Drug Saf (2013) 36:995–1006
DOI 10.1007/s40264-013-0061-7

ORIGINAL RESEARCH ARTICLE

Illustration of the Weibull Shape Parameter Signal Detection Tool Using Electronic Healthcare Record Data

Odile Sauzet · Alfonso Carvajal · Antonio Escudero · Mariam Molokhia · Victoria R. Cornelius

Published online: 15 May 2013

Abstract
Background The WSP tool has previously been proposed as a method to detect signals for adverse drug reactions utilising time-to-event data without the need for a reference population. The aim of this study was to assess the performance of the tool on two well-known and two suspected adverse drug reactions for bisphosphonates that varied in both frequency and accuracy of reporting time.
Methods The use of the WSP tool was investigated on data from a matched population cohort study involving data from UK primary care patients exposed to oral bisphosphonates. Four listed/suspected ADRs were selected for investigation: headache, musculoskeletal pain, alopecia and carpal tunnel syndrome. For each suspected ADR, a graphical exploratory analysis was performed and the WSP tool was applied for two censoring periods each.
Results Both of the well-known and common ADRs (headache and musculoskeletal pain) were detected using the WSP tool, and the signals were present regardless of the censoring intervals used. A signal was also detected when the event was uncommon and the timing was likely to be an accurate reflection of onset time (alopecia). This signal was only present for some of the censoring intervals. As anticipated, no signals were raised in the control groups for these events regardless of the censoring interval used. The suspected ADR, which was uncommon and where reporting times may not reflect onset time accurately (carpal tunnel syndrome), was not detected. A signal was raised in the control group but its false-positive nature was visible in the exploratory graphical analysis, which led to it (frequent but for only a limited number of consecutive dates).
Conclusion This study illustrates the usability and examines the reliability of the WSP tool as a method for signal detection in electronic health records. When the events are uncommon the success of this method may depend on the reporting time accurately reflecting the true event onset time. The study has shown that further work is required to define the censoring periods. The addition of a control group is not required but may enhance causal inference by showing that other causes than the exposure may lead to a signal.

1 Introduction

Spontaneous reports (SR) of suspected adverse drug reactions have been widely used for post-market monitoring of drug harms, but increasingly electronic healthcare records (EHRs) are being proposed as a resource for routine drug surveillance [1–3]. These records provide valuable sources of longitudinal drug cohort data for all patients and can be used to examine patterns of adverse events. EHRs have key advantages over SR, as they contain population-based data including more comprehensive information on exposure for all patients taking the drug and not just the subset of patients where the suspected adverse drug reaction has been reported. As a consequence linked EHRs (potentially with larger

O. Sauzet (✉)
AG Epidemiologie and International Public Health, Universität Bielefeld, Pf 10 01 31, 33501 Bielefeld, Germany
e-mail: odile.sauzet@uni-bielefeld.de

A. Carvajal · A. Escudero
Department of Pharmacology, Faculty of Medicine, University of Valladolid, Valladolid, Spain

M. Molokhia · V. R. Cornelius
Department of Primary Care and Public Health Sciences, King's College, London, UK

numbers of events) maybe more suitable to detect signals for rare adverse drug reactions [4].

Cornelius et al. [5] have recently proposed a method of signal detection based on the analysis of time-to-event data: the Weibull Shape Parameter tool. It is a method of signal detection for the possible presence of adverse drug reactions when time-to-event data are available and does not require a reference (control) population.

If only background adverse events (not due to the drug) are present, it can be expected these events occur uniformly along the observation time, i.e. no temporal pattern is visible during the observation period. However, if there is a time window following the beginning of treatment in which there is an increased risk of having the event, then this could be an indication that there is a temporal causal relationship, linking drug and event. The WSP tool will raise a signal if there is statistical evidence that the events do not occur uniformly during the observation period. Both WHO [6] and Naranjo [7] used criteria that examine the temporal relation between the drug and adverse event to assign causality, but have limitations [8].

The power of the WSP tool has been evaluated using simulations [5]. In order to examine the validity of the WSP tool in a real setting, we analysed data from a matched cohort study involving patients in the UK primary care data: THIN database [9] exposed to oral bisphosphonates, which are commonly used in the treatment of osteoporosis. Four events known or suspected to be ADRs with different expected incidence rates were selected for an illustrative investigation: headache, musculoskeletal pain, alopecia and carpal tunnel syndrome (CTS). Headache and musculoskeletal pain are well recognised side effects (1–10 %) of bisphosphonates; however, rarer events such as alopecia and carpal tunnel syndrome (1 %) are less well studied. The aim of this study is to explore the use of the WSP tool in practice on a real data set and to provide guidelines on the methodology to fit the tool and interpret its results.

2 Methods

2.1 Statistical Analysis of Time-to-Event Data

Time-to-event data are available when a group of patients are observed for a certain duration and the timing of the occurrence of a particular event is recorded. The time-to-event data consists of two parameters; the first one records the presence of an event (typically 1 if an event occurred and 0 if not) and a time-to-onset parameter. This time to onset is calculated from the beginning of the observation (e.g. time of a subject's first prescription) to either the occurrence of the event or the end of the subject's observation time. The subjects' observation time will either be the end of the defined study period (e.g. 1 year) or the time at which the exposure to the drug stops or the patient is lost to follow-up. Therefore, after an event has been reported, a patient is no longer observed. When the end of the observation period for a participant is earlier than the defined study period, then the observation is said to be censored.

Statistical methods for time-to-event data, on which the WSP test is based, are models for instantaneous failure (or hazard) rates (rates of events over a given unit of observation time). This is an instantaneous measure of the relative number of events to the actual number of subjects observed. The actual object modelled related to instantaneous rates is the hazard function, which is the limit (as the time unit of observation becomes smaller) of the probability of experiencing an event during a small time interval, given that no events were experienced until the beginning of the interval.

A constant hazard function over time equates with an equal 'risk' of having an event over all of the observation period. However, if the hazard function varies over time then the 'risk' of having an event is time dependent. With regard to monitoring drug safety, we infer the 'risk' of having an event has a temporal relationship to the beginning of the observation, usually defined as the start of treatment. The use of the Weibull method means that there is limited need to adjust the model for confounding factors, as even if the cohort contained a disproportional number of 'higher risk' patients and therefore had an increased number of events of interest, we would still expect the occurrence of these events to be evenly distributed over time (so long as they are not caused by the drug).

2.2 WSP Tool for Time-to-Event Data

The WSP tool is based on the WSP test, which consists of fitting a Weibull model to time-to-event data. The Weibull distribution for time-to-event data has a shape parameter that, when different from the value one, indicates that the hazard is not constant, i.e. that the hazard (or risk) of having an event varies in function of the time-to-onset from the beginning of treatment. A statistical test with a null hypothesis that the shape parameter is equal to one exists and can be used to test the possibility that there is a temporal causal relationship between the beginning of treatment and event. A p value less than a specified significance level (say 0.05) obtained after fitting a Weibull model to the time-to-event data will be an indication that the hazard function is not constant and that the occurrence of the event is temporally connected to the date of first exposure.

Simulations have shown that, when the increased risk appears towards the middle of the observation period, the

hazard function has a symmetric shape. In such cases, the WSP test does not detect a variation in hazard easily [5]. The WSP tool addresses this by censoring the data set at selected time points and applying the WSP test at each time point. If for one of these periods a significant shape parameter is obtained (p value for the shape parameter is less than the defined significance level, e.g. 0.05), then a signal is raised indicating the possible presence of ADRs. In practice, one decides the censoring period, for example 1 month, and the WSP test is applied to the data censored at 1 month, then 2 months, 3 months, etc., until the end of the observation period. For each execution of the WSP test, a p value is returned and a signal is raised if at least one of them is less than the specified significance level. The WSP tool is implemented with the R package Survival [10].

2.3 Role of the Control Group in this Study

All the chosen adverse events included in this study occur in the general population (background events) independent of any relation with the drug. The aim of the WSP tool is to distinguish between the two sets of events (background events vs. true ADR) in the exposed population. In the control group, all events occurring are considered to be background events and we expect to observe a constant rate of events over time.

The WSP tool has been developed for use in data sets for which no control group is available. The role of the control group in this study is to observe the behaviour of the WSP tool when no relationship is to be expected.

2.4 Exploratory Analysis of Time-to-Event Data

Exploratory analyses allow the examination of distribution of events and risk (ratio of events to the number of observation) and the assessment of the constancy of the hazard over the observation period. We suggest three plots for the exploratory analysis.

(1) Histogram: a histogram of the time-to-event data will show the distribution (absolute number of events and the time of their occurrence) independent of the number of participants at the time the events occur.

(2) Instantaneous hazard rates: the instantaneous hazard rate is defined as the rate of events in a defined observation period starting at time t, conditional upon the patients not having the event until time t. The observation period is divided in time units and a rate is computed for each of these units. It is a way of measuring the variability in the rate of events for various lengths of time periods.

(3) Hazard function: smoothed estimate of the hazard function, which is derived from the data. The estimates are obtained with the R function *muhaz* (package of the same name) [11].

2.5 Observation Time and Censoring Periods

The observation time for the occurrence of an ADR related to a drug will be determined by the nature of the ADR and with consideration of the underlying causal mechanism, taking into account failing to detect an ADR and avoiding false positives. For example, a headache occurring 3 years after the first administration of the drug is unlikely to be related to this first administration. Examples of immediate, intermediate and delayed ADRs include anaphylaxis (seconds–minutes), skin rash (hours–days) and liver toxicity (a few days to several months). The censoring period defines the dates at which the data will be censored during the observation time. A censoring period that is too small (the WSP test will be performed on the data many times) will increase the risk of obtaining a false-positive result and a censoring period that is too large (the WSP test will be performed on the data few times) will increase the risk of missing a true ADR. For each event two censoring periods are presented based on the clinical nature of the event. This will illustrate how the duration of the censoring periods affects the results of the WSP tool.

2.6 Data set

The Health Improvement Network (THIN) is a database containing the electronic healthcare records of over 550 UK general practices covering almost 6 % of the UK population. The records started in 2002 and provide longitudinal prescribing and reported event data for each patient of the participating practices [9].

The data set used in this study consists of female patients in the THIN data set who have been prescribed bisphosphonates for the treatment of osteoporosis. Exclusion criteria were history of malignant cancer or Paget's disease. From each individual patient all the relevant medical information from primary and secondary care, as well as additional information, including demographics and drug prescriptions, is recorded in the health care medical record. Each of these patients were matched on age (year of birth) and practice (to control for social background) with two women unexposed to bisphosphonates. The index date for each patient is the date of first prescription of a bisphosphonate. Controls and exposed subjects are matched on the index date to allow for similar follow-up times. Unless the patient died or left the practice, all observations ended in August 2008. The READ dictionary was used to identify diagnostic and symptom codes for selected events, and British National Formulary (BNF) codes were used to select drugs of interest.

Patients prescribed one of the following drugs—alendronate, etidronate, ibandronic acid, palmidronate,

zolidronic acid, risedronate, tiludronate and clodronate—were included from the start of their first prescription. Patients who were newly registered at a practise had a 3-month window following registration to allow sufficient time for practice data entry of baseline characteristics and existing medical information prior to bisphosphonate prescription. Events on the index date (first prescription date in exposed) were excluded. Any event recorded after the index date is included in the analysis. To include a lag time would be mere guess work and we would rather not include an arbitrary lag time for all events. The end date was defined by the earliest of transfer out of the practice date, death date or last data collection date. A cohort of patients who newly started bisphosphonates was identified.

Events were selected to illustrate the performance of the WSP tool for events with varying incidence rates and reporting pattern by patients. These were the only events analysed for this study and therefore what is presented may not represent patterns more generally observed. Four events were examined: headache, musculoskeletal pain, alopecia and CTS. For headache and musculoskeletal pain, an observation time was limited to 1 year and the WSP tool is applied for censoring periods (choice related to the nature of the events) of 3 and 15 days and of 7 and 15 days respectively. For alopecia and CTS an observation time of 3 years was considered, and the results for censoring periods of 1 and 3 months are presented.

2.7 Presentation of the Work

In Sect. 3, for each adverse event, the data are presented in a three-step analysis process:

(1) The absolute number of events, rates for person's years observed by groups and rate of events in the exposed group as a percentage increase of the rate of background cases (control group) are provided.

(2) The data are explored using the three plots suggested above.

(3) The results of the WSP tool are described for the whole observation period and presented in tables only for the first ten censoring periods.

3 Results

There were 19,817 exposed patients for 39,634 unexposed controls (2 controls for each exposed patient matched on age). The mean age at the time of first prescription was 72.7 years, ranging from 52 to 103 years, and the median was 73.3 years. In this data set, 76 % of observations are of 1 year or more and 36 % of 3 years or more. The median number of prescriptions is 5 (range 1–26).

3.1 Headache

3.1.1 Rate of Events

A total of 581 cases of first occurrence of headache are reported among 19,817 exposed patients and 885 among 39,634 controls. The incidence rate of events per year of observation is 11.2 ‰ person year for exposed and 8.5 ‰ person year for control (background rate). The incidence rate of event per year for the exposed group corresponds to a 32 % increase of the rate per year of the exposed group.

3.1.2 Exploratory Analysis

Figure 1 shows the exploratory plots for the headache data for the exposed group (plots a.1, b.1, c.1) and the control group (plots a.2, b.2, c.2). Histogram a.1 indicates an increased number of cases of headache between 2 weeks and 3 months after prescription of the drug. Histogram a.2 for the control data shows a more regular occurrence of events with occasional peaks. The reduction of cases later in the observation period is in part due the reduction of the number of patients being observed at this time. For the control group the number of cases decreases proportionally with the number of patients under observation over time.

Plots b.1 and b.2 represent the instantaneous hazard rates for the headache data for time windows of 3 and 15 days for an observation period of 1 year. Increased hazard rates for the exposed group (b.1) are visible until 3 months of observation and also at around 7 months then the shapes of the plots for exposed and controls are similar.

Plots c.1 and c.2 represent the smoothed estimated hazard function for both groups. There is a decrease in hazard for both. A slightly increased risk is also visible for the control group at the beginning of the observation period. Whether the hazard is significantly decreasing is tested by the WSP tool.

3.1.3 Applying the WSP Tool

The results of the application of the WSP tool (given as p values) for censoring periods of 3 and 15 days are each reported in Table 1 for ten censoring intervals covering respectively a period of 33 and 150 days.

A signal is raised for headache in the exposed group with all p values being smaller than 0.05 between 15 and 24 days of follow-up and then all p value are above 0.05. This means that there is some evidence that the hazard function for all censored data is non-constant. For the control group a signal is raised for a censoring period of 3 days for one censoring time only as all other p values are greater than 0.05. For the whole observation period (results not shown here), no other p value under 0.05 was obtained.

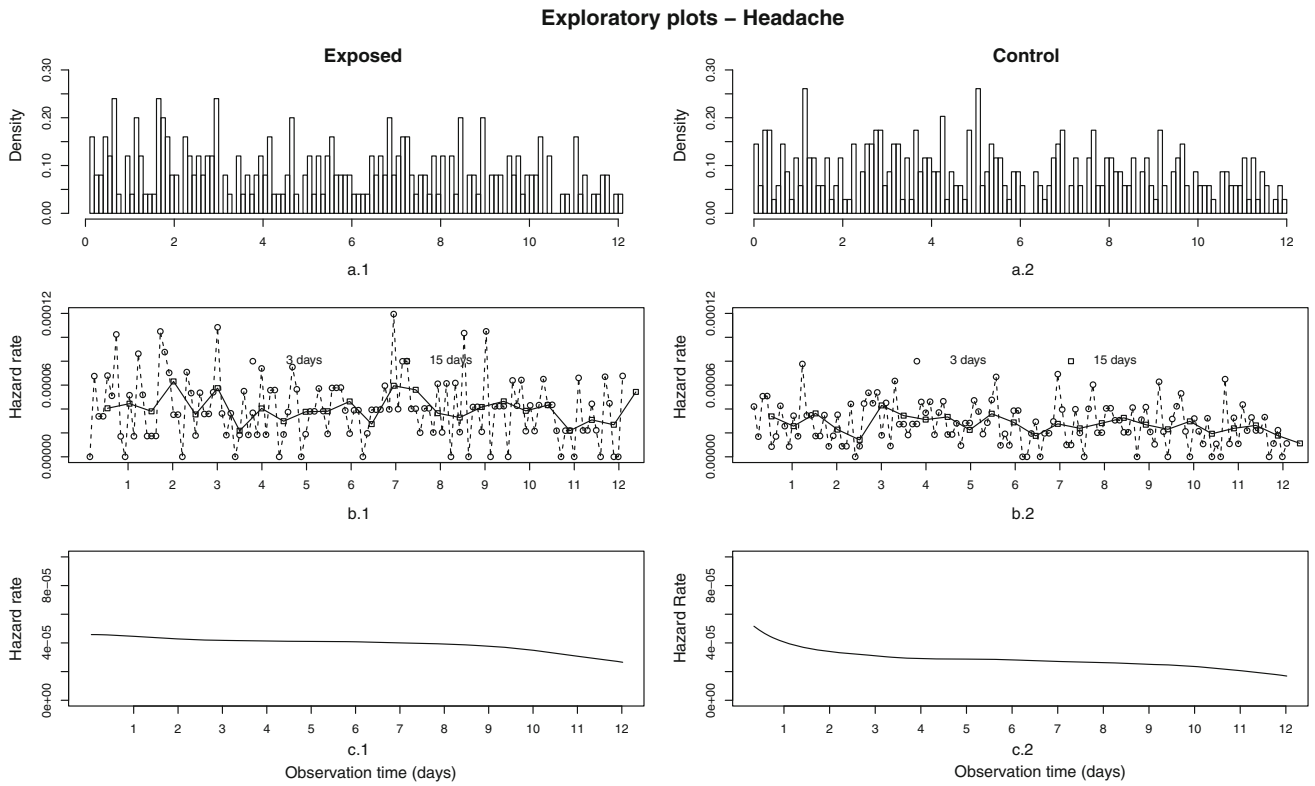

Fig. 1 Graphs for the exploratory analysis of headache over a year with bisphosphonates (month): *a.1–a.2* Histograms of cases; *b.1–b.2* instant hazard rates; *c1–c2* estimated hazard functions

Table 1 Results of the Weibull shape parameter tool applied to the headache data using censoring intervals of 3 and 15 days

Headache: *p* values for the Weibull shape parameter										
Censoring period of 3 days										
Censoring (days)	6	9	12	15	18	21	24	27	30	33
Exposed	<0.01	0.04	0.07	0.03	0.04	0.04	0.04	0.19	0.17	0.32
Control	0.35	0.08	0.04	0.29	0.55	0.35	0.43	0.77	0.67	0.85
Censoring period of 15 days										
Censoring (days)	15	30	45	60	75	90	105	120	135	150
Exposed	0.03	0.17	0.38	0.12	0.30	0.17	0.58	0.65	0.92	0.991
Control	0.89	0.69	0.94	0.66	0.22	0.75	0.96	0.98	0.88	0.87

3.2 Musculoskeletal Pain

3.2.1 Rate of Events

After removal of cases recorded on the index date, a total of 914 cases of musculoskeletal pain are reported among 19,817 exposed patients and 923 among 39,634 controls. The incidence rate of events is 18.4 ‰ person year for exposed and 8.9 ‰ person year for control (background rate). The incidence rate of event per year for the exposed group corresponds to a 108 % increase of the rate per year of the exposed group.

3.2.2 Exploratory Analysis

Figure 2 shows the exploratory plots for the musculoskeletal pain data for the exposed group (plots a.1, b.1, c.1) and the control group (plots a.2, b.2, c.2).

Histogram a.1 indicates an higher number of cases of musculoskeletal pain on the first few months after

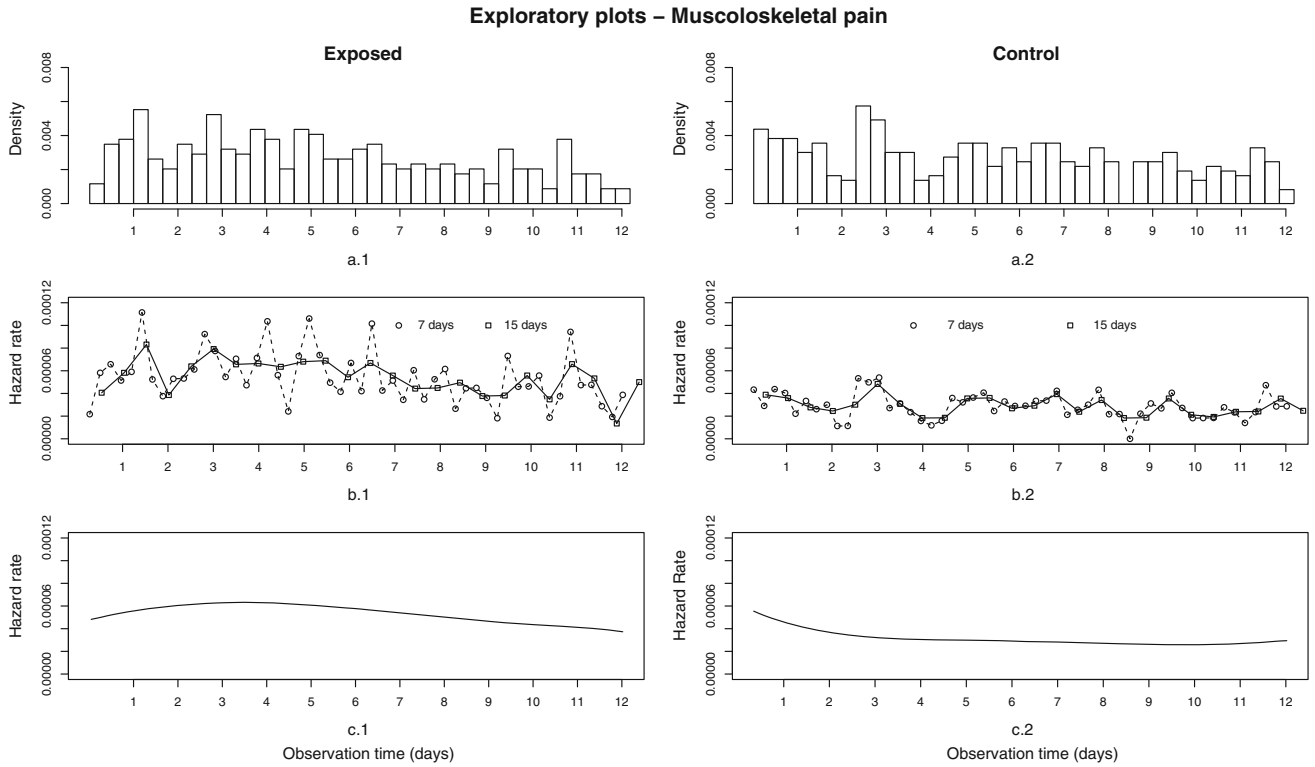

Fig. 2 Graphs for the exploratory analysis of musculoskeletal pain over a year with bisphosphonates (month): *a.1-a.2* Histograms of cases; *b.1–b.2* instant hazard rates; *c1–c2* estimated hazard functions

Table 2 Results of the application of the Weibull shape parameter tool for censoring periods of 7 and 15 days for musculoskeletal pain data

Musculoskeletal pain: p values for the Weibull shape parameter Censoring period of 7 days										
Censoring (days)	7	14	21	28	35	42	49	56	63	70
Exposed	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	0.01	0.03	0.04
Control	0.01	0.17	0.11	0.15	0.49	0.53	0.75	0.86	0.66	0.35
Censoring period of 15 days										
Censoring (days)	15	30	45	60	75	90	105	120	135	150
Exposed	<0.01	<0.01	0.01	0.04	0.04	0.01	0.01	0.01	0.02	0.02
Control	0.08	0.28	0.73	0.79	0.82	0.51	0.64	0.84	0.47	0.64

prescription of the drug (this higher number is not present on histogram a.2 for the control data). The reduction of cases later in the observation period is in part due the reduction of the number of patients being observed at this time. For the control group the number of cases decreases proportionally with the number of patients under observation over time.

Plots b.1 and b.2 represent the instantaneous hazard rates for the musculoskeletal pain data for time windows of 7 and 15 days for an observation period of 1 year. Increased hazard rates for the exposed group (b.1) are visible until 7 months of observation, then the shapes of the plots for exposed and controls are similar.

Plots c.1 and c.2 represent the smoothed estimated hazard function for both groups. A very clear increased hazard is visible for the exposed group from the very beginning of the observation time and then decreases during the whole of the observation period (1 year) to reach the hazard level of the control group.

3.2.3 Applying the WSP Tool

The results of the application of the WSP tool (given as p values) for the censoring period of 7 and 15 days are each reported in Table 2 for ten censoring intervals covering respectively a period of 70 and 150 days.

A clear and unambiguous signal for musculoskeletal pain is raised for the exposed group with all p values being less than 0.05. This means that there is strong evidence that the hazard function for all censored data is non-constant. No signal is raised for the control group as all p values are greater than 0.05 apart for the data censored at 7 days (see Sect. 4). The same results were found for the whole observation period (results not shown here).

3.3 Alopecia

3.3.1 Rate of Events

A total of 172 cases of alopecia are reported among 19,817 exposed patients and 227 among 39,634 controls. The incidence rate of events is 3.4 ‰ person year for exposed and 2.2 ‰ person year for control subjects (background rate). The incidence rate of event per year for the exposed group corresponds to a 57 % increase of the rate per year of the exposed group.

3.3.2 Exploratory Analysis

Figure 3 shows the exploratory plots for the CTS data for the exposed group (plots a.1, b.1, c.1) and the control group (plots a.2, b.2, c.2).

Histograms a.1 and a.2 indicate that the cases of alopecia are reported irregularly over the time periods in the histograms, which is to be expected with rare cases. However, the exposed group shows an increased number of reported cases at the beginning of the observation period then between the 5th and 13th months. Plots b.1 and b.2 represent the instantaneous hazard rates for the time windows of 30 and 91 days for an observation period of 3 years. Those plots confirm the presence of an increased risk between the 5th and 13th months of observation for the exposed group. Plots c.1 and c.2 represent the smoothed estimated hazard function for both groups. A constant hazard is visible for the exposed group until the 14th month of observation followed by a decrease.

3.3.3 Application of the WSP Tool

The results of the application of the WSP tool (given as p values) for censoring periods of 30 and 91 days are each reported in Table 3 for ten censoring intervals covering respectively a period of 330 and 910 days. A signal is raised for alopecia in the exposed group for the censoring intervals ending at 270, 300 and 330 for the censoring period of 30 days as well as for 360 and 390 (not reported in Table 3) and at 273 and 364 for the censoring period of 91 days. No signal is raised for the control group. No other

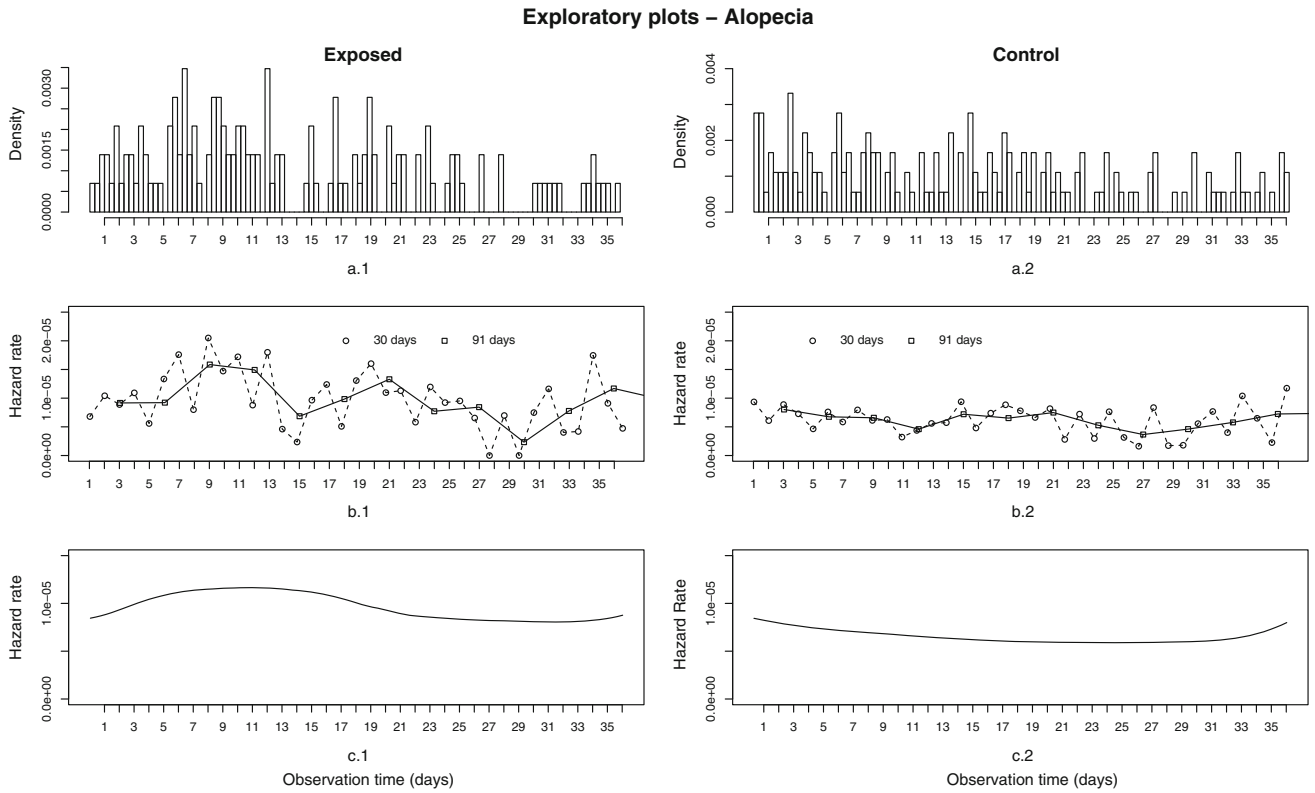

Fig. 3 Graphs for the exploratory analysis of alopecia over 3 years with bisphosphonates (month): *a.1–a.2* histograms of cases; *b.1–b.2* instant hazard rates; *c1–c2* estimated hazard functions

Table 3 Results of the Weibull shape parameter tool applied to alopecia for censoring periods of 30 and 91 days

Alopecia: *p* values for the Weibull shape parameter Censoring period of 30 days										
Censoring (days)	60	90	120	150	180	210	240	270	300	330
Exposed	0.24	0.32	0.25	0.53	0.28	0.09	0.17	0.04	0.03	0.02
Control	0.65	0.92	0.88	0.56	0.69	0.59	0.70	0.63	0.58	0.37
Censoring period of 91 days										
Censoring (days)	91	182	273	364	455	546	637	728	819	910
Exposed	0.21	0.31	0.03	0.02	0.18	0.29	0.21	0.41	0.59	0.89
Control	0.87	0.74	0.65	0.29	0.45	0.48	0.62	0.48	0.26	0.19

signal is raised for the rest of the observation period for the two censoring periods (results not shown).

3.4 Carpal Tunnel Syndrome

3.4.1 Rate of Events

A total of 233 cases of CTS are reported among 19,817 exposed patients and 321 among 39,634 controls. The incidence rate of events is 4.5 ‰ person-years for exposed and 3.1 ‰ person-years for control subjects (background rate). The incidence rate of event per year for the exposed group corresponds to a 46 % increase of the rate per year of the exposed group.

3.4.2 Exploratory Analysis

Figure 4 shows the exploratory plots for the CTS data for the exposed group (plots a.1, b.1, c.1) and the control group (plots a.2, b.2, c.2).

Histograms a.1 and a.2 indicate that the cases of CTS are reported irregularly over the time periods in the histograms. However, the exposed group shows an increased number of reported cases until the 400th day. Plots b.1 and b.2 represent the instantaneous hazard rates for the for time windows of 30 and 91 days for an observation period of 3 years. Those plots confirm the presence of an increased risk of the 13th month of observation for the exposed group. Plots c.1 and c.2 represent the smoothed estimated hazard function for both groups. A decreasing hazard is visible for the exposed group over the 3 years of observations and reaches the level of the control group.

3.4.3 Application of the WSP Tool

The results of the application of the WSP tool (given as *p* values) for censoring periods of 30 and 91 days are each reported in Table 4 for ten censoring intervals covering a period of 330 and 910 days, respectively. No signal is raised for CTS in the exposed group and a signal is raised for the control group at 120 and 150 days for the censoring period of 30 and 637 days for the censoring period of 91 day. Not shown here is that for the 30-day censoring period a signal is obtained at 570 days and then another at 630 days for the control group.

4 Discussion

4.1 Summary of Findings

The WSP tool has been applied to four adverse events related to the use of bisphosphonates prescribed for the prevention of osteoporosis. Two of them, headache and musculoskeletal pain, are well-known side effects of bisphosphonates (described in Summary Product Characteristics, SPC); there is some evidence that carpal tunnel syndrome could be associated to bisphosphonates [12, 13], but no published evidence of bisphosphonate association with alopecia (included in the SPC only).

An increased risk of headache was detected using a very short censoring period (3 days) and was detectable until 24 days after the first prescription. Headache may be a common event but is rarely reported unless extreme or an unrelated visit to the doctor occurs. In spite of this the WSP tool did raise a signal.

When an increased risk of an adverse event occurs at the beginning of the observation period, it seems to be easily detected as for musculoskeletal pain, giving a signal for all censoring intervals provided they include the period of increased risk.

Alopecia is a rare condition (3.4 ‰ persons year for the exposed group). A small indication of an increased hazard between the 5th and 14th months of observation was shown by the instantaneous hazard rates, contrary to the estimated hazard function, which showed only a decrease in hazard with a sharper slope between after 15 months of observation, not showing the increased risk occurring just before

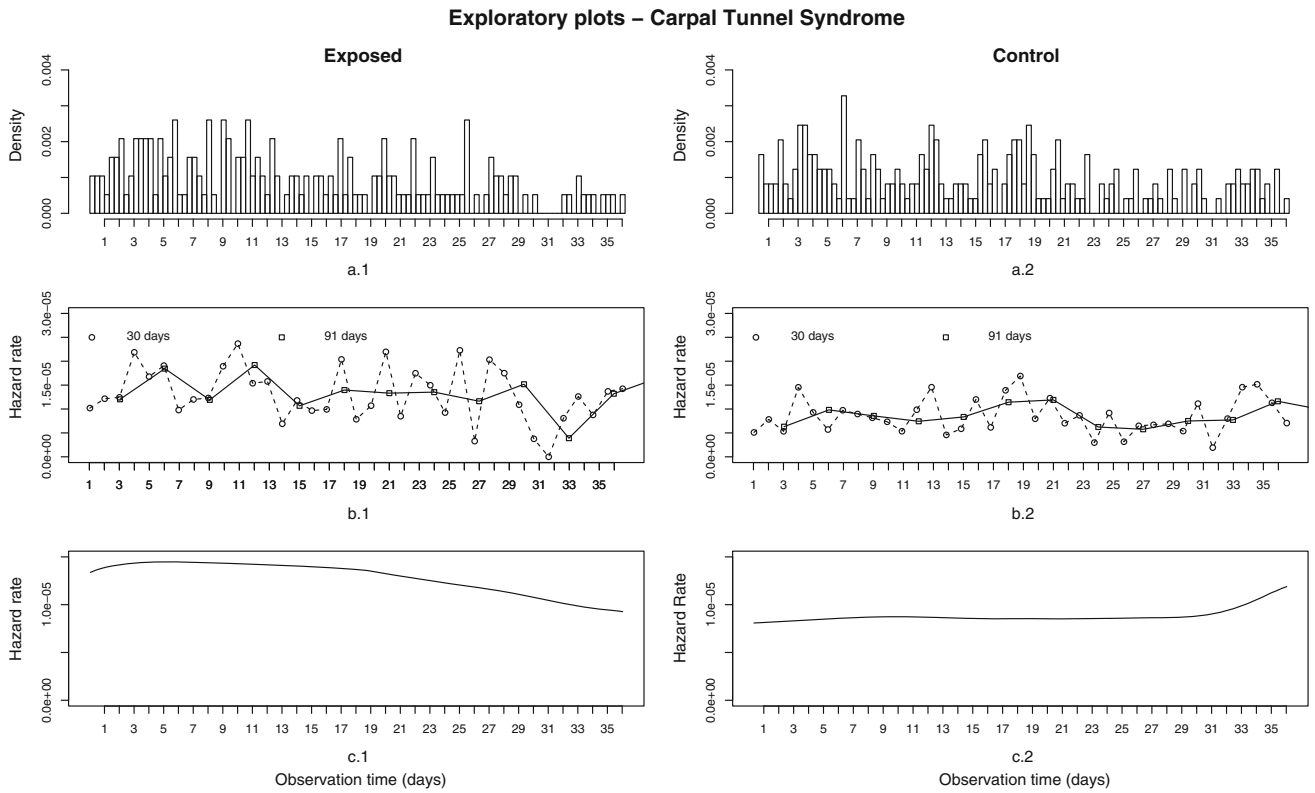

Fig. 4 Graphs for the exploratory analysis of carpal tunnel syndrome over 3 years with bisphosphonates (month): *a.1–a.2* histograms of cases; *b.1–b.2* instant hazard rates; *c1–c2* estimated hazard functions

Table 4 Results of the Weibull shape parameter tool applied to the carpal tunnel syndrome for censoring periods of 30 and 91 days

Carpal tunnel syndrome: p values for the Weibull shape parameter Censoring period of 30 days										
Censoring (days)	60	90	120	150	180	210	240	270	300	330
Exposed	0.50	0.57	0.14	0.12	0.08	0.22	0.33	0.44	0.33	0.16
Control	0.08	0.25	0.01	0.02	0.09	0.08	0.09	0.12	0.17	0.31
Censoring period of 91 days										
Censoring (days)	91	182	273	364	455	546	637	728	819	910
Exposed	0.54	0.10	0.42	0.17	0.51	0.57	0.65	0.71	0.88	0.83
Control	0.18	0.08	0.15	0.29	0.35	0.12	0.04	0.14	0.33	0.42

the 5th month. The WSP tool has been able to give a signal for a non-constant hazard for censoring times ending between 9 and 13 months of observation (between 270 and 400 days) corresponding to what is observed in the plots for instantaneous hazard rates. This suggests that there is a delayed temporal relationship between the beginning of observation (first prescription) and the occurrence of alopecia.

Carpal tunnel syndrome (CTS) is also an uncommon event (4.5 ‰ persons-years for the exposed group), and the observation of hazard plots was unable to identify a period of increased risk using the WSP tool (although this cannot be excluded, which will be discussed later).

4.2 Methodology to Use the WSP Tool

4.2.1 Exploratory Analysis

Using multiple plots relative to time-to-event data for adverse events allows the display of the temporal distribution of hazard and the localisation of a period of increased hazard. A histogram of time to events displays

the time of occurrence of events and can demonstrate periods for which there is an absolute increase in the numbers of cases. However, this plot does not allow for a global consideration over the complete observation period as it is a representation of the absolute numbers of cases at a given time interval with no relation to the number of observations. Also, if the number of cases is very small this histogram may be uninformative as there will be too few cases in each plotting interval for meaningful interpretation. If the duration of the observation period is large in comparison with the duration of increased risk—as was the case for headache—then a histogram might not be very informative.

Instantaneous hazard rates reflect the hazard function, which is the limit of these hazard rates when the duration of the time units tends to zero. Despite showing a great amount of variability when there are few events in each observation period, they may suggest times of increased risk. Smaller observation units seem to better reflect the presence of an increased risk for a shorter period of time, but the plot will be less smooth than with longer periods because of the variability in the number of cases during shorter units.

Estimated hazard function: The WSP tool actually tests the "flatness" of the hazard function, i.e. it will return a high p value if it is unlikely that the hazard is constant over the observation period. However, the appearance can vary greatly with the parameters chosen for the estimation of the plotted function and it can be misleading if the hazard function is 'over smoothed'. It is common for the estimated hazard function to show a slight increase of risk at the beginning of the observation period, which does not reflect the data. Displaying confidence intervals for the hazard plots will inform on the precision of the estimate.

4.2.2 Censoring Period

The censoring period needs to be determined according to two principles: avoiding false negatives (for example for selecting a period that is too wide) and avoiding false positives by detecting short-lived increases in risk due the variability in the data (with too short an interval). It would be advantageous if the decision about the choice of a censoring period is informed by the biological mechanism leading to the event under consideration. A common and well-reported ADR appearing a few hours after the beginning of treatment (e.g. skin rash) can be detected with a broad (any sort of) censoring interval (the WSP test should be sufficient). An ADR due to cumulative drug effects will appear later with greater variability and will require a longer censoring period than an ADR, which would be a consequence of just one dose of medication.

For the alopecia data, the mechanism is currently not well understood, but it is unlikely to occur soon after taking one dose of medication, possibly because of an immune-mediated effect. Therefore, it would not be very meaningful to take a censoring period of just a few days as the increase in risk happens much later. In this case, the actual choice of a censoring period does not have an effect on the results within a range of possible periods (as long as they include the risk period), with an apparent increased risk between 9 and 13 months.

4.2.3 False Positives

A p value under 0.05 was obtained for musculoskeletal pain when the data were censored at 7 days. Simulations have shown that when a censoring point occurs soon after the beginning of the observation period, a non-constant hazard is very often falsely reported because of the low number of cases. Therefore, it might be preferable to start censoring after a reasonable duration.

Simulations [5] have shown that for rare events false-positive signals are detected in data sets between 7 and 12 % of the time. A false positive (because obtained for the control group) was found in CTS as well as for headache. The distribution of time to events has shown irregularity in the number of reported cases in each period presented in the histogram (in the case of CTS seasonal distribution is a possibility) [14]. This can be explained by the small number of cases and other unobserved factors. Here both groups had a similar pattern of time-to-event distributions but the exposed group was half the size of the control group and therefore had fewer events. The signals obtained are short (a signal is obtained by censoring at 4–5 months and later at 22 months of observation for CTS and just at 12 days for a 3-day censoring period) and obtained when a sudden small increase in risk is seen at the very end of the censored period but the signal is lost very soon when the censoring point increases. The likely reason for those increases is that the few cases are not uniformly reported over time. For exposed groups, there was no signal, which could be suspected as a false-positive result due to the brevity of the signal (however, this could also be due to other factors related to the condition).

When a signal is raised, the possibility of a false positive should always be investigated. The WSP tool primarily tests for increased risk. Methodology for assessing whether a signal is likely to be due to the drug needs to be included in the analysis, based on additional epidemiological and ADR causality criteria [6, 7]. The exploratory analysis should give an indication that cases are reported regularly or that there are short-lived sudden increases. Here the first three events have reported cases rather regularly whereas the last one seems to be erratic and might be due to random

error. There may not be any reason to explain that an increased risk is for a very short period of time or that increased risk appears at several intervals. There are currently a number of bodies examining potential ADR signals for further monitoring/validation on a broad level: the WHO monitoring centre in Uppsala [15], MHRA UK [16], FDA Sentinel [17], EMA [18] and European Network of Centres for Pharmacoepidemiology and Pharmacovigilance [19], in addition to academic and industry collaborator networks for drug safety studies.

4.2.4 Interpretation of Results

A signal is raised if, for at least one censoring point, a *p* value under 0.05 is obtained [5]. If there is a true temporal relationship between prescription and events then the non-constant hazard should be detected during a sufficiently large duration (relative to the biological mechanism); otherwise, this is likely to be a false positive.

If a control group is available and for both groups, *p* values under 0.05 are obtained, then this is an indication that the signal could be due to the irregularity of background case reporting (creating local increases in risk). A signal obtained for several consecutive censoring points in the exposed group and just for one of those in the control group can be interpreted as a temporal relationship between prescription and events where there are indications that background events are irregularly reported.

4.3 Limitations and Further Work

4.3.1 Limitations of the Study

This study is an illustration of the use of the WSP tool in practice with an in-depth look at graphical summary methods. Four pre-selected events and one type of drug class were considered; therefore, only limited conclusions can be drawn about the validity of the method. All patients that have been prescribed once are considered at risk during the whole observation period, even if they did not take the drug or took the drug for a limited time, which might either artificially produce a reduction of risk over time and lead to a false-positive result or hide an increased risk later in the observation period and lead to a false-negative result. The controls in the data set used are unexposed to the conditions that led to the prescription; therefore, this group had only a limited role in assessing causality.

4.3.2 Limitations of the WSP Tool

There is some evidence that carpal tunnel syndrome is associated with bisphosphonates [12, 13]. The WSP tool did not detect an association. One reason for this might be the lack of power as carpal tunnel syndrome is a rare condition. Even through the risk was twice as great as that for the control group, the pattern of reporting was constant over time. It may be that the mechanism is not time dependent or the previous associations are due to confounding by indication [20] as CTS is associated with menopause [21]. This first application of the WSP tool in practice has highlighted a dependence of the results on the choice of the censoring periods leading to issues of multiple testing. These issues can be resolved by further development of the method (see below).

4.3.3 Further Development

The WSP tool needs to be validated on a large unselected set of events for a variety of drugs and compared with other signal detection methods currently in use. This large case study can incorporate solutions for a methodology to define the censoring periods. Further work is planned to include adequate controls for multiple testing and introducing Bayesian methods with informative priors to estimate the shape parameter of the Weibull distribution.

5 Conclusion

This study is the first to examine the usability and reliability of the WSP tool as a method for signal detection in electronic health records. When the events are uncommon the success of this method may depend on the reporting time accurately reflecting the true event onset time. The addition of a control group exposed to the conditions leading to prescription is not required but may enhance causal inference. Graphical methods suggested in this paper provide a means to further investigate the robustness of an ADR signal that has been raised and proves a valuable adjunct to assess drug safety. This study also highlighted the dependence of the results upon the choice of censoring periods. Future research is planned to investigate optimal methods for detection that are not reliant on the user or event. The WSP tool can now be validated in a large-scale study comparing existing methods and incorporating messages learned in the present study.

Acknowledgments V. Cornelius and M. Molokhia were supported by the National Institute for Health Research (NIHR) Biomedical Research Centre at Guy's and St. Thomas' NHS Foundation Trust and King's College London. The views expressed are those of these authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Dr. Molokhia has received grants from AstraZeneca, Pfizer and the Serious Adverse Events Consortium (SAEC; collaboration of academia and industry) for studies on drug safety. O. Sauzet, A. Carvajal, A. Escudero and V.R. Cornelius have no other conflict of interest to declare.

The authors would like to thank the reviewers of this paper for their insightful comments, which have greatly improved its quality.

References

1. Coloma PM, Trifiro G, Schuemie MJ, et al. Electronic healthcare databases for active drug safety surveillance: is there enough leverage? Pharmacoepidemiol Drug Saf. 2012; 21(6):611–21.
2. Johansson S, Wallander MA, de Abajo FJ, et al. Prospective drug safety monitoring using the UK primary-care General Practice Research Database: theoretical framework, feasibility analysis and extrapolation to future scenarios. Drug Saf. 2010;33(3): 223–32.
3. Park MY, Yoon D, Lee K, et al. A novel algorithm for detection of adverse drug reaction signals using a hospital electronic medical record database. Pharmacoepidemiol Drug Saf. 2011; 20(6):598–607.
4. Schuemie MJ, Coloma PM, Straatman H, et al. Using electronic health care records for drug safety signal detection: a comparative evaluation of statistical methods. Med Care. 2012;50(10):890–7.
5. Cornelius VR, Sauzet O, Evans SJ. A signal detection method to detect adverse drug reactions using a parametric time-to-event model in simulated cohort data. Drug Saf. 2012;35(7):599–610.
6. WHO-UMC causality assessment.docx; 2011. Available from: URL:http://who-umc.org/Graphics/24734.pdf. Accessed 12 Dec 2012.
7. Naranjo CA, Busto U, Sellers EM, et al. A method for estimating the probability of adverse drug reactions. Clin Pharmacol Ther. 1981;30(2):239–45.
8. Agbabiaka TB, Savovic J, Ernst E. Methods for causality assessment of adverse drug reactions: a systematic review. Drug Saf. 2008; 31(1).
9. The Health Improvement Network [cited 2012 Oct 10]. Available from: URL: www.thin-uk.com.
10. Therneau T. A Package for Survival Analysis in S.: R package version 2.36-14. 2012.
11. Hess K, Gentleman R. Package 'muhaz'; 2012 [cited 2012 Nov 16]. Available from: URL: http://cran.r-project.org/web/packages/muhaz/muhaz.pdf. Accessed 12 Dec 2012.
12. Carvajal A, Martin Diego I de, Matthews J, et al. Carpal Tunnel syndrome associated with bisphosphonates. A population-based cohort study. Pharmacoepidemiol Drug Saf. 2011; 47(12): 1863–72. doi:10.1016/j.ejca.2011.03.027.
13. Jones D, Savage R, Highton JG. Synovitis induced by alendronic acid can present as acute carpal tunnel syndrome. Br Med J. 2005;330(7482):74.
14. Gomes I, Becker J, Ehlers JA, et al. Seasonal distribution and demographical characteristics of carpal tunnel syndrome in 1039 patients. Arquivos de Neuro-Psiquiatria. 2004; 62(3A):596–9.
15. The Uppsala Monitoring Centre. Available from: URL:http://www.who-umc.org/. Accessed 12 Dec 2012.
16. Medicines and Healthcare products Regulatory Agency. Available from: URL: www.mhra.gov.uk/. Accessed 12 Dec 2012.
17. U.S. Food and Drud Administration. Available from: URL: http://www.fda.gov/safety/FDAsSentinelInitiative/ucm2007250.htm. Accessed 12 Dec 2012.
18. European Medicines Agency. Available from: URL:http://www.ema.europa.eu/ema/. Accessed 12 Dec 2012.
19. Eurepean Network of Centres for Pharmacoepidemiology and Pharmacovigilance. Available from: URL: http://www.encepp.eu/. Accessed 12 Dec 2012.
20. Salas M, Hofman A, Stricker BH. Confounding by indication: an example of variation in the use of epidemiologic terminology. Am J Epidemiol. 1999;149(11):981–3.
21. Kaplan Y, Kurt SG, Karaer H. Carpal tunnel syndrome in postmenopausal women. J Neurol Sci. 2008;270(1–2):77–81.