EUROCAT Statistical Monitoring Protocol 2008

(Uploaded to EUROCAT website April 2010)

EUROCAT Central Registry
University of Ulster
Newtownabbey, Co Antrim
Northern Ireland, BT37 0QB

Tel: +44 28 9036 6639
Fax: +44 28 9036 8341
Email: eurocat@ulster.ac.uk
Web: www.eurocat-netork.eu

EUROCAT receives funding from the European Union, in the framework of the Public Health Programme
WHO Collaborating Centre for the Epidemiology Surveillance of Congenital Anomalies

EUROCAT PROTOCOL FOR STATISTICAL MONITORING AND INVESTIGATION OF TRENDS AND CLUSTERS

CONTENTS

Acknowledgements:

Alan Kelly and Conor Teljeur (Trinity College Dublin) adapted and implemented the scan technique for statistical monitoring, in discussion with Helen Dolk (EUROCAT Project Leader, University of Ulster) with first results presented to the Registry Leader's Meeting (RLM) in Heidelberg 2003. Subsequently, the statistical code was incorporated into the EUROCAT Data Management Programme (EDMP) by James Densem (Biomedical Computing Ltd). Statistical monitoring was co-ordinated by Araceli Busby (LSHTM) until 2005, and by Maria Loane (Central Registry, UU) since then. In 2006, data analysis using the standard software was transferred from Dublin to Central Registry. This protocol has been written by all those involved.

1. Background: monitoring strategy and timetable

Statistical Monitoring relates to two of EUROCAT's objectives:

- to provide essential epidemiologic information on congenital anomalies in Europe
- to co-ordinate the detection of, and response, to clusters and early warning of teratogenic exposures

EUROCAT was established in 1979 in the wake of the thalidomide epidemic, and statistical monitoring still aims at early detection of any new teratogenic drug. Since then, interest has also widened to other environmental chemicals as potential teratogens (see EUROCAT Special Report: a Review of Environmental Risk Factors for Congenital Anomaly 2004 http://www.eurocat-network.eu/content/Special-Report-Env-Risk-I-and-II.pdf and http://www.eurocat-network.eu/content/Special-Report-Env-Risk-III.pdf). Statistical monitoring is only one part of surveillance of teratogenic exposures, to identify potential cause for concern where there is no specific prior hypothesis about the exposure. It is essentially a screening method to scrutinise data regularly and systematically, to detect any previously unrecognised increases in frequency. Where there are specified hypotheses about new teratogens (e.g. which drug, when, where), other direct approaches for analysis should be undertaken which are not the subject of this protocol.

The EUROCAT Coding and Classification Committee select EUROCAT congenital anomaly subgroups for statistical monitoring. All subgroups are defined in EUROCAT Guide 1.3 http://www.eurocat-network.eu/content/EUROCAT-Guide-1.3.pdf.

Currently, statistical monitoring is conducted to detect changes in time within each registry and to detect trends across all registries. Future developments will refine data analysis across registries and incorporate the spatial dimension.

Statistical methods have been chosen which are relatively straightforward for public health authorities to understand and communicate, and which can be supplied in the EUROCAT software (EDMP) to member registries for use by non-statisticians.

The elements of the current monitoring strategy are:

- Common user-friendly statistical software for use centrally and locally (EDMP)
- Annual statistical monitoring for trends and clusters at central level, 15 months after last date of birth (e.g. year 2008 births included in monitoring in March 2010).
- More frequent and/or earlier statistical monitoring locally (by member registries)
- Use of EUROCAT communications and special data analyses to respond to news about clusters identified locally or outside the monitoring system
- A clear and prompt system of investigation and reporting of results
- Use of statistical monitoring additionally as a data quality control system

All full member registries able to meet the annual February data transmission deadline (e.g. Feb 2010 for 2008 births) participate in cluster detection monitoring. All full member registries that have data no more than one year behind (e.g. 2007 complete in Feb 2010) are included in trend analysis. Associate members can also request to be included in trend analysis. Associate member registries are invited to use EDMP locally for cluster detection and report their results to Central Registry.

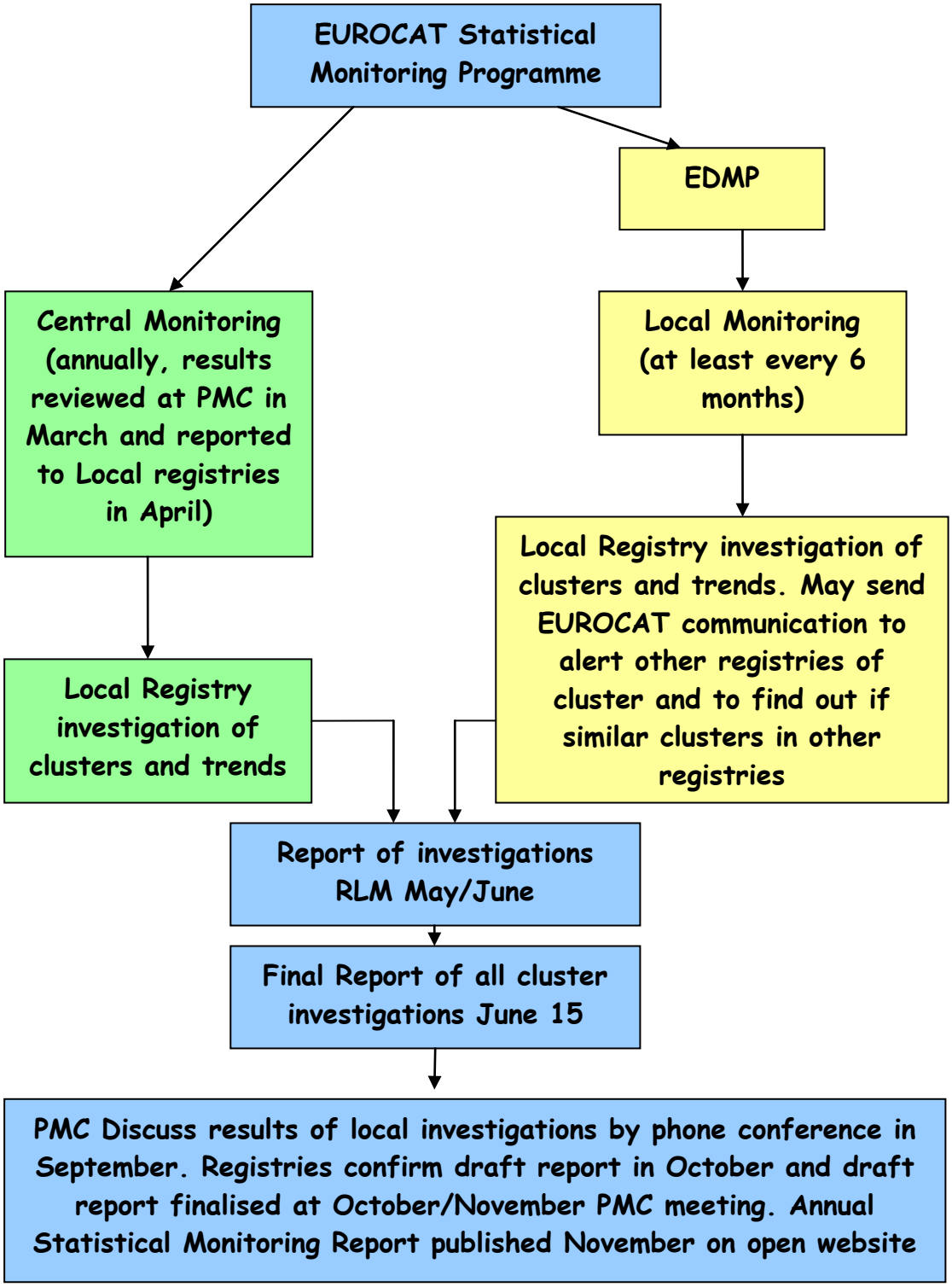


Figure 1: Monitoring and reporting centrally and locally.

2. Statistical Methods

2.1 Statistical methods for the detection of trends

Trend tests are performed for each anomaly subgroup (see Appendix 3) for each registry. Currently, Central Registry performs a trend test for the most recent five years of data, based on the number of cases per year, as well as a trend test for the most recent 10 years (or 8 years if 10 years are not available), grouping each two year interval. A trend test is not performed if the expected number of cases per year (or two year interval) is less than 5 or if the observed number of cases per in any one year (2 year interval) is less than 2.

Change over time is tested with a chi square test for heterogeneity, divided into the trend component ("chi square test for trend") and the non-linear component ("chi square test for non-linear change"). The analysis is based on the number of cases per year of birth and the number of births per year. The Chi square test for trend identifies evidence of an increasing or decreasing trend in prevalence. The significance level (p-value) and direction (upward or downward) and slope of the trend are given in the output. The Chi square test for non-linear change identifies significant change over time (i.e. the prevalence changes from year to year), that does not show an increasing or decreasing trend. The p-value is given in the output.

Where $p<0.05$ for trend component and $p>0.01$ for non-linear component, the results are identified as an "increasing or decreasing trend". Since overall directional trend is of most concern for investigation, a chi square for trend p-value less than 0.05 is interpreted as a trend even where the p-value for non-linear change is weakly significant also (between 0.05 and 0.01).
Where $p<0.05$ for trend component and $p<0.01$ for non-linear component, the results are identified as "non-linear change".
Where $p>0.05$ for trend component and $p<0.05$ for non-linear component, the results are identified as "non-linear change".
Where $p>0.05$ for trend component and $p>0.05$ for non-linear component, the results are interpreted as showing no significant change over time.

Since the Chi squared test is based on conventional probabilistic statistics, at a significance level of $p<0.05$, 5% of the test results will be statistically significant by chance. This should be kept in mind in interpretation (see protocol for investigation).

2.2 Adjusting Down syndrome trends for in-utero survival and maternal age

For each registry, the observed number of terminations of pregnancy for Down syndrome (DS) at less than 20 weeks gestation are corrected for the probability of survival to 20 weeks gestation. The denominators are adjusted for change in maternal age distribution over time. Both unadjusted and adjusted results are given. Details of the adjustment methodology are given in Appendix 1.

2.3. "Pan-Europe" trend detection.

The "Pan-Europe" analysis repeats the procedures above, but adds up data across all eligible registries. This is particularly useful for very rare anomalies which frequently have too few cases per year/2-year interval in individual registries for chi square analysis. The "Pan-

Europe" analysis is also useful for an indication of the overall trend across Europe. More sophisticated techniques for meta-analysis across registries are under development.

2.4. Statistical methods for the detection of clusters

EUROCAT defines a cluster as:

> "An aggregation of cases of congenital anomaly in time and/or space which appears to be unusual". [1]

This definition includes space as defined by a common activity such as a place of work/ education/ recreation etc. and not just space as defined by residence.

Currently, central statistical monitoring detects temporal clusters within each registry area (region or nation). The EDMP software allows registries locally to detect temporal clusters within subareas of the area covered by the registry.

Cluster detection is based on a moving window test described by Naus and Nagarwalla[2] (see Appendix 2 for statistical details). The method uses a moving window of a given number of cases (window size), measuring the length of time between the first and last case. The method detects whether the given number of cases has occurred in a shorter time than would be expected by chance. The method is not robust with a window size of less than 5 cases. A minimum of 7 cases over the study period of interest is needed to run the analysis. All window sizes from a minimum of 5 to a maximum of the total number of cases minus 2 are tested.

Each registry and anomaly subgroup is tested independently and the analysis takes into account the number of tests with different window sizes being performed within each registry/anomaly combination to allocate a likelihood to each cluster. The method gives each cluster identified a scan statistic called "lambda", from which a p-value for significance is derived (see Appendix 2).

Many clusters may overlap in time, the inclusion or exclusion of individual cases changing their significance. All significant clusters are identified. The "most significant" cluster (lowest p-value) is then identified. All other significant clusters ($p<0.05$) for which at least 75% of cases overlap with the "most significant" cluster are considered to belong to the same cluster group. The method then looks for the second most significant cluster not already allocated to a cluster group and proceeds similarly identifying the second cluster group. The output identifies the most significant cluster in each cluster group, and the first and last case in time belonging to any cluster within the cluster group.

Following identification of all cluster groups over the time period of analysis, cluster groups overlapping with the last two years of data, and in which the most significant cluster is less than 18 months in length, are chosen for output and investigation. An option is also provided

---

[1] EUROCAT Working Group on the Management of Clusters and Environmental Exposure Incidents, 2003

[2] Naus, J.I. (1965). The distribution of the size of the maximum cluster of points on a line. Journal of the American Statistical Association 60, 532-538

Nagarwalla, N. (1996). A scan statistic with a variable window. Statistics in Medicine 15, 845-850

in the software for Central Registry to scan the most significant cluster in all years of data scanned, rather than just the last two years, for research purposes. Similarly, options are available to reset the percentage acceptable for missing gestational age, and cluster lengths greater than 18 months.

Currently Central Registry performs annual cluster analysis using the most recent 5 years of data. More than 5 years may tend to identify trends rather than clusters, and will be computationally slower. Less than 5 years may fail to detect if the most recent years are unusual compared to preceding years.

Since it is exposure during early pregnancy (organogenesis) that is relevant, it is preferable to use estimated date of conception rather than date of birth. Thus cases of different gestational ages (and terminations of pregnancy of low gestational age) are related to a common time when they passed through organogenesis. Cluster detection uses date of conception where gestational age is recorded for more than 90% of cases (for any one anomaly subgroup and registry) allowing its estimation. Where gestational age is missing, it is estimated on the basis of the average gestational age in the registry, by year, anomaly subgroup, and outcome of pregnancy. Gestational age is not estimated if it is missing for more than 10% of cases for the registry and anomaly subgroup, in which case cluster detection is based on date of birth.

Where date of conception is used as a basis for cluster detection, the conception period for statistical monitoring must end 9 months before the last birth month where data collection is complete. Where full years of data are used for surveillance, this means ending the period of the scan at 31 March (date of conception) of the last year of data collected. After this date, some conceptions may result in births in the next year which are not yet included in the dataset. Similarly, the start of the monitoring period must include a complete cohort of conceptions i.e. to detect clusters by date of conception for cases born (or with estimated date of delivery) in the approximate period 2007-2008, the scan routine includes cases with date of conception between 1st April 2006 and 31st March 2008 (24 months). The default monitoring period is set to start with estimated dates of conceptions from 1 January of the beginning of the five year period (for simplicity, and to ensure complete coverage of eligible conception outcomes). This means that cluster detection by date of conception is run on 51 months of data (4 complete years: 4 x 12 months = 48 months; and 3 months of the last year included in surveillance). Cluster detection by date of birth is run on 5 complete years of data (60 months).
The Scan method identifies both significant excesses and deficits of cases. Excesses are marked in the output as “clusters”, deficits as “deficits”.

The Scan method is based on case counts, and results would not be valid if there is large underlying change in population (births) size. There are two types of population change: change in geographical area that the registry covers, and large change in birth rate within the same geographical area. When basing cluster detection on conception cohorts, the latter could be easily taken into account by statistical modifications. However, it is the former type of population change that is more frequent, and a decision was therefore made to exclude registries where population (no. births) change is more than 10% between any two years within the five year period. This is under review.

The data requirements are:

- Full individual dataset including full date of birth (day, month, year), outcome of pregnancy (live/still/TOPFA), gestational age, malformation codes and their derived anomaly subgroups (see EUROCAT Guide 1.3).
- Population change (no. births) must be less than 10% between any two years.

Each dataset used for monitoring should be archived, both by Central Registry and by local registries, as the database is dynamic. Registries should archive a copy of the file transmitted to Central Registry in February of the year of monitoring, and to rename the file eg edmpdata_statmon08. With this file, registries should be able to reproduce Central Registry statistical monitoring results using the EDMP.

2.5 History of changes to the statistical software

- A new subgroup called "Severe CHD" was included in the list of subgroups (year 2008 monitoring).
- Adjusting Down syndrome trends for probability of survival to 20 weeks gestation of terminations of pregnancy and maternal age (year 2008 monitoring)
- Output of trend analysis. Significant increasing or decreasing trends are only reported where the p-value is <0.05 for the trend component and >0.01 for the non-linear component. Where p<0.05 for trend component and p<0.01 for non-linear component, the results are identified as 'non-linear change'. Where p>0.05 for trend component and <0.05 for non-linear component, the results are identified as 'non-linear change' (year 2008 monitoring onwards)
- Pan-Europe analysis whereby a trend analysis for all registries combined is performed (year 2007 monitoring onwards)
- Trend analysis: in addition to the expected number of cases criterion (i.e. at least an average of 5 per year or 2-year interval), trend analysis is now run if the observed number of cases in each year or 2 year interval is at least 2 (year 2007 monitoring onwards)
- The output of the trend analysis now includes a separate test for trend and a test for non-linearity (Chi square). All significant results are shown (p <0.05) (year 2007 monitoring onwards)
- The Adobe files showing graphical representation of clusters now includes a case list showing estimated date of conception, and ordered by date of conception (year 2006 onwards).
- Trend detection is now run using 10 years of data (where available), as opposed to 5 years of data. (year 2005 monitoring onwards)
- The output of the trend analysis shows observed cases grouped in 2 year intervals, e.g. 1996-97, 1998-99, 2000-01, 2002-03 and 2004-05. A minimum of 5 EXPECTED cases in each 2 year period is needed to be included in the trend analysis. (year 2005 monitoring onwards)
- The cluster output now includes a summary of trend test results run on the same five years of the cluster monitoring, to establish if a cluster could also be described as a trend. (year 2005 monitoring onwards)
- Chromosomal anomalies excluded from the statistical monitoring of non-chromosomal subgroups (year 2004 monitoring onwards)
- Cluster detection run by estimated date of conception, instead of date of birth/delivery, unless date of conception is unavailable. See EUROCAT Statistical

Monitoring Protocol http://www.eurocat-network.eu/content/Stat-Mon-Protocol-2004.pdf (year 2004 monitoring onwards)

- Instead of just showing the most significant cluster in each anomaly subgroup, we now show all significant clusters, which may overlap in time (year 2004 monitoring onwards)
- Heterogeneous large subgroups (e.g. digestive system) are not analysed for clusters. (year 2004 monitoring onwards)
- We only report clusters within or overlapping the last two years of data sent by your registry (i.e. “new” clusters), and only clusters of less than 18 months length.(year 2004 monitoring onwards)

3. Use of the EUROCAT Data management Program (EDMP) at local level

Registries should be performing monitoring earlier, and may need to consider smaller geographical areas for monitoring. The EDMP has been designed to facilitate this.

Trend tests.
Registries may wish to use the EDMP to expand the trend tests run centrally e.g. use a different period of years (e.g. since beginning of registry or including a more recent year), or congenital anomaly subgroups outside the selected subgroups for monitoring (these "user defined subgroups" can be defined within the EDMP)
Cluster detection
Local monitoring for clusters can be run more frequently using the most recently ascertained data. As a guide, local registries should run the EDMP statistical monitoring program at least every 6 months and as early as possible after data collection is complete (apart from late diagnosed cases).

Local registries can run cluster detection in geographical sub-areas of their registry area, by using the EDMP "user defined LOCATION subgroup", located under the REPORTS function. Registries must also enter the total births for the user defined registry sub-area.

Local registries can also choose to run cluster monitoring on non-standard anomaly subgroups of interest to them ("user defined subgroups" can be defined within the EDMP).

The time period of the scan is important and may affect the results. The reasons Central Registry has chosen a five year period are given above. Registries scanning more recent data every 6 months are recommended to add the extra months available to the basic 5 year period. Remember that dates of conception end 9 months before the last birth month available. Which clusters are detected and their statistical significance will depend on the number of years scanned, which is an a priori protocol decision to be made before running the software (i.e. do not make the cluster "disappear" or "appear" by trying many different time periods, as this is statistically invalid).

Clusters and trends identified locally should be investigated in the same way as those identified by central monitoring (see section 4), and reported at the following Registry Leaders Meeting and in the Annual Statistical Monitoring Report.

Following preliminary investigations, plausible clusters and trends not resulting from data quality errors can be reported to central registry for communication to all local registries for further investigation (ie. are similar clusters and trends occurring elsewhere). In this way the situation across Europe may be monitored for early detection of possible new teratogens.

4. How to investigate trends and clusters identified by statistical monitoring

See Appendix 5 for examples of the output available describing the results of the statistical monitoring.

The results of all cluster investigations in local registries will be part of our report to the EC and will be available on the open website http://www.eurocat-network.eu/CLUSTERSAndTRENDS/StatisticalMonitoring/StatisticalMonitoring-2007

Note that clusters arising from statistical monitoring are a special situation. Guidelines for investigating clusters more generally (e.g. those reported to you from the community or health professionals) are part of the EUROCAT Cluster Advisory Service http://www.eurocat-network.eu/CLUSTERSAndTRENDS/ClusterAdvisoryService/HowUnusualisanObservedCluster?

4.1 Guidelines for the preliminary investigation of trends

Concentrate your investigations on the increasing trends to report at the RLM and to Central Registry in order to detect issues of urgent concern. However, note that downward trends and heterogeneity are also of interest in relation to risk factors and public health action, and that investigation of all change over time can help you with data quality monitoring. You will be asked to report on a selection of downward trends where similar trends are occurring across Europe.

Consider the following, consulting with clinicians or others where appropriate:

- Case verification
    - Confirmed and accurate diagnoses?
    - Duplicates?
    - Resident within region? (truly population-based?)
- Print out a graph of the yearly prevalence (see Appendix 5) - is the trend gradual or a steep change? When does the trend appear to begin?
- Has there been a change in definition or diagnosis, or diagnostic methods e.g. increasing use of prenatal or postnatal ultrasound?
- Has there been a change in inclusion or exclusion criteria?
- Have there been changes in how cases are reported to the register?
- Is the trend found across all reporting hospitals?
- Is there a similar trend for any other anomalies?
- Is the trend found in isolated or multiply malformed cases?
- Have there been changes in register population?
- Are there any known changes in the risk factors for this anomaly?
- Have any other registers experienced a similar trend for this anomaly?
- For upward trends; is the prevalence in your registry at the end of the time period above the EUROCAT average?
- For downward trends; is the prevalence in your registry at the end of the time period below the EUROCAT average?

4.2 Guidelines for the preliminary investigation of clusters (see also Appendix 4)

- Case verification
  - Confirmed and accurate diagnoses?
  - Duplicates?
  - Confirmed to be resident within region? (truly population-based?)
- Diagnostic dimension:
  - How heterogeneous are the diagnoses? Are cases isolated, multiply malformed, syndromes? Any family history recorded?
  - Do any other anomalies have clusters at the same time?
- Space dimension:
  - Are they clustered near each other within region?
  - Do they come from a single hospital?
  - Do other regions have a cluster at a similar time? (use EUROCAT communication to query other registries).
  - NB. The aim is to describe the cluster in terms of its spatial characteristics, giving clues as to possible causes for further investigation. For example, a new drug on the market may not show spatial concentration, whereas a local chemical pollution accident would be expected to show spatial concentration.
- Time dimension:
  - Is the cluster part of a longer term trend identified by the trend analysis? (if so, investigate as trend rather than as cluster)
  - When does the increased risk appear to start and end – a longer period than the dates of the most statistically significant cluster itself? Look at other clusters in the cluster group to get an idea of the extent of the cluster.
  - If cluster is based on date of birth rather than date of conception, is it likely, making assumptions about gestational age of cases within and outside the cluster, that the cluster would also appear if analysed by date of conception?
- Diagnostic & reporting factors:
  - Could a change in diagnostic methods, training, personnel or reporting practice have caused the cluster? This might be particularly suspected if only one hospital is involved, and for anomaly subgroups which vary widely in severity
  - Does the registry have a lower rate before the cluster compared to other registries? Check EUROCAT website for prevalence in other registries (http://www.eurocat-network.eu/ACCESSPREVALENCEDATA/PrevalenceTables)
  - If there are a very large number of cases in the cluster in a very short time period, it is unlikely to be due to diagnostic factors.
- Aetiological factors:
  - Which factors have been investigated?– choice of factors to investigate depends on type of anomaly
  - Which factors have been looked at within the registry database (list variables) and outside the registry database and do any of these appear to explain the cluster?
- Local context:
  - Was there local awareness of the cluster before it was found by central statistical monitoring, either by local EDMP monitoring or other means?
  - Had anyone outside the registry in your region (e.g. local community or health professional) previously been aware of the cluster?

- Are there any local concerns about environmental exposures which may need investigation?

If investigation of clusters identifies data errors (e.g. incorrect diagnoses, incorrect dates of birth) these errors should be corrected and updated data included in the next data transmission to Central Registry. A record of these errors should be included in the local cluster report sent to Central Registry (MA.Loane@ulster.ac.uk).

Recommendation:
Follow up all cluster cases to their current age for further diagnostic information and family history. For clusters of chromosomal anomalies exact karotype should be reported for all cases in the cluster. If it is not already normal practice, registries should go back to original medical records for cluster cases. When a cluster is confirmed and recommended for continued surveillance registries should organise to do this surveillance earlier than the CR system.

4.3 Reporting the results of investigations to Central Registry

The following templates have been designed to assist registries with initial investigations into clusters and trends.

Cluster Template

Statistical Monitoring Investigations (*up to 2008 data*)
Run at Central Registry, *April 2010*

"Registry Name"

	Summary of Clusters for your registry	
	Anomaly	
3	Neural Tube Defects	By date of conception
79	Situs inversus	By date of conception

Please provide a written report for *each* cluster and send report to Maria Loane ma.loane@ulster.ac.uk, following the suggested format below:
The questions are explained in the "guidelines for the preliminary investigation of clusters" (Section 4.2 of Statistical Monitoring Protocol available at http://www.eurocat-network.eu/CLUSTERSAndTRENDS/StatisticalMonitoring/StatisticalMonitoring-2008 or attached at the bottom of this document for your convenience).

Please type as much as is necessary in the tables below. The boxes will expand to accommodate the text.

Has your registry used the EDMP statistical monitoring programme in the last year to look for clusters or trends in more recent data, or for different anomaly subgroups, or any other purpose? Yes/No. If yes, please give details. N.B. A summary of local monitoring will be included in the next Annual Statistical Monitoring Report.

Last year you had 1 cluster in the subgroup (e.g. Spina Bifida). We would be interested to know what the Public Health Authority response was to this.

Neural Tube Defects	**Methods and results of case verification**
	Dimensions Diagnostic dimension: Spatial dimension: Time dimension:
	Methods and results of any investigations as to whether changes in diagnostic or reporting practices might have contributed to the cluster
	Aetiological factors examined and result:
	Local context
	Conclusion: **Do you consider the cluster 'explained' by your preliminary investigation? Yes/No.** **If yes, give a summary of your explanation.** **If no,** **-Does the cluster require a further period of surveillance before a decision is made to investigate further? Why?** **OR** **-Is there going to be further aetiological/other investigation? Please give details.**
	Which public health authorities have been or will be notified about this cluster? Please give details.

Increasing Trend Template

Statistical Monitoring Investigations (*up to 2008 data***)**
Run at Central Registry, *April 2010***.**

Please investigate all the increasing trends (/) in your registry using the Guidelines for the preliminary investigation of trends, section 4.1 of the Statistical Monitoring Protocol http://www.eurocat-network.eu/CLUSTERSAndTRENDS/StatisticalMonitoring/StatisticalMonitoring-2008.

1. Please select an appropriate code from the list below and enter into the appropriate column

   A: Changes in case ascertainment (data quality)
   B: Changes in local or central registry methods e.g. definitions and inclusion criteria
   C: Changes in diagnostic methods
   D: Trend confirmed, due to known demographic changes (e.g. maternal age)
   E: Trend confirmed, investigation ongoing
   F: Trend confirmed, further surveillance proposed before more detailed investigation
   G: Trend not confirmed when additional years added, or trend concerns a heterogeneous subgroup with unlikely common cause.

2. Add an explanation
3. Fill in the column 'To which public health authority will the result be reported?'

Please type as much as is necessary in the questionnaire attached – the boxes will expand to accommodate the text.

Increasing Trend Template

"Registry Name" Trends 1999-2008

			To be completed by registry	To be completed by registry	To be completed by registry
	Anomaly	**"Increasing Trends"**	**Code A, B, C, D, E, F, G**	**Summary Explanation (please attach full report)**	**To which Public Health authority will the results be reported? (Please give details of how you will be reporting these results regionally and/or nationally)**
19	Transposition of great vessels	/			
66	Club foot - talipes equinovarus	/			
69	Syndactyly	/			

Statistical Monitoring Investigations (up to 2008 data)
Run at Central Registry, April 2010.

This year, in addition to the increasing trends investigation, we have selected 6 subgroups that showed a decreasing trend in 4 or more registries and also in the pan-Europe (all registries combined) analysis for further investigation (see below):

- **Severe CHD**
- **Ano-rectal atresia**
- **Neural Tube Defects**
- **Spina Bifida**
- **Congenital hydronephrosis**
- **Hip dislocation**

Please investigate the decreasing trends (\) in your registry. We need a brief indication as to whether you consider the trends *real* decreasing trends (if so, what are the known or possible reasons?), or due to registration or diagnostic factors (if so, which?). This will help EUROCAT comment on the general tendency for a decrease across Europe. Please type as much as is necessary in the questionnaire attached – the boxes will expand to accommodate the text. .There will be an opportunity to discuss these declining trends in Dublin.

Decreasing Trend Template

“Registry Name” Trends 1999-2008

			To be completed by registry
	Anomaly	**Decreasing Trends**	**Summary Explanation**
97	Severe CHD	\	
55	Congenital Hydronephrosis	\	

4.4 Next steps in investigation

Guidance on protocols for further cluster investigation can be found at http://www.eurocat-network.eu/CLUSTERSAndTRENDS/ClusterAdvisoryService/ClusterInvestigationProtocols. Note that it may be useful to expand the study beyond the original cluster rather than uniquely to focus on the cluster itself.

APPENDIX 1:

Adjustment of trends in Down syndrome for survival and maternal age

- Adjust each DS termination for probability they won't survive to 20 weeks
- Calculate total number DS fetuses at 20 weeks for each year (Observed DS in year I = ODS_i)
- Calculate expected number of DS live births for each year using numbers of population births according to 5 year maternal age groups for each year plus rate from published paper[3] (Expected DS in year I = EDS_i)
- Take the expected number of DS live births for all the years and sum to calculate the total number of expected Down syndrome live births (TOTEDS).
- Take the population data for all the years and sum to calculate the total number of population births (TOTPOP).
- Calculate the adjusted population in each year as the ratio of EDS in year I divided by TOTEDS and multiplied by TOTPOP
- Round this adjusted population figure to the nearest integer (as it will be large this will be a negligible adjustment)
- Calculate adjusted rate as observed number of DS fetuses divided by the adjusted population for each year.
- Calculate the chi-square for trend using the observed number of DS fetuses and the adjusted population in the same way as before

Detailed Calculations

1. For each Down syndrome termination apply the following weights according to the mothers age :

Table 1 Probability of survival to 20 weeks given mother's age and gestation
(From G Savva, personal communication)

Maternal age (years)	Gestation (weeks)										
	10	11	12	13	14	15	16	17	18	19	20
25	0.92	0.93	0.94	0.95	0.98	0.98	0.98	0.99	0.99	1.00	1.00
26	0.91	0.93	0.94	0.95	0.97	0.98	0.98	0.99	0.99	1.00	1.00
27	0.91	0.93	0.94	0.95	0.97	0.98	0.98	0.99	0.99	1.00	1.00
28	0.91	0.93	0.94	0.95	0.97	0.97	0.98	0.99	0.99	1.00	1.00
29	0.91	0.92	0.94	0.94	0.97	0.97	0.98	0.99	0.99	1.00	1.00
30	0.90	0.92	0.93	0.94	0.97	0.97	0.98	0.99	0.99	1.00	1.00
31	0.90	0.92	0.93	0.94	0.97	0.97	0.98	0.99	0.99	1.00	1.00
32	0.90	0.92	0.93	0.94	0.97	0.97	0.98	0.99	0.99	1.00	1.00
33	0.89	0.91	0.93	0.94	0.97	0.97	0.98	0.99	0.99	1.00	1.00

[3] Morris JK, Mutton DE, Alberman E. Revised estimates of the maternal age specific live birth prevalence of Down's syndrome. Journal of Medical Screening 2002; 9:2-6.

34	0.89	0.91	0.93	0.93	0.97	0.97	0.98	0.98	0.99	1.00	1.00
35	0.89	0.91	0.92	0.93	0.97	0.97	0.98	0.98	0.99	1.00	1.00
36	0.88	0.91	0.92	0.93	0.96	0.97	0.98	0.98	0.99	1.00	1.00
37	0.88	0.90	0.92	0.93	0.96	0.97	0.98	0.98	0.99	1.00	1.00
38	0.88	0.90	0.92	0.93	0.96	0.97	0.98	0.98	0.99	1.00	1.00
39	0.87	0.90	0.91	0.92	0.96	0.96	0.98	0.98	0.99	1.00	1.00
40	0.87	0.89	0.91	0.92	0.96	0.96	0.98	0.98	0.99	1.00	1.00
41	0.86	0.89	0.91	0.92	0.96	0.96	0.98	0.98	0.99	1.00	1.00
42	0.86	0.89	0.91	0.92	0.96	0.96	0.97	0.98	0.99	1.00	1.00
43	0.86	0.88	0.90	0.91	0.96	0.96	0.97	0.98	0.99	1.00	1.00
44	0.85	0.88	0.90	0.91	0.95	0.96	0.97	0.98	0.99	1.00	1.00
45	0.85	0.88	0.90	0.91	0.95	0.96	0.97	0.98	0.99	1.00	1.00
46	0.84	0.87	0.89	0.91	0.95	0.96	0.97	0.98	0.99	0.99	1.00
47	0.84	0.87	0.89	0.90	0.95	0.95	0.97	0.98	0.99	0.99	1.00
48	0.83	0.87	0.89	0.90	0.95	0.95	0.97	0.98	0.99	0.99	1.00
49	0.83	0.86	0.88	0.90	0.95	0.95	0.97	0.98	0.99	0.99	1.00
50	0.82	0.86	0.88	0.89	0.95	0.95	0.97	0.97	0.99	0.99	1.00
Mean	0.89	0.91	0.92	0.93	0.97	0.97	0.98	0.98	0.99	1.00	1.00

Using the mean when maternal age is missing and assuming in mothers < 25 the rates for 25 year olds.
For cases with gestational age missing, assume a gestational age of 1.00

2. Take no of births (unaffected) in each register for each year according to maternal age in 5 year age groups.
3. Within each register for each year calculate the expected numbers of DS live births as the observed numbers of (unaffected) births in 5 year age groups in that year times “DS live birth rate in table 2” for that age group and then total this up for all mothers for each year.

Table 2 : DS Live birth rate according to 5 year age groups (different classifications according to population data that is available)

Age group	DS live birth rate per 1000 births	Age group	DS live birth rate per 1000 births	Age group	DS live birth rate per 1000 births
< 20	0.6678	<15	0.6590	<20	0.6678
20-24	0.6992	15-19	0.6679	20-24	0.6992
25-29	0.8395	20-24	0.6992	25-29	0.8395
30-34	1.4764	25-29	0.8395	30-34	1.4764
35 +	6.4980	30-34	1.4764	35-39	4.7054
		35-39	4.7054	40+	15.9289
		40-44	15.1828		
		45-49	30.2248		
		50+	37.9947		

Table 2 is calculated using formula in Morris et al for birth rate plus weighting by births in England and Wales from 1989-2007

4. Within each register calculate the weighted number of DS pregnancies achieving 20 weeks gestation = DS Live births + DS Still births + DS weighted terminations (excluding miscarriages < 20 weeks) for all mothers (ie ignore maternal age).
5. Take the expected number of DS live births for all the years and sum to calculate the total number of expected Down syndrome live births (TOTEDS).
6. Take the population data for all the years and sum to calculate the total number of population births (TOTPOP).
7. Calculate the adjusted population in each year as the ratio of EDS in year I divided by TOTEDS and multiplied by TOTPOP
8. Round this adjusted population figure to the nearest integer (as it will be large this will be a negligible adjustment)
9. Calculate adjusted rate as observed number of DS fetuses divided by the adjusted population for each year.
10. Calculate the chi-square for trend using the observed number of DS fetuses and the adjusted population in the same way as you did before

APPENDIX 2.

Scan statistic – procedure for calculation

Scan statistic formula based on the formula defined by Nagarwalla (1996).

Data
The date for each case is provided. A start and end date are specified and each case date is represented as a fraction of the period of interest.

$$fraction_{case} = \frac{date_{case} - date_{start}}{date_{end} - date_{start}}$$

Methods

Calculation of lambda
For a given subset of cases, it is possible to calculate the test statistic lambda as follows:

$$\lambda = \left(\frac{n}{r}\right)^n \left(\frac{r-n}{n}\right)^{r-n} \left(\frac{1}{d}\right)^n \left(\frac{1}{1-d}\right)^{r-n}$$

Where r is the total number of cases, n is the number of cases in the subset, also called the scanning window, and d is the date fraction spanned by the subset of cases. Typically a minimum scanning window of 5 is used. Lambda is calculated for every possible subset of n consecutive cases for $n = 5,\ldots, r$

The first step is to generate simulated datasets with random numbers. For each simulation, the lambda for each subset is calculated and the largest lambda is recorded. After, for example, 999 iterations the lambdas are ordered by size from smallest to largest and the 95th percentile is noted as λ_{sig}. In order to save computing time, and to give accurate p-values based on sufficient iterations, a look-up table was created for all r between 7 and 700 cases, using 100,000 iterations for r under 200 cases and 50,000 iterations for r over 200 cases. For r of more than 700 cases, Monte Carlo simulations (999 iterations) are used to create an array of lambdas.

For the real dataset, the lambda for each subset of cases is calculated. If the lambda is greater than or equal to λ_{sig} then the cases are designated as a significant cluster and the details of the cluster, such as start date and duration, are recorded. The p-value is recorded as the position of lambda within the simulated lambdas.

As a first step, a 'quick' check is performed to test whether the most significant cluster is significant at p<=0.1. If no significant cluster, then move on to next anomaly subgroup. If significant cluster is found then a full check is performed for all clusters with a p-value less than 0.06.

Group allocation
The final step is to allocate clusters to groups based on their overlap. The cluster with the largest lambda is identified and is allocated to the first group. Each ungrouped cluster is then tested for overlap with the first cluster. Overlap is calculated as the number of cases appearing in both

clusters divided by the number of cases appearing in either cluster. If the overlap is greater than or equal to 0.75 then both are given the same group label. If any clusters remain ungrouped then the cluster with the largest lambda is identified and given a new group label and the procedure is repeated until all clusters are grouped.

Nagarwalla, N. (1996). A scan statistic with a variable window. Statistics in Medicine 15, 845-850

APPENDIX 3.

List of congenital anomaly subgroups for monitoring.

The EUROCAT congenital anomaly subgroups are defined in EUROCAT Guide 1.3, Chapter 3.3 (http://www.eurocat-network.eu/content/EUROCAT-Guide-1.3.pdf). The following list shows which subgroups are to be analysed in the following ways:

1. Prevalence by outcome of pregnancy, by registry and year (or combined registries/years). All cases, and all cases excluding chromosomal cases.
2. Analysis of trends, all outcomes of pregnancy combined. Chromosomal cases excluded from all subgroups except chromosomal subgroups.
3. Detection of clusters, all outcomes of pregnancy combined. Chromosomal cases excluded from all subgroups except chromosomal subgroups.

	Prevalence by pregnancy outcome, registry, year	Included in cluster analysis	Included in trend analysis
All anomalies	X		X
Nervous system	X		X
Neural Tube Defects	X	X	X
Anencephalus and similar	X	X	X
Encephalocele	X	X	X
Spina Bifida	X	X	X
Hydrocephaly	X	X	X
Microcephaly	X	X	X
Arhinencephaly/ holoprosencephaly	X	X	X
Eye	X		X
Anophthalmos/ micropthalmos	X	X	X
Anophthalmos	X	X	X
Congenital cataract	X	X	X
Congenital glaucoma	X	X	X
Ear, face and neck	X		X
Anotia	X	X	X
Congenital heart disease	X		X
Common arterial truncus	X	X	X
Transposition of great vessels	X	X	X
Single ventrical	X	X	X
Ventricular septal defect (VSD)	X	X	X
Atrial septal defect (ASD)	X	X	X
Atrioventricular septal defect (AVSD)	X	X	X
Tetralogy of Fallot	X	X	X
Tricuspid atresia and stenosis	X	X	X
Ebstein's anomaly	X	X	X
Pulmonary valve stenosis	X	X	X
Pulmonary valve atresia	X	X	X
Aortic valve atresia/stenosis	X	X	X
Hypoplastic left heart	X	X	X
Hypoplastic right heart	X	X	X
Coarctation of aorta	X	X	X
Total anomalous pulmonary venous return	X	X	X
Respiratory	X		X

Choanal atresia	X	X	X
Cystic adenomatous malf of lung	X	X	X
Oro-facial clefts	X		X
Cleft lip with or without palate	X	X	X
Cleft palate	X	X	X
Digestive system	X		X
Oesophageal atresia with/ without tracheo-oesophagal fistula	X	X	X
Duodenal atresia or stenosis	X	X	X
Atresia or stenosis of other parts of small intestine	X	X	X
Ano-rectal atresia and stenosis	X	X	X
Hirschsprung's disease	X	X	X
Atresia of bile ducts	X	X	X
Annular pancreas	X	X	X
Diaphragmatic hernia	X	X	X
Abdominal wall defects	X		X
Gastroschisis	X	X	X
Omphalocele	X	X	X
Urinary	X		X
Bilateral renal agenesis including Potter syndrome	X	X	X
Cystic kidney disease	X	X	X
Congenital hydronephrosis	X	X	X
Bladder extrophy and/or epispadia	X	X	X
Posterior urethral valve and/or prune belly	X	X	X
Genital	X		X
Hypospadias	X	X	X
Indeterminate sex	X	X	X
Limb	X		X
Limb reduction	X	X	X
Upper limb reduction	X	X	X
Lower limb reduction	X	X	X
Complete absence of a limb	X	X	X
Club foot - talipes equinovarus	X	X	X
Hip dislocation and/or dysplasia	X	X	X
Polydactyly	X	X	X
Syndactyly	X	X	X
Arthrogryposis multiplex congenita	X	X	X
Musculo-skeletal	X		X
Thanatophoric dwarfism	X		X
Jeunes syndrome	X		X
Achondroplasia	X		X
Craniosynostosis	X	X	X
Congenital constriction bands/amniotic band	X	X	X
Other malformations	X		X
Asplenia	X	X	X
Situs inversus	X	X	X
Conjoined twins	X	X	X
Disorders of skin	X	X	X
Teratogenic syndromes with malformations	X		X
Fetal alcohol syndrome	X	X	X
Valproate syndrome	X	X	X
Warfarin syndrome	X	X	X
Maternal infections resulting in malformations	X	X	X
Genetic syndromes & microdeletions	X		X

Chromosomal	X		X
Down's syndrome	X	X	X
Patau syndrome/ trisomy 13	X	X	X
Edward syndrome/ trisomy 18	X	X	X
Turner's syndrome	X	X	X
Klinefelter's syndrome	X	X	X
Cri-du-chat syndrome	X	X	X
Wolff-Hirschorn syndrome	X	X	X

APPENDIX 4

Investigating a cluster of birth defects
Elisabeth Robert-Gnansia, Lyon FRANCE

Investigations vary according to the anomaly of interest and local conditions. Elisabeth Robert has prepared notes to guide investigations of clusters arising from statistical monitoring, based on long experience in the Central East France register and elsewhere.

A cluster can be defined as a space-time aggregation of occurrences, in this context of birth defects. An example of such a cluster is the occurrence of 5 infants with a given malformation born in a geographic area during 4 months and with 3,000 annual births. If the overall rate for this malformation in the region is 1/6,000, it means that the expected number of cases during the 4 months is 0.17, and 5 were found. Such aggregations can have many causes. The reason for being interested in them is that the cases may have been caused by some local source of exposure. One often thinks of industrial pollution as a cause, but other factors may explain the cluster, including many artefactual events.

Clusters may be identified as a local observation by people living in a specific area, by health care workers, or by the routine surveillance of congenital malformations or other pregnancy outcomes. In rare cases, the detection of a cluster is forwarded from the news media, or a pressure group who may perceive that a specific source of pollution may cause birth defects and therefore a cluster must have occurred around that source. It then becomes our responsibility to investigate the identified cluster. Some review articles and papers on statistical methodology have questioned the usefulness of investigating clusters, and of monitoring programs in general for the detection of a newly introduced teratogen.

There are many arguments against cluster investigations. First, there is a great statistical probability that many clusters occur by chance alone, because there are many serious congenital malformations monitored routinely and repeated evaluations. Second, in evaluating situations in which an environmental agent has been evidenced as harmful, it appears that most clusters have occurred in areas where no registry was operative or where the hypothesis came from observations made by astute clinicians. Third, clusters are often alleged in relationship to a perceived environmental hazard, such as a toxic waste site, and the investigating team often must extend the investigation far beyond what is indicated scientifically in order to respond to community concerns.

Among the possible explanations for a cluster, one thinks first of a local harmful factor, but the actual cause may be different: local temporal variations in diagnosis, in ascertainment, local aggregation of genes, e.g. due to population migrations, demographic or socioeconomic variations, and finally a random phenomenon. Clusters are different to long time trends: a steep increase of urinary flow anomalies has occurred from the early 1980s, obviously due to better and earlier diagnosis through fetal ultrasonography. After having excluded all artefactual explanations, sometimes one nevertheless may be unable to answer the question "Is there a cluster?"

Examples can be given of surveillance events that may or may not be clusters: a marked increase in the prevalence at birth of hypospadias was observed in several registries in the world from the end of the 1980s, but no clear cut explanation can be given: improved ascertainment of minor forms or real increase? An sharp increase of the incidence of gastroschisis has been observed as well in many western registries since the beginning of the 90s, the reality of which is much less questionable than the one of hypospadias, but no explanation has been found so far.

When a cluster is considered as likely, other questions are to be answered:

1. To follow up or not to follow up? The decision to follow-up means spending time and money, and run the risk of spreading rumours.

2. How to follow up?

This should be performed in 4 stages, which might be called (a) nosodetermination, (b) chronodetermination, (c) geodetermination, (d) etiodetermination.

The first stage (*nosodetermination*) consists in confirming the case status, with exact diagnoses, and definition of inclusion and exclusion criteria: it should be checked that the cluster is not due to inclusion of minor forms of the studied malformations, which were not included in the baseline data. This is especially true for limb reduction defects (inclusions of missing phalanges?) or hypospadias (distal forms). Also diagnostic methods or medial personnel might have changed (e.g. more trained ultrasonographist or neonatologist), and explain an increased number of cases.

The second stage (*chronodetermination*) consists in checking the clustering in time: dates of birth or of termination of pregnancy are to be clearly assessed. Gestational ages are to be taken into consideration, especially when dealing with a cluster arising over a short period of time.

The third stage (*geodetermination*) should check the clustering in space. Some malformed babies might be transferred in utero for a better management of a surgical malformation at birth. One should then eliminate the possibility of a concentration of cases in university hospitals that attract patients with rare malformations, by assessing correctly the place where the child's family actually lived during pregnancy.

The fourth stage (*etiodetermination*) is the most difficult one, and consists in collecting detailed exposure information on cluster cases. A more detailed form than the routine one is to be used in order to collect all this information. If an environmental exposure is suspected, one should try to study similarly exposed areas, and to set up a case-control study within the area (3). In any case, it is essential to continue the surveillance.

The experience of registries that have investigated numerous clusters of congenital malformations shows that:

- Clusters will occur and will be found if looked for
- Most clusters have technical explanations or are artefacts
- Most clusters will go away when you look them “in the eyes”
- Nevertheless, clusters may be important for the identification of hazards in the environment and we have to investigate them

To decide whether one should follow up a cluster or not is difficult and depends mainly on available resources. If one investigates a cluster, most likely no explanation will be found. A unique example of a routine cluster investigation that led to an environmental factor as a possible explanation is given by a study led in the county of Östergötland, in Sweden. A higher prevalence at birth of congenital malformations was noticed (1), and looking at data revealed that cardiovascular malformations (CVM) explained the differences in rates. A case-referent study was performed (2), which retrieved several known risk factors for CVM (maternal diabetes, body mass index over 29, involuntary childlessness, previous spontaneous abortion, first trimester exposure to thyroid drugs and NSAIDs). All the significantly high ORs were higher in the studied county than in the rest of Sweden. Geographic comparisons of CVM were made within the county (3), and the hypothesis arose that the clustering in time and space showed that drinking water chlorination might be a mild risk factor for CVM (4).

Acknowledgement
This summary was adapted from a lecture by Pr Bengt Källén (Lund), with permission.

References
1. Blomberg M, Selbing A, Kallen B.Congenital malformations in the southeast of Sweden--a registry study with validation. Acta Paediatr. 2000 Oct;89(10):1238-43 2. Cedergren MI, Selbing AJ, Kallen BA. Risk factors for cardiovascular malformation--a study based on prospectively collected data. Scand J Work Environ Health. 2002;28:12-7.
3. Cedergren M, Selbing A, Kallen B. Geographic variations in possible risk factors for severe cardiac malformations. Acta Paediatr. 2002;91:222-8.
4. Cedergren MI, Selbing AJ, Lofman O, Kallen BA. Chlorination byproducts and nitrate in drinking water and risk for congenital cardiac defects. Environ Res. 2002;89:124-30.

APPENDIX 5.

EUROCAT Data Management Program (EDMP) output – an explanation

For each registry, the results are output in:

I. An EXCEL file containing 5 worksheets (Read me, Summary, Trend, Trend2, Clusters). This must be examined first. See Figure 1a.
II. An Adobe file ("Registry name_tr graphs") containing graphs of anomaly subgroups with significant trends. See Figure 2
III. An Adobe file ("Registry name_cluster cases") containing a full case list of cases within the subgroup where a cluster was identified. See Figure 3.

I. EXCEL file:

It is important to read the EXCEL "Read me" instructions (first Sheet, see Figure 1a), as this explains some of the key elements of the Statistical Monitoring exercise.

The second sheet "Summary" lists all the significant results detected for the registry (see Figure 1b):

- Significant trend results (5 year period)
- Significant trend results (10 year period)
- Significant clusters (5 year period)

The third sheet "Trend" shows the five year trend results. Output is in the same format as explained below for "Trend2".

Long-term (10 year) trend analysis results are in the fourth sheet "Trend2" (see Figure 1c) – output shows:

i. the number of cases per subgroup per 2 year period, and the total number of cases over the time period.
ii. A Chi square test for trend giving the direction (upward or downward) and slope of any linear trend and the statistical significance (p value).
iii. A Chi square test for non-linear change reporting the Chi square statistic and the statistical significance (p-value).
iv. The last column presents the overall description of the identified trend/linear change.

Cluster analysis results are in the fifth sheet "Clusters" (Figure 1d).

i. Type of cluster analysis run (date of conception or date of birth), the number of cases in the most significant cluster, the start and end date of most significant cluster, the number of expected cases in the time period of the most significant cluster, and the level of statistical significance (p-value and lamda)
ii. The total number of valid cases within the subgroup (i.e. must have valid date of birth) and the proportion of cases with missing gestational age is presented.
iii. The next column explains if an excess of cases was reported (cluster) or if less cases than expected were reported (deficit)

iv. The last column shows the trend test result run on the same five years of the cluster monitoring, to establish if the cluster may be part of an overall trend.

II. Adobe file " Registry name_tr graphs" shows graphical displays of significant trends. See Figure 2 below.

III. Adobe file "Registry name_cluster cases" gives more detail on all significant clusters. See Figure 3.

i. ALL cases in the anomaly subgroup both within and outside cluster are listed with local ID, date of conception, date of birth, gestational age (GA), estimated GA, and day case occurred.

ii. Lambda values for the most significant cluster are listed, along with the lambda values for all possible overlapping clusters. The highest lambda value indicates the most unusual cluster.

iii. Many clusters may overlap in time, with the inclusion or exclusion of individual cases changing their significance. The "most" significant cluster (highest "lambda" value) is listed first, followed by all significant clusters, with "cluster group" showing groups of highly overlapping clusters.

iv. A timeline shows the occurrence of all cases within the anomaly subgroup over the time period. A solid line displays the most significant cluster within the time period with a dotted line on either side showing the longest extent of the cluster group.

v. It is important to remember that if a cluster is detected using DATE OF CONCEPTION, the timeline represents the date of conception of those cases.

Figure 1a: Brief explanation of the key elements of Statistical Monitoring.

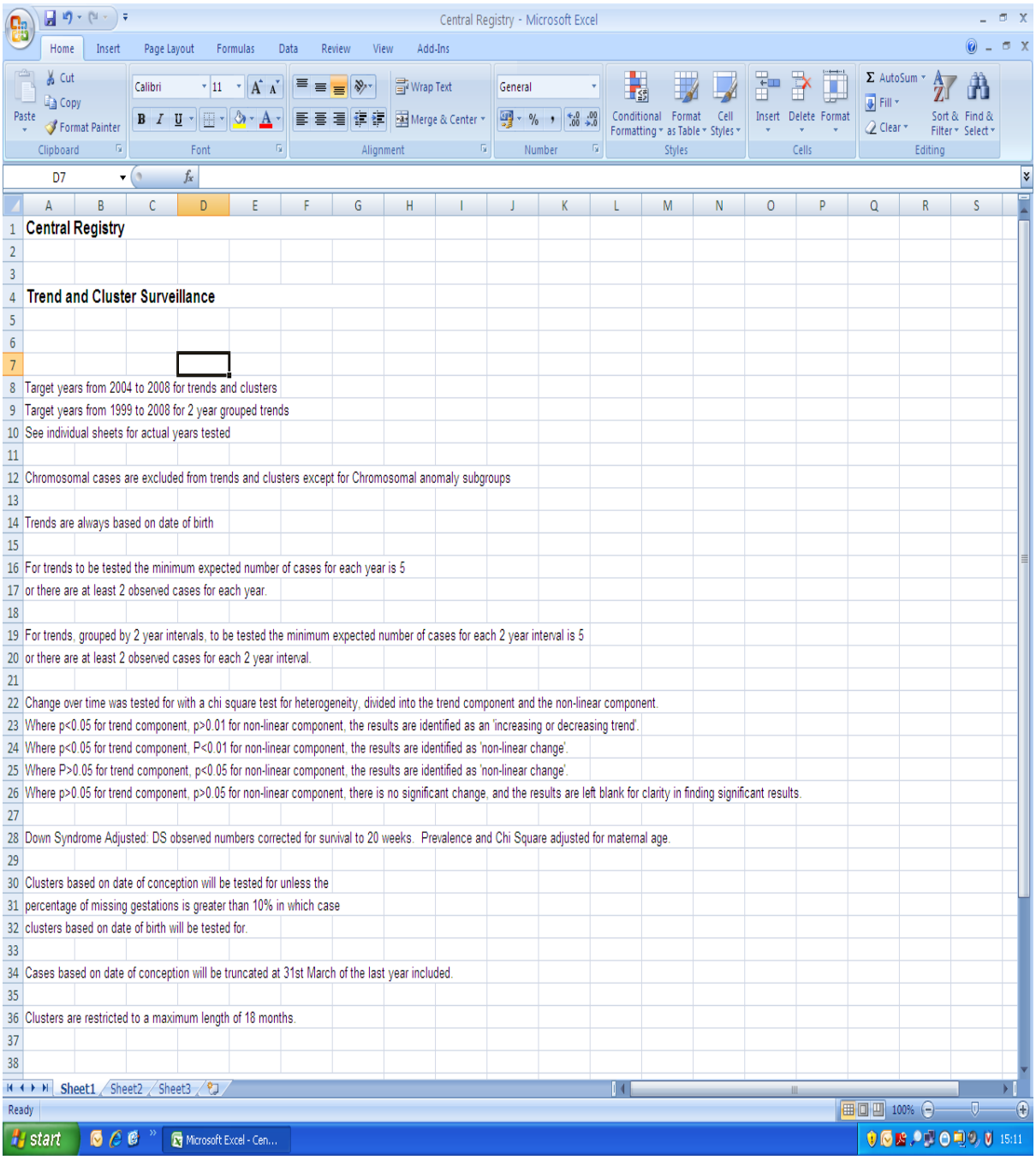
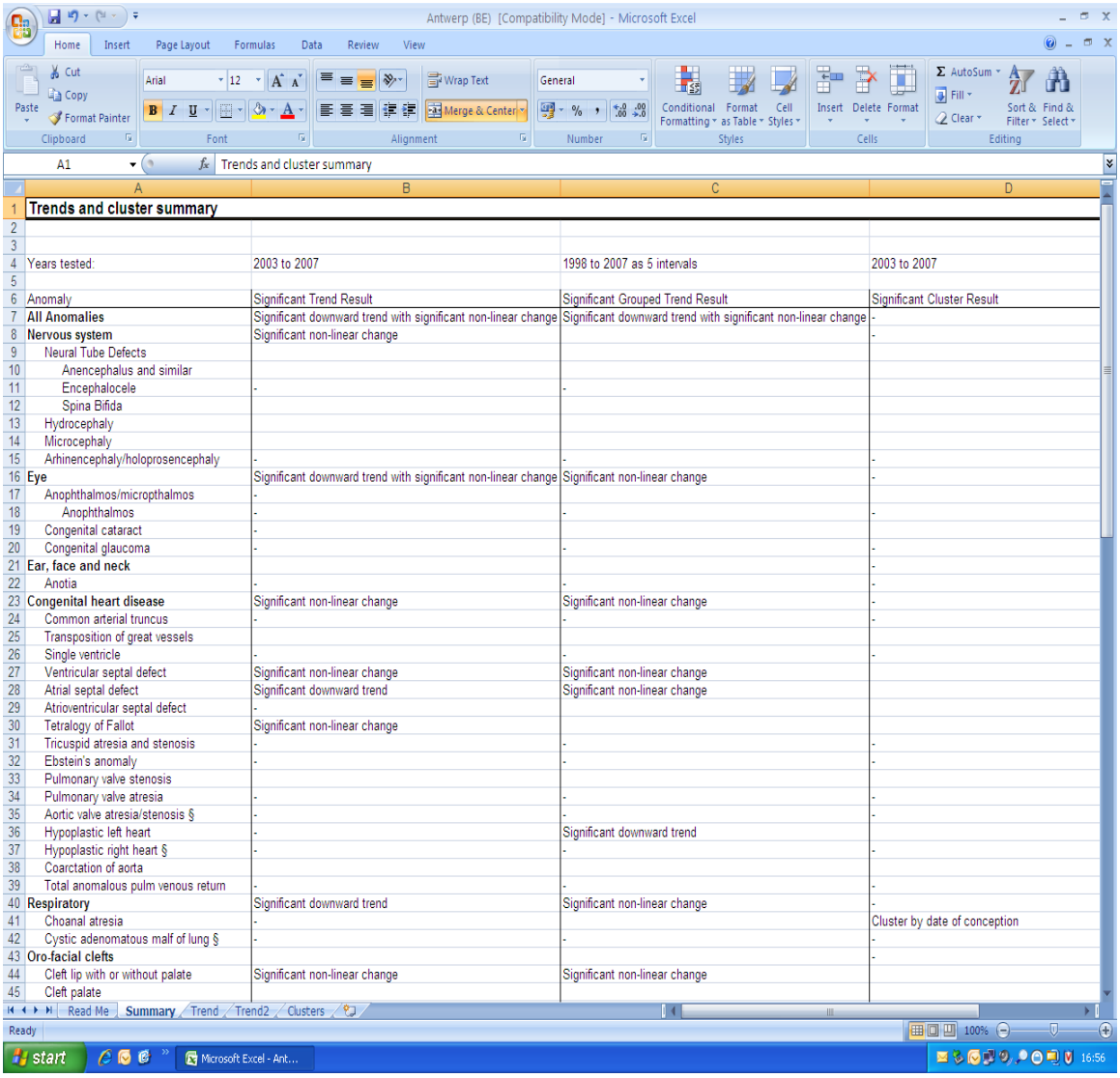


Figure 1b: Output lists all the significant trends and clusters identified in Statistical Monitoring.



Antwerp (BE) [Compatibility Mode] - Microsoft Excel

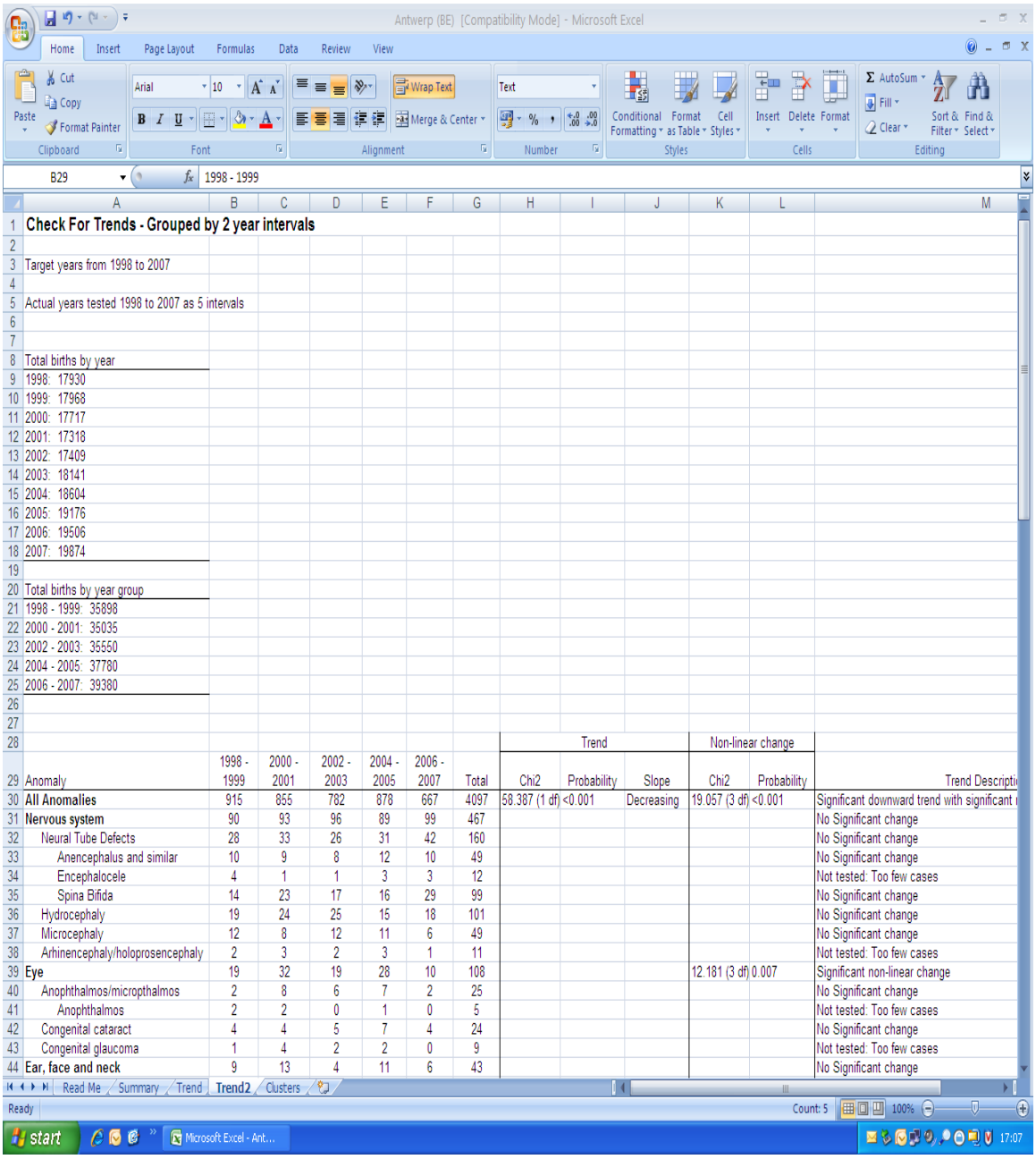
A1: Trends and cluster summary

	A	B	C	D
1	**Trends and cluster summary**			
2				
3				
4	Years tested:	2003 to 2007	1998 to 2007 as 5 intervals	2003 to 2007
5				
6	Anomaly	Significant Trend Result	Significant Grouped Trend Result	Significant Cluster Result
7	**All Anomalies**	Significant downward trend with significant non-linear change	Significant downward trend with significant non-linear change	-
8	**Nervous system**	Significant non-linear change		-
9	Neural Tube Defects			
10	Anencephalus and similar			
11	Encephalocele	-	-	
12	Spina Bifida			
13	Hydrocephaly			
14	Microcephaly			
15	Arhinencephaly/holoprosencephaly	-	-	-
16	**Eye**	Significant downward trend with significant non-linear change	Significant non-linear change	-
17	Anophthalmos/micropthalmos	-		
18	Anophthalmos	-	-	-
19	Congenital cataract	-		
20	Congenital glaucoma	-	-	-
21	**Ear, face and neck**			-
22	Anotia	-	-	-
23	**Congenital heart disease**	Significant non-linear change	Significant non-linear change	-
24	Common arterial truncus	-	-	-
25	Transposition of great vessels			
26	Single ventricle	-	-	-
27	Ventricular septal defect	Significant non-linear change	Significant non-linear change	
28	Atrial septal defect	Significant downward trend	Significant non-linear change	
29	Atrioventricular septal defect	-		
30	Tetralogy of Fallot	Significant non-linear change		
31	Tricuspid atresia and stenosis	-	-	-
32	Ebstein's anomaly	-	-	-
33	Pulmonary valve stenosis			
34	Pulmonary valve atresia	-	-	-
35	Aortic valve atresia/stenosis §	-	-	-
36	Hypoplastic left heart	-	Significant downward trend	
37	Hypoplastic right heart §	-	-	-
38	Coarctation of aorta			
39	Total anomalous pulm venous return	-	-	-
40	**Respiratory**	Significant downward trend	Significant non-linear change	-
41	Choanal atresia	-		Cluster by date of conception
42	Cystic adenomatous malf of lung §	-	-	-
43	**Oro-facial clefts**			-
44	Cleft lip with or without palate	Significant non-linear change	Significant non-linear change	
45	Cleft palate			

Read Me | Summary | Trend | Trend2 | Clusters

Ready

Figure 1c: Output showing the results of the 10 year trend test analyses.



Check For Trends - Grouped by 2 year intervals

Target years from 1998 to 2007

Actual years tested 1998 to 2007 as 5 intervals

Total births by year

1998:	17930
1999:	17968
2000:	17717
2001:	17318
2002:	17409
2003:	18141
2004:	18604
2005:	19176
2006:	19506
2007:	19874

Total births by year group

1998 - 1999:	35898
2000 - 2001:	35035
2002 - 2003:	35550
2004 - 2005:	37780
2006 - 2007:	39380

							Trend			Non-linear change		
Anomaly	1998 - 1999	2000 - 2001	2002 - 2003	2004 - 2005	2006 - 2007	Total	Chi2	Probability	Slope	Chi2	Probability	Trend Descripti
All Anomalies	915	855	782	878	667	4097	58.387 (1 df)	<0.001	Decreasing	19.057 (3 df)	<0.001	Significant downward trend with significant
Nervous system	90	93	96	89	99	467						No Significant change
Neural Tube Defects	28	33	26	31	42	160						No Significant change
Anencephalus and similar	10	9	8	12	10	49						No Significant change
Encephalocele	4	1	1	3	3	12						Not tested: Too few cases
Spina Bifida	14	23	17	16	29	99						No Significant change
Hydrocephaly	19	24	25	15	18	101						No Significant change
Microcephaly	12	8	12	11	6	49						No Significant change
Arhinencephaly/holoprosencephaly	2	3	2	3	1	11						Not tested: Too few cases
Eye	19	32	19	28	10	108				12.181 (3 df)	0.007	Significant non-linear change
Anophthalmos/micropthalmos	2	8	6	7	2	25						No Significant change
Anophthalmos	2	2	0	1	0	5						Not tested: Too few cases
Congenital cataract	4	4	5	7	4	24						No Significant change
Congenital glaucoma	1	4	2	2	0	9						Not tested: Too few cases
Ear, face and neck	9	13	4	11	6	43						No Significant change

Figure 2: Graphical representation of trend surveillance

Significant trends found between 1998 and 2007

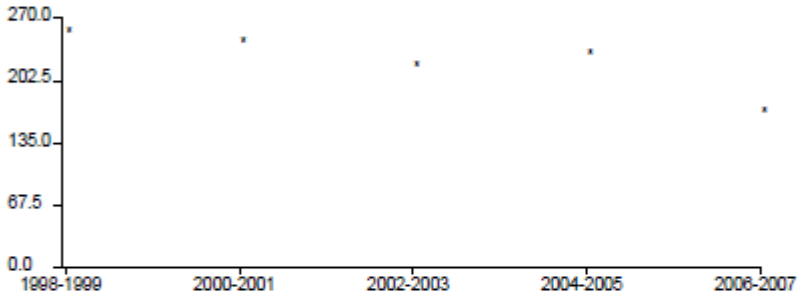
Anomaly Subgroup: **All Anomalies**

Significant downward trend with significant non-linear change

Test for Trend			Test for non-linear change		
Chi2	58.387	(1 df)	Chi2	19.057	(3 df)
p value	<0.001		p value	<0.001	
Slope	Decreasing				

Year groups:	Cases:	Total births:	Rate (per 10,000 births)
1998 - 1999	915	35898	254.89
2000 - 2001	855	35035	244.04
2002 - 2003	782	35550	219.97
2004 - 2005	878	37780	232.40
2006 - 2007	667	39380	169.38

Trend plot Y axis: Rate per 10,000 births X axis: Year of birth



Anomaly Subgroup: **Eye**

Significant non-linear change

Test for Trend			Test for non-linear change		
Not significant			Chi2	12.181	(3 df)
			p value	0.007	

Year groups:	Cases:	Total births:	Rate (per 10,000 births)
1998 - 1999	19	35898	5.29
2000 - 2001	32	35035	9.13
2002 - 2003	19	35550	5.34
2004 - 2005	28	37780	7.41
2006 - 2007	10	39380	2.54

Trend plot Y axis: Rate per 10,000 births X axis: Year of birth

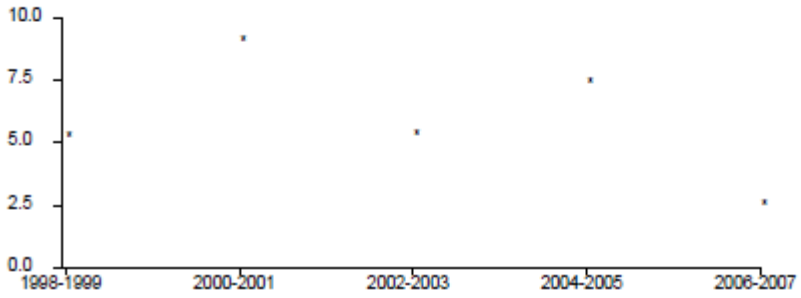
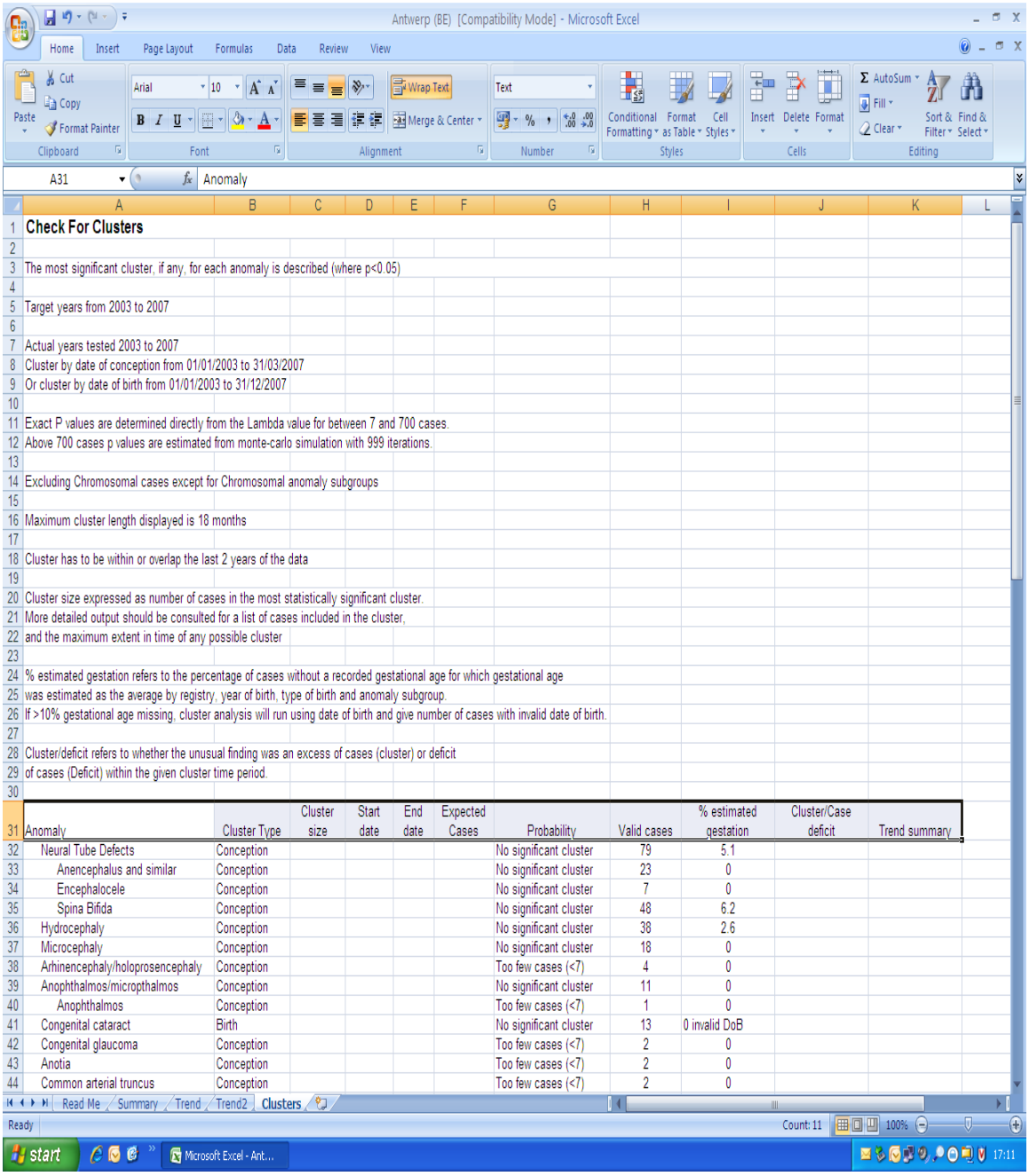


Figure 1d: Output showing the results of the cluster monitoring



Check For Clusters

The most significant cluster, if any, for each anomaly is described (where p<0.05)

Target years from 2003 to 2007

Actual years tested 2003 to 2007
Cluster by date of conception from 01/01/2003 to 31/03/2007
Or cluster by date of birth from 01/01/2003 to 31/12/2007

Exact P values are determined directly from the Lambda value for between 7 and 700 cases.
Above 700 cases p values are estimated from monte-carlo simulation with 999 iterations.

Excluding Chromosomal cases except for Chromosomal anomaly subgroups

Maximum cluster length displayed is 18 months

Cluster has to be within or overlap the last 2 years of the data

Cluster size expressed as number of cases in the most statistically significant cluster.
More detailed output should be consulted for a list of cases included in the cluster,
and the maximum extent in time of any possible cluster

% estimated gestation refers to the percentage of cases without a recorded gestational age for which gestational age
was estimated as the average by registry, year of birth, type of birth and anomaly subgroup.
If >10% gestational age missing, cluster analysis will run using date of birth and give number of cases with invalid date of birth.

Cluster/deficit refers to whether the unusual finding was an excess of cases (cluster) or deficit
of cases (Deficit) within the given cluster time period.

Anomaly	Cluster Type	Cluster size	Start date	End date	Expected Cases	Probability	Valid cases	% estimated gestation	Cluster/Case deficit	Trend summary
Neural Tube Defects	Conception					No significant cluster	79	5.1		
Anencephalus and similar	Conception					No significant cluster	23	0		
Encephalocele	Conception					No significant cluster	7	0		
Spina Bifida	Conception					No significant cluster	48	6.2		
Hydrocephaly	Conception					No significant cluster	38	2.6		
Microcephaly	Conception					No significant cluster	18	0		
Arhinencephaly/holoprosencephaly	Conception					Too few cases (<7)	4	0		
Anophthalmos/micropthalmos	Conception					No significant cluster	11	0		
Anophthalmos	Conception					Too few cases (<7)	1	0		
Congenital cataract	Birth					No significant cluster	13	0 invalid DoB		
Congenital glaucoma	Conception					Too few cases (<7)	2	0		
Anotia	Conception					Too few cases (<7)	2	0		
Common arterial truncus	Conception					Too few cases (<7)	2	0		

Figure 3: Adobe File “Registry name_cluster cases”. Graphical representation of detected clusters.

Anomaly Subgroup: **Choanal atresia**

Cluster type: Cluster by date of conception

Date range: Clusters between 01/01/2003 and 31/03/2007

Most significant cluster

Number of cases: 6 Expected number of cases: 0.89 p value: 0.021

Start date: 14/08/2006

End date: 07/12/2006

Record	Date of conceptior	Days	Date of birth	Gestation	GA estimated	Local ID
1	08/05/2003	127	15/01/2004	36	☐	2004027
2	06/09/2003	248	10/04/2004	31	☐	2004252
3	25/02/2004	420	08/12/2004	41	☐	2004695
4	22/05/2004	507	26/02/2005	40	☐	2005061
5	26/06/2004	542	02/04/2005	40	☐	2005784
6	07/04/2005	827	29/12/2005	38	☐	2005763
7	14/08/2006	1321	28/05/2007	41	☐	2007482
8	15/08/2006	1322	15/05/2007	39	☐	2007427
9	22/09/2006	1360	11/05/2007	33	☐	2007402
10	26/10/2006	1394	19/07/2007	38	☑	2007233
11	17/11/2006	1416	10/08/2007	38	☐	2007652
12	07/12/2006	1436	23/08/2007	37	☐	2007244

Lambda:	No cases:	expected:	Start case:	Start date:	End date:	p value:	cluster group:	Type
2324.46	6	0.9	7	14/08/2006	07/12/2006	0.021	1	Cluster

Distribution of cases

* = cases with gestation entered, ? = cases with estimated gestation

Tick marks for the 1st of each month.
Solid line represents span of most significant cluster, dotted lines indicate span of cases in clusters of the same group.