries and Notary; ONSA: National Health Observatory (currently Epidemiology Department of INSA).
on a given day was the expected. Although this model did not have all the required statistical properties it seemed to accurately predict the real situation.

Another feature was tested: adding to the expected number of deaths those predicted by the ICARO model. Results demonstrated that it was possible to model the delay in registered mortality including excess mortality predicted by the ICARO model. But at the time the ICARO surveillance system relied on an extrapolated index for Lisbon and not on the current national weighted index (which relies on indexes for four regions of Portugal weighted according to resident population).

Several estimates of excess deaths due to heat episodes have been made using the mortality surveillance system (original VDM). For example, heatwave data analysis for 2003 was done comparing several reference periods of mortality data with the observed data in that particular heatwave period, deriving the respective confidence intervals and calculating exact Poisson probabilities to obtain p-values [3].

Although this mortality system was in place and functional, the detection and calculation of excess mortality due to an event was not straightforward. The main limitation of this system was the use of daily number of death registrations instead of daily number of deaths. The use of the former is heavily dependent on the weekday, with mortality on weekends and holidays registered later (Figure 2). Consequently, it was only possible after the experience of three summers (2003-2005) to establish a rationale that allowed modelling the expected delayed mortality registration.
The new VDM system (in place since 2007)
The new VDM system was only possible because of a technical change in 2006 within the Ministry of Justice through a system called SIRIC (Integrated system for civil registries and identification) that aimed at connecting all civil CROs and collecting their data in a centralised way. SIRIC included information from the Institute for Registries and Notary (IRN), which is responsible for the CROs and was implemented and maintained by the Institute for Information Technologies of Justice (ITIJ).

Although SIRIC included only half of the existing CROs (those that were already computerised), IRN asked INSA in mid-2006 to test an automated version of the VDM system using exchange between INSA and ITIJ of data collected centrally by SIRIC. For this test, only of few variables (date of death, age, gender, and geographical code for location of death registration) were circulated by e-mail. The experience was satisfactory and the automated system was deemed feasible and useful as it would save resources at both ends.

The two systems, old and new, coexisted from September 2006 to May 2007 when the original VDM system was discontinued. By June 2007, all Portuguese CROs were reporting to SIRIC. Data flow consisted of a single e-mail, containing information for the previous day, sent by ITIJ on a daily basis, including weekend days and holydays (Figure 3). With this paper we aim to demonstrate that the new VDM system, fully automated since early 2008, allows quick detection of events and rapid estimates of their impact using reduced resources.
Methods
Definition of events
Influenza activity
The definition of influenza epidemic periods relied on information of influenza activity that consisted on weekly estimates of influenza-like illness (ILI) incidence rates obtained by the Portuguese General Practitioner (GP) Sentinel Network (Rede Médicos-Sentinela) [8]. The study period comprised the seasons 2006-7, 2007-8 and 2008-9. The epidemic periods were defined as the set of consecutive weeks with estimates of ILI incidence rates above the 95% upper confidence limit of the baseline levels [9] (see Table 1).

Heat periods
For the determination of the heat periods, data of observed maximum temperatures in a given district were considered. Temperature data were made available by the Meteorological Institute in Portugal. Heat periods, summarised in Table 2, were defined according to one of two criteria:

1. Two or more consecutive days with temperature above 35°C in two (of 18) districts (or temperature above 36°C in one district)
2. Two or more consecutive days with ICARO index above zero.

Mortality
Mortality data were generated by the new VDM system from January 2006 to June 2009. Depending on the event studied, data was aggregated on a daily (heat event) or weekly (influenza activity event) basis.
Data was stratified by gender, age group (65-74 and ≥75 years) and by Nomenclature of Territorial Units for Statistics (NUTS) 2 regions (North, Centre, Lisbon and Tagus Valley, Alentejo and Algarve).

**Determination of VDM system delay**
To determine the VDM system delay, daily data by date of registration, date of death and the respective number of days of delay were considered. The period from 5 June 2007 (1st day with all CROs included in the system) to 4 June 2009 was covered. Analysis by weekday and month of death was performed using relative percentages for each weekday or month.

**Statistical methods**
For confirmation and detection of impact, cyclical regression models [10] were fitted to complete mortality data (January 2007 to June 2009) using the Flubase package [11]. For the confirmation approach, models were fitted to data excluding the event periods (Tables 1 and 2), and for event detection, all data was used without considering known event periods.

The mortality predicted by the model was considered as the baseline mortality. A confirmed excess mortality period was defined as a set of consecutive days or weeks (depending on data level being used) that began with two values of observed number of deaths above the upper 95% confidence limit of the baseline and ended with two consecutive mortality values below this limit. For heat periods impacts, a single day or week of mortality above the 95% confidence interval, within a marked heat event period, was also regarded as a confirmed excess mortality.

**Results**

**New VDM system overview**
Figure 4 shows mortality series, identifying the influenza epidemics and heat periods as well as the evolution of the ICARO system. The evolution of ILI incidence rate per 100,000 inhabitants is presented elsewhere [10]. From January 2006 to June 2007, several CROs were added to the SIRIC system, resulting in several steps visible in Figure 4 as abrupt increases in the number of reported deaths. Mortality peaks identi-
fied by the system coincide with important events of heatwaves and influenza activity.

The information collected by the new VDM system, by sex, age group and region of death registration is summarised in table 3 below.

**VDM system delay**

In order to understand the potential and applicability of the generated data, it was important to determine the system delay (time in days from death until inclusion in the VDM system). Overall, 14.1% of deaths were included on the day of occurrence, 58.0% within one day, 77.9% within two days, 94.2 within three to four days, and 98.0% were included within seven days of occurrence.

While each weekday had a specific information delay (Figure 5), the different patterns seemed to be the simple consequence of the weekend when only a reduced number of CROs are available. The greatest proportion of mortality included on the day of occurrence was observed on Fridays. Fridays were also the weekdays on which it took longer to include all information in the system, with only 87.3% of complete information available within four days.

**Confirmation and detection of an event’s impact**

**Detection**

Figures 6 and 7 show the results of a cyclical model fitting total data, stratified at various levels (sex, age group and two major Portuguese regions) both daily and weekly. It is noteworthy that the main known events (influenza activity and heat) were identified by the system as expected.

The influenza periods 2006-7 and 2008-9 and the heat periods of 2007 and 2008 were identified using both daily and weekly data. However, daily data seemed to generate better evidence of the occurrence of these events than the VDM data aggregated by week.

There was a consistent lack of evidence for any impact on mortality of the observed influenza activity during the season of 2007-8. This was expected, since the estimated incidence rates were low and within the sentinel influenza surveillance 95% confidence interval limits. The consistent increase of mortality identified in the daily data series (Figure 6) at the end of year of 2007 did not correspond to any known event.

**Confirmation**

Figures 8 and 9 demonstrate the impact on mortality of known events, namely the influenza seasons 2006-7 and 2008-9, the heat periods of 2007 and 2008, and provide some evidence of an impact on mortality of the heat period observed in the summer of 2009.

Both approaches (daily and weekly) were suitable for the confirmation of events. The weekly aggregated data confirmed an impact of influenza activity in 2006-7 and 2008-9 and the heat periods of 2007 and 2008. The daily data on the other hand showed less excess mortality related to the 2008 heat period and better evidence of impact of the 2009 heat period. Also here, both approaches did not reveal any impact of the influenza activity of 2007-8 and showed increased mortality at the end of 2007.

Figure 8. Cyclical regression on daily mortality data generated by the daily mortality surveillance system (VDM), with defined event periods (respective data omitted in model fitting), by total mortality, sex, age group and two major regions in Portugal, January 2007-June 2009

**Discussion and conclusion**

Our results show the advantages of the VDM system: lightness, timeliness, rapidity and completeness. By lightness we mean that very little information on a registered death is exchanged, consisting only of date of death and registration, sex, age and region of death registration. The residence of the individual was not considered in order to enforce confidentiality of data.
The system obtains information very quickly and in a very complete way. It collects 94.2% of the total number of deaths within four days of their occurrence and 98% within seven days; it allows detecting impacts on mortality in a very timely manner; and it seems to be sensitive to phenomena that are generically considered for the construction of such mortality surveillance systems, i.e. it is sensitive to extreme weather events and to winter influenza activity [7,10]. In fact, in a year with moderate summer temperatures as in 2007, the VDM system data without additional information could point out even some of the events with low impact in mortality within a framework of event detection (Figures 6 and 7).

We showed that the reporting delay was dependent on the weekday, with deaths occurring on Fridays entered into the VDM system with the greatest delay (three or more days), although mortality registration on the same day of death occurrence was highest on Fridays. This pattern might indicate that mortality monitoring is least influenced by delay on Thursdays, when most of the weekend mortality is already integrated. This would probably be the optimal day for monitoring of mortality data from the preceding week.

Mortality data delay also showed some dependence on the month of registration (data not shown). January to March had the minimum delay with at least 14.7% of the data registered on the day of the mortality event and more than 80% within two days. Greater delays occurred in April, August and December, which was expected because these months are associated with holiday periods. The minimum mortality registered on the day of death occurrence was 12.2% in December, followed by the summer months with about 13%. In months associated with holidays (April, August and December), arrival of data within two days was always below 75%, with December reaching only 71.5%.

The system delays related to weekdays and months are basically a reflection of society’s organisation of time. This time structure remains unaltered through the years, therefore detection and determination of event impacts should be done using within-system data for the definition of baselines and thresholds. Overall, the identified system delays are not expected to dramatically change monitoring or surveillance schemes using VDM system information. Empirically, past experience of VDM system management, both in the old and the new version, showed that residual impacts on

**Figure 4**

Daily observed mortality data collected by the new daily mortality surveillance system (VDM) and ICARO Index, Portugal, 2006-2009

Blue line: new VDM system; grey line: ICARO Index; light blue bars: heat periods; medium blue bars: influenza epidemic periods.
mortality (like those of the moderate heat periods of 2007 and 2008) are usually only identifiable about four or five days after the event.

It seems widely accepted to study winter/influenza activity using weekly aggregated mortality data and heatwaves of moderate proportion using daily mortality data. For heatwaves, the daily level is very important in order to quickly confirm the impact and initiate mitigating measures. Our results, albeit contradictory, indicated that daily data may be better suited to confirm and detect events in general. The contradiction lay in the fact that best confirmation of an event, both influenza activity and moderate or mild heat periods, was obtained from daily data, while these daily data were of limited use for the purpose of detecting heatwaves. This latter limitation may be due to the fact that the system so far only covers a short time series of 2.5 years. It is reasonable to expect that as the VDM time series gets longer, its ability to model cyclical patterns, detect and confirm relevant events will improve.

The apparently unexplainable increase in mortality at end of the year 2007 may in fact have been a result of the short length of the VDM data series at the time. The increased mortality may reflect a cyclical pattern that has not yet been picked up because many other events occurred in the relatively short studied period.

The current status of the new VDM system allows us to look forward convinced that it can be used in a timely and useful manner no matter what public health problems arise (provided they have a detectable impact on mortality). But there were challenges during the development of the current system: The first version of the automated VDM system was not fully automatic and relied on human intervention. With the increase of data and work, it had to be improved and the used XML data structure had to be changed. That posed problems because the simple solutions previously adopted and implemented in Microsoft Excel were not able to deal with the new data structure. The problem could not be solved by the use of statistical packages. Basically, the system received partial records of different tables, relating to the same individual but not connected. A solution was found using an open source database management system called MySQL. Although the solution was not complex, it took more than six months to have the system completely operational and re-establish of the data flow.

In early 2009, when trying to check if the number of records for full years was correct, about 5% of data was found to be missing. The explanation for this was that only records with registration and confirmation on the same day were being sent to VDM. Records registered on one day, but with administrative confirmation on a different day were not sent. What must be stressed is that both parties intended to exchange all the information, but the implementation was done in a different way. This was easily corrected and now all is working correctly. Such quality control must be implemented in early phases of these systems.

The construction of an automated system for mortality surveillance requires strong inter-institutional will and awareness of the potential downsfalls to ensure that such a system is fully operational all year round. This seems to be more important than financial resources, seeing as the successive versions of the Portuguese

---

**Table 3**

Number of deaths registered in the daily mortality surveillance system (VDM) by sex, age group and region of death registration, Portugal, January 2006-June 2009

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td>264,427</td>
<td>100%</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>135,832</td>
<td>51,4%</td>
</tr>
<tr>
<td>Female</td>
<td>128,475</td>
<td>48,6%</td>
</tr>
<tr>
<td>Unknown</td>
<td>120</td>
<td>0,045%</td>
</tr>
<tr>
<td><strong>Age group (years)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 - 14</td>
<td>1,462</td>
<td>0,6%</td>
</tr>
<tr>
<td>15 - 24</td>
<td>1,366</td>
<td>0,5%</td>
</tr>
<tr>
<td>24 - 44</td>
<td>9,393</td>
<td>3,6%</td>
</tr>
<tr>
<td>45 - 64</td>
<td>34,991</td>
<td>13,2%</td>
</tr>
<tr>
<td>65 - 74</td>
<td>43,896</td>
<td>16,6%</td>
</tr>
<tr>
<td>75+</td>
<td>169,717</td>
<td>64,2%</td>
</tr>
<tr>
<td>Unknown</td>
<td>3,602</td>
<td>1,4%</td>
</tr>
<tr>
<td><strong>Region (of death registration)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norte</td>
<td>76,880</td>
<td>29,1%</td>
</tr>
<tr>
<td>Centro</td>
<td>65,017</td>
<td>24,6%</td>
</tr>
<tr>
<td>Lisboa</td>
<td>69,075</td>
<td>26,1%</td>
</tr>
<tr>
<td>Alentejo</td>
<td>22,441</td>
<td>8,5%</td>
</tr>
<tr>
<td>Algarve</td>
<td>12,022</td>
<td>4,5%</td>
</tr>
<tr>
<td>Açores</td>
<td>5,814</td>
<td>2,2%</td>
</tr>
<tr>
<td>Madeira</td>
<td>6,546</td>
<td>2,5%</td>
</tr>
<tr>
<td>Other/foreign</td>
<td>6,632</td>
<td>2,5%</td>
</tr>
</tbody>
</table>

**Figure 5**

Delay of data inclusion in the daily mortality surveillance system (VDM) from date of death occurrence to registration, by weekdays, Portugal, 5 June 2007 to 4 June 2009
Figure 6
Cyclical regression of daily mortality data generated by the daily mortality surveillance system (VDM) without definition of event periods (for detection purposes) by total mortality, sex, age group and two major regions in Portugal, January 2007-June 2009.

Black line: mortality baseline; blue line: observed mortality; grey line: baseline 95% confidence interval upper limit. In the x axis, .0 stands for beginning of the year, .5 for half year.
Figure 7
Cyclical regression on data generated by the daily mortality surveillance system (VDM), aggregated by week, without defined events periods (for detection purposes), by total mortality, sex, age group and two major regions in Portugal, January 2007-June 2009

Black line: mortality baseline; blue line: observed mortality; grey line: baseline 95% confidence interval upper limit. In the x axis, .0 stands for beginning of the year, .5 for half year.
Figure 8

Cyclical regression on daily mortality data generated by the daily mortality surveillance system (VDM), with defined event periods (respective data omitted in model fitting), by total mortality, sex, age group and two major regions in Portugal, January 2007-June 2009.

Black line: baseline as defined by cyclical regression representing expected mortality; grey line: 95% Confidence interval upper limit. Light blue bars: known events periods; medium blue bars: confirmed excess mortality periods (two or more consecutive days). In the x axis, .0 stands for beginning of the year, .5 for half year.
Figure 9
Cyclical regression on data generated by the daily mortality surveillance system (VDM), aggregated by week, with defined events periods (respective data omitted in model fitting), by total mortality, sex, age group and two major regions in Portugal, January 2007-June 2009

Black line: baseline as defined by cyclical regression representing expected mortality; grey line: 95% Confidence interval upper limit. Light blue bars: known events periods; medium blue bars: confirmed excess mortality periods (two or more consecutive weeks). In the x axis, .0 stands for beginning of the year, .5 for half year.
VDM system were developed without any specific budget using only the available manpower and computer technology. It should be possible to establish a timely mortality surveillance system such as the VDM at European level. It would probably need goodwill and coordination between institutions within and among the countries rather than resources, and faces potential technical pitfalls that would need to be solved.

Currently, the new version of the VDM system has been separated from the ICARO surveillance system, having its own bulletin that is sent on every weekday to national and regional health authorities and also to the Portuguese civil protection authorities. Some products based on this system are being studied for future application. A system for surveillance of infant mortality is currently being tested with promising results.

References