EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

12 February 2019
EMA/662979/2018
European Medicines Agency

Angiotensin II receptor blockers and risk of cancer after contamination with N-nitrosodimethylamine (NDMA) and N-nitrosodiethylamine (NDEA)

Feasibility analysis for study on excess risk of cancer associated to sartans

DRAFT

Investigators	Alexandra Pacurariu, Karin Hedenmalm, Jim Slattery, Maria Gonzalez Martinez
Sign off	Xavier Kurz
First date:	28th September 2019
Last update:	11th February 2019

30 Churchill Place ● Canary Wharf ● London E14 5EU ● United Kingdom
Telephone +44 (0)20 3660 6000 **Facsimile** +44 (0)20 3660 5555
Send a question via our website www.ema.europa.eu/contact

Table of Contents

1. Background

In June 2018, a MAH found that valsartan API supplied by Zhejiang Huahai (ZH) was contaminated by N-nitrosodimethylamine (NDMA), which is a "probable human carcinogen" as classified by the International Agency for Research on Cancer (IARC) (Class 2a carcinogen) (1). NDMA is an unexpected process impurity for valsartan. A referral was started to investigate the issue and, as precautionary immediate measure, national authorities across the EU recalled from pharmacies all medicines containing valsartan supplied by ZH. In addition, ZH is no longer authorised to manufacture the valsartan active substance for EU medicines.

In September 2018, a further N-nitroso impurity N-nitrosodiethylamine (NDEA) has been found in earlier valsartan batches manufactured by ZH, and more recently also in a batch of losartan manufactured by Hetero Labs, India. Subsequently, as a precautionary measure, the referral was extended to all sartans authorised in the EU which contain a tetrazole ring: candesartan, irbesartan, losartan, olmesartan and valsartan.

The contamination is thought to have occurred for valsartan manufactured by ZH between 2011-2015 for NDEA and NDMA and between 2016-2018 for NDMA only. (2) For the period 2012-2015 the two contaminations are potentially overlapping.

Based on preclinical toxicology studies, the theoretical excess lifetime cancer risk for valsartan exposure was calculated to be 21.5: 100,000 patients if taking 320 mg/day Valsartan contaminated with 24.1 µg for 6 years.(2) For NDEA, the theoretical excess lifetime cancer risk for valsartan exposure was calculated to be 8:100,000 patients if taking 320 mg/day Valsartan contaminated with 3.7 µg for 4 years.

A potential worst case scenario for valsartan would be an exposure to NDEA for 4 years (2011 – 2015) and NDMA exposure for 6 years (2012 – 2018). The cumulative theoretical excess risk would then be 29.5/100,000.

For the other sartans, the maximum level of contaminants is much lower than for valsartan and in some cases unknown or below detection limit, see Table 1. Therefore a theoretical excess risk per substance will not be calculated, as it is not considered relevant.

Table 1 Maximum level of contaminants per substance

	Maximum level (ppm)	
	NDMA	NDEA
Candesartan	0.272	0.08
Irbesartan	Below detection	Below detection
Losartan	0.092	0.054
Olmesartan	Unlikely to be contaminated	
Valsartan	61.3	11.53

Of note, NMDA might also be present in contaminated food and beverages and in the environment in general.

From the existing preclinical data, it appears that the liver and the gastrointestinal tract might be potential sites for carcinogenic effect, however the target organs for humans are still under investigation by the Safety Working Party (SWP).(8)

2. Objectives

The objectives of this feasibility analysis are:

- to estimate the sample size (number of exposed patients) that would be needed to identify an excess risk of cancer associated with sartans under different assumptions about the background rate of cancer in Europe and assumptions of the relative risk.
- To identify European population databases that would be appropriate to analyse this risk and estimate whether their size in terms of estimated exposure to sartans would be adequate for the analysis.

3. Epidemiological considerations of study on association between sartans and cancer

Due to comparatively small study populations and short follow up in clinical trials, the evaluation of an association between drug exposure and cancer occurrence for a theoretical small increase in cancer risk arising from preclinical studies cannot be done using data from clinical trials, and instead other population-based data sources need to be considered, including observational epidemiological studies.(3) Studying the association between a particular treatment and occurrence of cancer in epidemiological studies is complicated by characteristics of this association such as:

- Long induction and latency periods, requiring long follow up. The exact induction time is usually not known and differs depending on the cancer type and the mechanism of carcinogenesis (the drug being an initiator or promotor) but is usually at least a few years.
- Cancer often starts with insidious symptoms leading to difficulty in determining the date of onset of cancer, and hence whether the treatment preceded the cancer.
- The types of cancers to be included in the study need to be considered. For rare cancer types large or very large populations may be needed, requiring multi-database and/or multinational studies in order to acquire enough power for the association.
- Recording of the outcome - since existing electronic healthcare records are generally collected for purposes other than research, they may have inadequate or incomplete information about this diagnosis. Cancer-specific registries are the preferred source of information, as they are usually better documented with compulsory reporting in some countries, maximizing the completeness, but information on drug exposure prior to the cancer may be missing.(4) However cancer outcomes are also recorded, with various levels of detail and accuracy, in primary care or claims databases and these can be used especially if the outcome was validated.
- Cancer encompasses a complex aetiology, therefore it is essential to adjust for a wide range of confounders, many of them environmental or lifestyle related (smoking, diet, sun-exposure, rural vs. urban living) which might not be available in electronic healthcare records.
- Most cancers are dependent on cumulative amount and duration of use. A dose-response relationship is considered to support causality and should be performed if possible.

4. Sartans exposure patterns in Europe

The overall exposure of sartans in Europe was evaluated, in order to inform feasibility of studies in Europe. Information about the individual brand names sold in each country and respective market share was not available at this stage. Information from the following countries is publicly available: Sweden, Norway, Denmark, Finland, Netherlands and France. In addition, the number of patients with a prescription for sartans was collected from a sample of general practitioner (GP) practices in France and Germany.

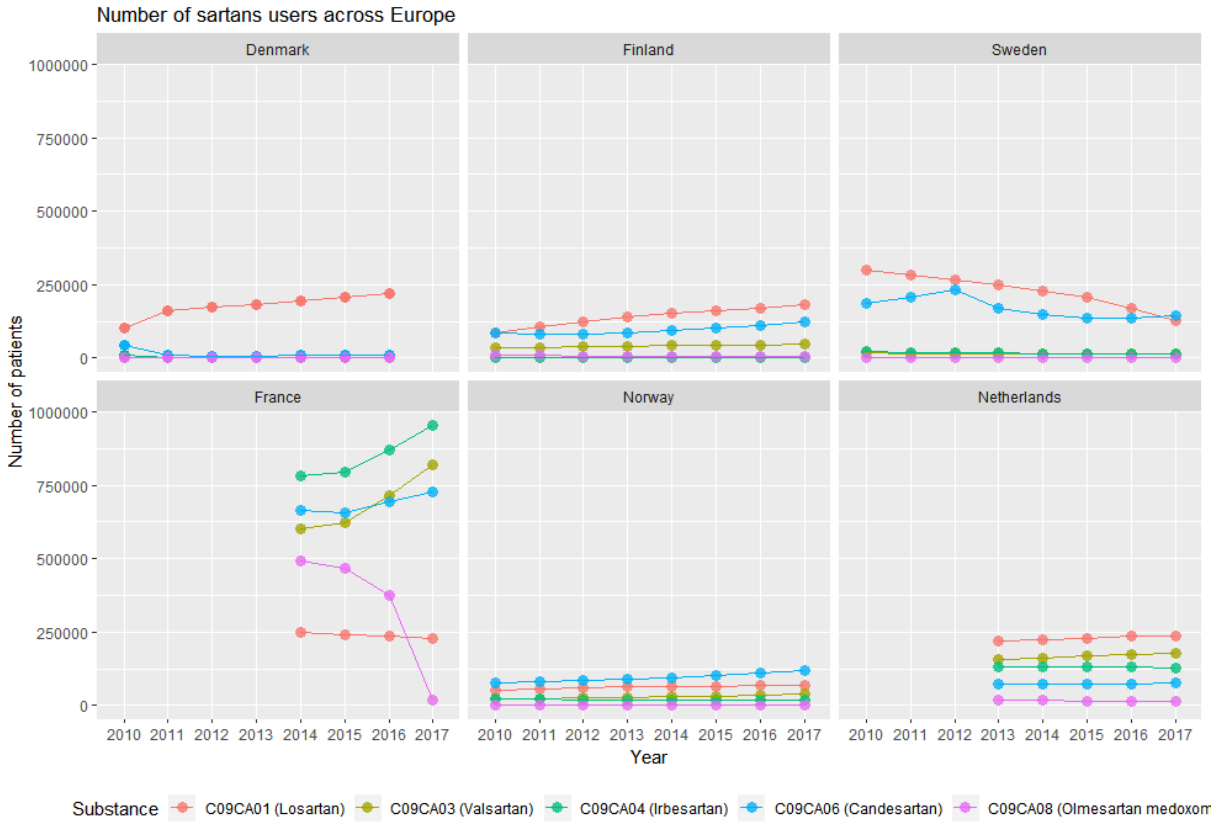
The total use of the five sartans across all countries and the entire period has an average of 363,223 persons/year and a total of 13,246,803 (all sartans, five countries, 2010-2017).

Table 2 Number of patients treated with all sartans in some European countries.

	2010	2011	2012	2013	2014	2015	2016	2017
Denmark	343,004	349,232	364,432	388,330	413,424	443,976	470,404	
Finland	210,939	232,216	252,768	275,270	295,197	317,919	333,879	359,848
Netherlands				599,892	605,244	616,249	629,928	634,641
Norway	173,953	181,557	191,686	201,305	209,604	220,010	231,695	243,306
Sweden	309,217	343,255	376,649	412,723	447,729	484,128	521,373	561,821
France					2,789,675	2,781,549	2,884,617	2,745,720

The lowest consumption is registered for olmesartan in Denmark (average 584 patients across all years) and the highest consumption is for irbesartan in France (average 849,049 patients across all years). In general, losartan and candesartan are the most used, with some regional variations, for example, in France, irbesartan is the most used from the five sartans. Please see Annex 1 for individual country data.

Figure 1 Number of sartans users, per country and individual substance



In order to estimate the adequacy of data sources needed to identify an excess risk of cancer in the available databases, Table 2 provides an estimate of the percentage of the population exposed to all sartans and to valsartan (if the analysis is restricted to this substance) in the six countries.

Table 3. Percentage of the total population being exposed to all sartans and to valsartan in six countries

	Population 2017 *	All sartans 2017 *	Valsartan 2017 *	% all sartans	% valsartan
Denmark	5,711,870	470,404	1,853	8.24	0.03
Finland	5,503,000	359,848	48,090	6.54	0.87
Netherlands	17,080,000	634,641	179,250	3.72	1.05
Norway	5,258,000	243,306	38,038	4.63	0.72
Sweden	9,995,000	561,821	15,892	5.62	0.16
France	67,120,000	2,745,720	818,480	4.09	1.22
All countries	**110,667,870**	**5,015,740**	**1,101,603**	**4.53**	**1.00**

* 2016 for Denmark

From Table 3, the feasibility analysis will need to consider that the percentage of the population exposed to any sartan is not higher than 4.5% (1.0% for valsartan), provided that the average

exposure to sartans in the countries with appropriate databases (see section 4) is comparable to the distribution of the six countries listed in Table 3.

5. Baseline risk of cancer

The baseline risk of cancer in a population is difficult to estimate precisely as it varies with ethnicity, gender, race, country and the type of cancer. During the referral, the overall cancer estimates in Germany were 50.3% in men and 43.5% in women, while in Italy they were 62% in men and 59% in women (during the entire lifespan). (2)

In the report from European Cancer Information System (ECIS) the age standardized incidence of cancer at all sites (except melanoma) was reported 569 per 100,000 persons per year. (5) In another publication, the cumulative risk of cancer for older adults (>60 years) is mentioned to be 13% over a ten year period. (6)

As sartans are used by an older population and the survival time is shorter, the cumulative risk of cancer will be lower for sartan treated patients than for the general population.

Considering all these reported estimates, figures of 5% to 30% baseline risk are considered in the sample size calculations.

6. Power calculations

6.1. Assumptions

There are many uncertainties around parameters as relative risk, percentage of contaminated product and duration of exposure to the contaminant. Therefore, the following assumptions are made for the sample size calculation. They may change when additional information becomes available.

- We assume that all the products in Europe are contaminated (100% contamination degree). This is because the list of contaminated products is likely to be different from country to country and there is a lack of information on which brands are contaminated and what is their market share in each country. Any patient exposed during the contamination period is considered likely to have been exposed to a contaminated product.
- The percentage of the total population being exposed to all sartans in the 6 countries for which public data was available is representative of the whole EU population. This is 4.5% exposure for the sartans currently under review and 1% for valsartan.
- The theoretical excess lifetime cancer risk for valsartan exposure is considered to be 29.5:100,000 (the worst case scenario estimate for taking valsartan contaminated to NDEA for 4 years (2011 – 2015) and NDMA exposure for 6 years (2012 – 2018).
- As the level of contamination for other sartans is much lower than for valsartan, there is no excess risk estimate calculated for other sartans. Therefore, this feasibility analysis is performed for valsartan only.

- The baseline risk of cancer in the population is 10%, with a range of 5% to 30%. The calculation is performed for a composite outcome of all type of cancers (except melanoma) grouped together.

6.2. Sample size needed to detect the estimated relative risk

Considering the upper bound of the estimated excess risk (0.029%), a background incidence rate of cancer of 10%, alpha of 5%, power of 80% and an equal split of exposed and unexposed persons, with a total about 33.6 million patients (16.8 million patients exposed). If we vary the background incidence rate, for 5% background incidence we would need a total of 0.89 million (out of which 0.44 million exposed to the contaminated product) and with 20% we would need 297 million (149 million exposed) for the same parameters.

These calculations are based on certain assumptions regarding the excess risk that might change when more information about the contamination is known.

6.3. Sample size needed to exclude a risk

An alternative sample size analysis could be made by calculating the sample size needed to allow exclusion of a pre-defined risk of cancer with a pre-defined power. Figure 2 shows the minimum excess rates per 100,000 patients in the exposed group that can be excluded with 95% confidence.

Figure 2 Sample size calculations, the minimum excess incidence rate in the exposed group that can be excluded

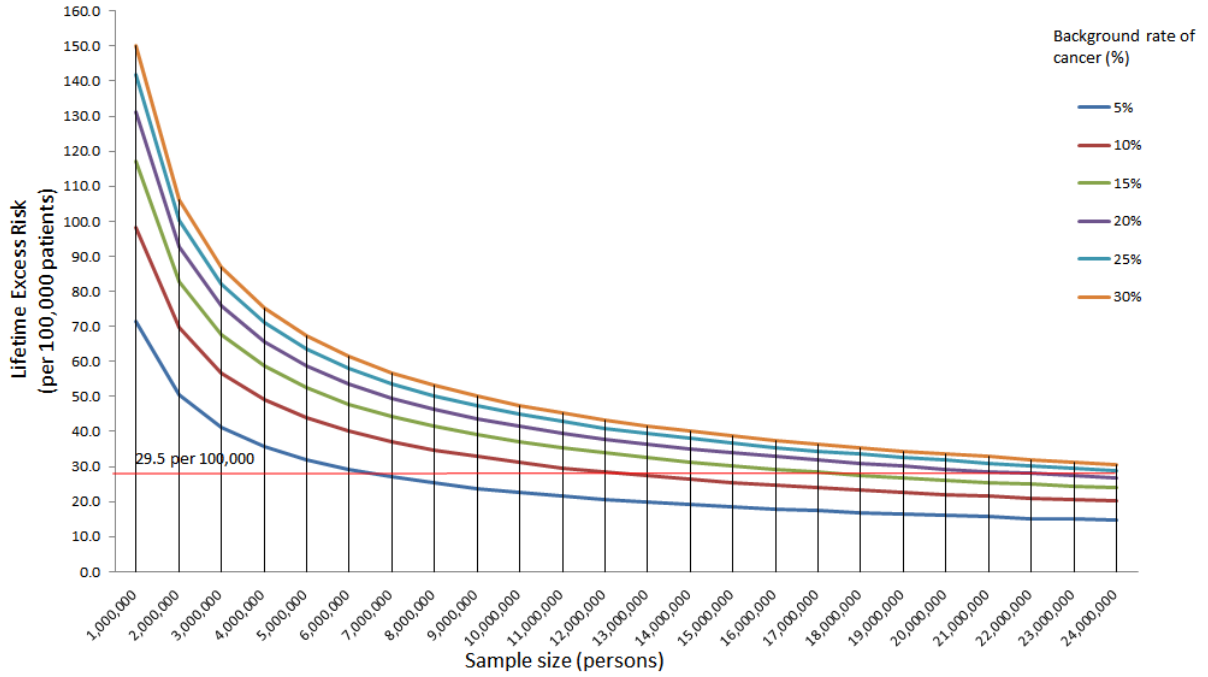


Table 4 Minimum excess incidence rate per 100,000 exposed persons that can be excluded for a certain background risk of cancer

Number of exposed persons in population	Background risk of cancer (%)					
	5%	10%	15%	20%	25%	30%
1,000,000	71.49	98.40	117.12	131.20	142.03	150.31
5,000,000	31.97	44.01	52.38	58.67	63.52	67.22
10,000,000	22.61	31.12	37.04	41.49	44.91	47.53

15,000,000	18.46	25.41	30.24	33.88	36.67	38.81
20,000,000	15.98	22.00	26.19	29.34	31.76	33.61

The grey shaded area represents the situations where the minimum detectable incidence rate is lower that the theoretical excess lifetime cancer risk (29.5/100,000).

If the background risk of cancer in the general population is higher than 25%, we need more than 20 million patients in order to detect the theoretical excess lifetime cancer risk of 29.5/100,000.

7. European databases suitable for study on association between sartans and cancer

Thirty-three (33) European databases were identified from ENCePP's databases repository and from a previous publication.[1] They were evaluated to see how suitable they are for the investigation of this potential association. The database information was obtained from publicly available information. The criteria used for evaluation were:

- Sufficient population capture, both in numbers and follow-up duration

- Sufficient capture of cancer related outcomes and acceptable quality of the recorded outcomes (cancer-related registries are considered gold standard due to the quality of outcome capture)

- Sufficient capture of exposure to sartans and an adequate level of detail regarding the exposure strength, dose, brand name and batch number

- Time window: if the database captures the relevant time window after the start of contamination

One database was excluded as it contains solely paediatric population (PediaNet). From the remaining 32 databases, there were 8 claims databases, 14 electronic medical records databases and 10 record linkage systems (with linked registries) across 13 European countries.

Twenty nine databases contain information on prescribed medicines and at least 12 on dispensed medicines (i.e. pharmacy records). The drug brand name is captured in 17 databases and batch number only in one (Danish registries). Dose information can be obtained from at least 25 databases.

The median size of the database in number of active patients is 4 million patients (range: 0.4-57 million). The coverage from the respective countries or regions is a median of 64%, while at least eleven databases have more than 85% of coverage in the respective country or region. For many databases, information on the average duration of follow up is not readily available, however three databases mention birth to death (or de-registration) follow-up and is assumed this is the case in all primary care databases at least.

Diagnosis is recorded in all of the databases but the completeness of the records and the validity of diagnoses differ according to the database. We consider cancer related registries to be the gold standard for capturing this type of outcome and at least eleven databases can be linked to cancer specific registries data. These are Danish registries, Finnish registries, Swedish registries, Norwegian registries, Icelandic registries, QResearch, Agencia Regionale di Sanita Tuscany database (ARS), Secure Anonymised Information Linkage (SAIL), Pharmo Database Network, Medicines Monitoring Unit Scotland (MEMO) and Clinical Practice Research Datalink - Primary care (CPRD).

[1] Pacurariu A, Plueschke K, McGettigan P, et al Electronic healthcare databases in Europe: descriptive analysis of characteristics and potential for use in medicines regulation BMJ Open 2018;8:e023090. doi: 10.1136/bmjopen-2018-023090

Information regarding the duration of follow-up was not readily available in the majority of cases, however three databases mentioned birth to death follow-up and is assumed this is the case in all primary care databases at least.

All databases cover the period when the contamination started and also a previous period, long enough to allow for a comparison group. There is usually a 1-2 year lag time in data update that has to be considered.

Published studies on cancer conducted within these databases were identified from Pubmed. Sixteen (16) databases had at least one publication on cancer outcomes, and at least one of them had one validation study on a cancer outcome (breast cancer) (i.e. Lazio Hospital Information System).

If we sum up the conditions of cancer registries being linked, adequate time window to capture the contamination moment, over 1 million active patients and information about exposure available at brand level, we have at least fourteen EU database that qualify. If we take the brand name criteria out, there will be at least twenty nine (29) of them (see Annex 2). The total number of active patients covered by the fourteen (14) databases is 168.4 million patients. Based on the figures of Table 2, the number of persons exposed to any sartan in these databases could be 7,6 million and the number of persons exposed to valsartan would be 1.7 million. These calculations did not consider the percentage of contamination among a specific product.

In conclusion, there are at least fourteen (14) EU databases where a study investigation the link between sartans and cancer can be performed. They consist of an adequately high population and are representative for Europe.

It is unfeasible to have batch number information in these databases.

Limitations: as the information was mainly obtained from publicly available sources, there might be some discrepancies with the current functionalities of the database. In order to solve this limitation, pre-selected databases might be individually contacted in order to obtain further information on duration of follow up as well as other particularities (e.g. coding of brand name, linkage to other databases/registries).

8. Conclusion

The study of this association is challenging due to a number of factors, the most important being the large sample size and duration of follow-up required.

The evaluation of European databases showed that we have at least fourteen (14) databases where the study can be conducted, with a total population of 168.4 million patients, out of which approximately 7,6 million will be exposed to any sartan under review and 1.7 million to valsartan only, based on estimated use.

The power calculations shown in section 4.3 predict that with 1.7 million exposed to valsartan we can exclude an excess risk between 50.5 (with a baseline risk of 5%) and 106.3 (with a baseline risk of 30%) new cancers per 100,000 persons.

The follow-up issue is harder to solve as the exposure to the potentially carcinogenic substance occurred in the recent past, and cancer has a long induction period. If the study starts in 2019 we will have at most 7 years of follow-up after the earliest exposure to the contaminant; it is unknown if this is enough to detect an effect.

These estimates are based on a high number of assumptions and they can change once more information is known.

Especially with regards to the level of contamination, this calculation assumed a 100% contamination of all valsartan products which is an overestimation of the actual (unknown) contamination. If an actual study will be pursued it is essential to have information about the national level of contamination and which products are contaminated.

On the other hand, with regards to the outcome, if a more specific type of cancer is suspected, the study will become more feasible, as it is easier to observe an effect in an isolated type of cancer.

With the current information, an observational study in Europe will not be sufficiently powered to detect the currently estimated excess risk but it can be performed in order to exclude a theoretical higher risk.

9. References

1. International Agency for Research on Cancer, Weltgesundheitsorganisation, editors. IARC monographs on the evaluation of carcinogenic risks to humans, volume 100 B, biological agents: this publication represents the views and expert opinions of an IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, which met in Lyon, 24 February - 03 March 2009. Lyon: IARC; 2012. 475 p.

2. EMA. Reader's guidance, Valsartan referral. 2018 Sep;

3. Pinheiro SP, Rivera DR, Graham DJ, Freedman AN, Major JM, Penberthy L, et al. Challenges in evaluating cancer as a clinical outcome in postapproval studies of drug safety. Ann Epidemiol. 2016 Nov;26(11):735–40.

4. Pottegård A, Friis S, Stürmer T, Hallas J, Bahmanyar S. Considerations for Pharmacoepidemiological Studies of Drug-Cancer Associations. Basic Clin Pharmacol Toxicol. 2018 May;122(5):451–9.

5. European Cancer Information System [Internet]. [cited 2019 Jan 20]. Available from: https://ecis.jrc.ec.europa.eu/explorer.php?$0-0$1-All$2-All$4-1,2$3-0$6-0,14$5-2008,2008$7-7$CEstByCountry$X0_8-3$X0_19-AE28E$X0_20-No$CEstRelative$X1_8-3$X1_9-AE28$X1_19-AE28E$CEstByCountryTable$X2_19-AE28E

6. Age and Cancer Risk: A Potentially Modifiable Relationship - ScienceDirect [Internet]. [cited 2018 Oct 3]. Available from: https://www.sciencedirect.com/science/article/pii/S0749379713006429

7. GmbH A-MDA. Valsartan: ZL findet NDMA in Tabletten [Internet]. Pharmazeutische Zeitung online. [cited 2018 Oct 3]. Available from: https://www.pharmazeutische-zeitung.de/ausgabe-302018/zl-findet-ndma-in-tabletten/

8. The Carcinogenic Potency Project (CPDB) [Internet]. [cited 2018 Oct 3]. Available from: https://toxnet.nlm.nih.gov/cpdb/

10. Appendices

Appendix 1. Data from publicly available drug consumption databases

Table 1A. Sartans use across European countries (Number of patients)

Number of patients	2010	2011	2012	2013	2014	2015	2016	2017
Denmark								
C09CA01 Losartan	104,393	159,769	172,163	183,763	194,250	207,691	219,702	na
C09CA03 Valsartan	9,912	2,013	1,493	1,586	1,720	1,778	1,853	na
C09CA04 Irbesartan	11,923	1,933	1,164	1,257	1,416	1,565	1,573	na
C09CA06 Candesartan	42,790	10,394	7,111	7,324	9,116	10,760	11,898	na
C09CA08 Olmesartan medoxomil	2,484	507	285	235	210	194	176	na
France								
C09CA01 Losartan	na	na	na	na	250,419	239,209	235,133	229,525
C09CA03 Valsartan	na	na	na	na	601,201	622,536	716,030	818,480

30 Churchill Place ● Canary Wharf ● London E14 5EU ● United Kingdom
Telephone +44 (0)20 3660 6000 **Facsimile** +44 (0)20 3660 5555
Send a question via our website www.ema.europa.eu/contact

Number of patients	2010	2011	2012	2013	2014	2015	2016	2017
C09CA04 Irbesartan	na	na	na	na	781,460	793,415	867,924	953,395
C09CA06 Candesartan	na	na	na	na	662,409	657,636	692,623	727,670
C09CA08 Olmesartan medoxomil	na	na	na	na	494,186	468,753	372,907	16,650
Norway								
C09CA01 Losartan	51,872	57,364	61,598	63,645	64,068	65,509	66,580	67,318
C09CA03 Valsartan	20,748	22,462	25,161	28,091	30,014	32,316	35,086	38,038
C09CA04 Irbesartan	21,419	20,347	19,740	19,084	18,709	18,440	18,207	17,849
C09CA06 Candesartan	78,234	79,742	83,545	88,847	95,213	102,192	110,280	118,563
C09CA08 Olmesartan medoxomil	1,680	1,642	1,642	1,638	1,600	1,553	1,542	1,538
Netherlands								
C09CA01 Losartan	na	na	na	220,540	223,860	228,630	234,610	236,650
C09CA03 Valsartan	na	na	na	158,010	162,600	168,040	175,010	179,250
C09CA04 Irbesartan	na	na	na	132,870	130,660	130,390	130,350	128,600
C09CA06 Candesartan	na	na	na	71,229	71,699	73,366	74,691	75,511
C09CA08 Olmesartan medoxomil	na	na	na	17,243	16,425	15,823	15,267	14,630
Sweden								
C09CA01 Losartan	125,980	170,852	207,721	230,293	248,406	266,252	282,914	299,508

Number of patients	2010	2011	2012	2013	2014	2015	2016	2017
C09CA03 Valsartan	18,181	16,183	15,627	15,398	15,402	15,514	15,707	15,892
C09CA04 Irbesartan	22,686	19,818	17,722	16,676	15,931	15,195	14,639	14,069
C09CA06 Candesartan	142,370	136,402	135,579	150,356	167,989	187,165	208,110	232,350
C09CA08 Olmesartan medoxomil	0	0	0	0	1	2	3	2
Finland								
C09CA01 Losartan	83,579	104,919	124,796	140,292	151,701	163,179	171,245	183,192
C09CA03 Valsartan	34,119	35,562	37,583	40,200	42,138	44,159	45,373	48,090
C09CA04 Irbesartan	0	0	0	0	0	0	0	0
C09CA06 Candesartan	84,808	83,323	82,480	87,313	94,092	103,641	110,654	122,303
C09CA08 Olmesartan medoxomil	8,433	8,412	7,909	7,465	7,266	6,940	6,607	6,263

Table 1B. Sartans use across European countries (Number of patients) – from in-house databases

In-house(EMA) databases*								
France -IMS Disease Analyser database	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**
C09CA01 Losartan	1,975	2,019	2,176	2,308	2,737	3,002	2,926	3,131
C09CA03 Valsartan	3,885	4,115	4,601	5,307	6,908	7,555	9,036	11,017
C09CA04 Irbesartan	4,567	5,022	5,759	6,672	8,694	9,812	10,997	13,015
C09CA06 Candesartan	4,594	5,060	5,623	6,552	8,282	9,251	9,553	10,653

In-house(EMA) databases*								
C09CA08 Olmesartan medoxomil	3,034	3,714	4,453	5,417	7,132	9,636	7,936	2,458
Germany -IMS Disease Analyser database								
C09CA01 Losartan	6,353	10,464	12,448	13,323	14,233	13,427	12,574	11,605
C09CA03 Valsartan	10,589	12,861	23,573	31,364	40,683	44,894	47,772	48,697
C09CA04 Irbesartan	4,845	5,106	5,787	6,967	8,039	7,886	7,374	6,861
C09CA06 Candesartan	20,224	23,202	30,780	41,309	56,822	63,930	68,898	71,236
C09CA08 Olmesartan medoxomil	10,041	11,423	12,708	13,957	13,031	2,762	2,206	2,647

* The numbers represents a sample size from the entire population. In 2013, there were 1090 (GP) physicians in IMS Disease Analyzer out of a total of 60,043 (GP) physicians in France, i.e. 1.8%. In Germany, 1486 out of 51,998 GPs without diabetologists were included in IMS Disease Analyzer, i.e. 2.9% (data from Physician's statistics BÄK/KBV 12/2012 and IMS Disease Analyzer PRO MAT 09/2015).

Appendix 2. Characteristics of the EU databases for the study of sartans and risk of cancer

Table 1. Eligible EU databases for the study of sartans and risk of cancer

Data source name	Country	Type of data source*	Number of active patients (milion patients)	Prescribed medicines	Dispensed medicines	Type of record	Brand name	Diagnose recorded	Type of record	Years of follow up
France										
QuintilesIMS Disease Analyser	France	EMD	4.4	Yes	No	Primary care	Yes	Yes	Primary care	Unknown
Securite Sociale de l'Assurance Maladie	France	Claims	57	Yes	Yes	Primary care/Secondary care	Yes	Yes	Primary/Secondary care	Lifespan
Germany										
QuintilesIMS Disease Analyser	Germany	EMD	15	Yes	No	Primary/Secondary care	Yes	Yes	Primary/Secondary care	Unknown
German Pharmacoepidemiological Research Database	Germany	Claims	13	Yes	Yes	Primary care	Yes	Yes	Primary/Secondary care	Unknown
Spain										
Base de Datos para la Investigación Farmacoepidemiológica en Atención Primaria	Spain	EMD	8.1	Yes	No	Primary care	Yes	Yes	Primary care	5.7

Data source name	Country	Type of data source*	Number of active patients (milion patients)	Prescribed medicines	Dispensed medicines	Type of record	Brand name	Diagnose recorded	Type of record	Years of follow up
National Health Fund	Poland	Claims	32.3	Yes	No	Primary care/Secondary care	Yes	Yes	Primary/Secondary care	Unknown
Sweden										
Swedish National Registries	Sweden	RLS	9.9	Yes	Yes	Primary care/Secondary care	Yes	Yes	Cancer Registry	Lifespan
Denmark										
Danish National and regional registries	Denmark	RLS	5.5	Yes	Yes	Primary care/Secondary care	Yes	Yes	Cancer Registry	Lifespan
Iceland										
Icelandic Registries	Iceland	RLS	1	Yes	Yes	Primary care	Yes	Yes	Cancer Registry	Lifespan
Norway										
Norwegian Registries	Norway	RLS	5.4	Yes	Yes	Primary care/Secondary care	Yes	Yes	Cancer Registry	Lifespan
United Kingdom										

Data source name	Country	Type of data source*	Number of active patients (milion patients)	Prescribed medicines	Dispensed medicines	Type of record	Brand name	Diagnose recorded	Type of record	Years of follow up
Hospital Treatment Insights	United Kingdom	RLS	2.1	Yes	No	Primary care	Yes	Yes	Secondary care	Unknown
Clinical Practice Research Datalink - Primary care	United Kingdom	EMD	10	Yes	No	Primary care	Yes	Yes	Cancer Registry	10*
Secure Anonymised Information Linkage	United Kingdom	RLS	1	Yes	No	Primary/Secondary care	Yes	Yes	Cancer Registry	20
Italy										
Agencia Regionale di Sanita Tuscany database	Italy	Claims	3.7	Yes	Yes	Primary care	Yes	Yes	Cancer Registry	Unknown

*EMD: electronic medical records; RLS: record linkage system