

A Systematic Review of Case-Identification Algorithms Based on Italian Healthcare Administrative Databases for Two Relevant Diseases of the Endocrine System: Diabetes Mellitus and Thyroid Disorders

Teresa Dalla Zuanna,¹ Gisella Pitter,² Cristina Canova,¹ Lorenzo Simonato,¹ Roberto Gnani³

¹ *Department of Cardiological, Thoracic and Vascular Sciences, University of Padua, Padua (Italy)*

² *Local Health Unit “Azienda ULSS 2 Marca Trevigiana”, Veneto Region, Treviso (Italy)*

³ *Epidemiology Unit, ASL TO3, Regione Piemonte, Grugliasco, Turin (Italy)*

Corresponding author: Teresa Dalla Zuanna; teresa.dallazuanna@studenti.unipd.it

Supplementary materials

Table S1. Characteristics of pertinent papers for diabetes mellitus included in the review.

Author, year	Journal	Setting	Analyzed period	Selected age range	Objective of the algorithm used*	Incident case identification	Algorithm ID	RESULTS Prevalence and/or incidence rate estimates (only articles with objective I) *
Articles reporting algorithms with Objectives I-III								
Brocco, 2007	Cardiovascular Diabetology	3 LHUs of Veneto Region	2001-2004	all ages	I	Yes (1 year back)	1	PR: M= 39-43-1000, F= 34,7-37,6-1000 (range annual prevalences) IR: M= 5,5-1000, F= 5-1000
Forastiere, 2007	Am J Ind Med	Rome	1996-1999	35+	III	-	2	
Giarrizzo, 2007	Epidemiol Prev	Viterbo	2004	all ages	I	-	3	PR: 6,3-8,8 (depending on linkage combination)
Gnavi, 2008a	Nutrition, Metabolism & Cardiovascular Diseases	Turin	1995-2003	all ages	I	-	4	PR: 4,79%
Gnavi, 2008b	Epidemiol Prev	4 Italian LHUs: Venezia, Thiene, Torino, Firenze	1998-2004	all ages	I	-	4	PR: 3,69%-4,18% in different towns
Bruno, 2008	Diabetologia	Turin	2003	all ages	II	-	5, 30	
Lonati, 2008	Epidemiol Prev	Brescia LHU	2004	all ages	I	-	6	PR: 36,6-1000
Gnavi, 2009	Diabetes Care	Turin	1995-2003	20+	II	-	4	
Nordio, 2009	Giornale Italiano di Nefrologia	Venice LHU (n°12)	1998-2002	all ages	II	Yes (1 year back)	7	
Baviera, 2011	Diabetes Research and Clinical Practice	Lombardy Region	2000-2008	all ages	II	-	8	

Author, year	Journal	Setting	Analyzed period	Selected age range	Objective of the algorithm used*	Incident case identification	Algorithm ID	RESULTS Prevalence and/or incidence rate estimates (only articles with objective I) *
Chini, 2011	BMC Public Health	Latium Region	2006	all ages	I	-	9	PR: 44,1-1000
Magoni, 2011	HIV Medicine	Brescia LHU	2003-2007	all ages	III	-	10	
Marchesini, 2011	Nutrition, Metabolism & Cardiovascular Diseases	22 LHUs of the italian ARNO observatory	1997-2006	all ages	I	Yes (1 year back)	11	PR: from 3,08% to 4,45% in different years IR: (given separately for treatments): Oral therapy=0,50 to 0,57%; Insulin=0,05 to 0,08%; combination: 0,02 to 0,03%
Monesi, 2011	Diabetic Medicine	Lombardy Region	2000-2007	all ages	I	Yes (1-8 years back, all the available previous years)	12	PR: 3,0%-4,2% (crude) in different years, (3,2-4,1 std). IR:4-1000 (crude and std)
Seghieri, 2011	Exp Clin Endocrinol Diabetes	Tuscany Region	2002-2008	30+	II	-	2	
Bruno, 2012	Nutrition, Metabolism & Cardiovascular Diseases	Turin	1989-2003	all ages	II	-	5, 30	
De Berardis, 2012a	Nutrition, Metabolism & Cardiovascular Diseases	21 Italian LHUs	2007-2008	all ages	II	-	13	
Fantini, 2012	Health Care Management Review	Bologna LHU	2006-2007	all ages	II	-	14	
Giorda, 2012	PLOS One	Turin	1997-2001	20+	II	-	4	
Degli Esposti, 2013	ClinicoEconomics and Outcomes Research	Arezzo and Ravenna LHUs	2009	45+	II	-	15	

Author, year	Journal	Setting	Analyzed period	Selected age range	Objective of the algorithm used*	Incident case identification	Algorithm ID	RESULTS Prevalence and/or incidence rate estimates (only articles with objective I) *
Gini, 2013	BMC Public Health	5 Regions of the VALORE Project	2003-2008	16-95	I	-	4	PR: 4,9%-7,1 % in different Regions
Valent, 2013	Journal of Diabetes Investigation	Friuli Venezia Giulia Region	2000-2010	all ages	I	-	16, 17	PR: 6,0%, 8,1% (with 2 different algorithms)
Visca, 2013	Health Policy	Patients of GPs with more than 3000 patients, from 6 Italian Regions	2003-2007	16+	II	-	4	
Ballotari, 2014	Diabetes Research and Clinical Practice	Reggio-Emilia province	2009	all ages	I	-	18, 30	PR: 4,4% (DM2), PR: 0,2% (DM1)
Baviera, 2014a	European Journal of Internal Medicine	Lombardy Region	2002-2006	40-89	II	Yes (2 years)	12	
Baviera, 2014b	Eur J Clin Pharmacol	Lombardy Region	2000 AND 2010	65+	II	-	19	
Buja, 2014	BMC Endocrine Disorders	21 Districts in 8 LHUs in 6 Regions of the VALORE project	2003-2008	16+	I	-	4	PR: 5,43%M, 5,05%F
Corrao, 2014	Diabetes Care	Lombardy Region	2003-2010	40-80	III	Yes (No DPD OR HDD in any diagnosis, in the 3 years before the index event)	20	
Gini, 2014	PLOS One	Veneto, Emilia-Romagna, Tuscany, Latium, Sicily Regions	2003-2008	16-95	II	-	4	
Giorda, 2014a	Lancet Diabetes Endocrinol	Piedmont Region	2008-2012	41+	II	-	21	
Giorda, 2014b	Nutrition, Metabolism & Cardiovascular Diseases	Turin	up to 6-2003 (follow back like Gnani)	36-80	II	-	22	

Author, year	Journal	Setting	Analyzed period	Selected age range	Objective of the algorithm used*	Incident case identification	Algorithm ID	RESULTS Prevalence and/or incidence rate estimates (only articles with objective I) *
Marchesini, 2014	Nutrition, Metabolism & Cardiovascular Diseases	30 LHUs of the italian ARNO observatory	2010	all ages	I	-	11	PR: 5,03% italians, 1,96% foreigners
Monesi, 2014	Nutrition, Metabolism & Cardiovascular Diseases	Lombardy Region	2002-2007	40-89	II	Yes (2 years)	12	
Scalone, 2014	PLOS One	Lombardy Region	2000	all ages	I	-	4	PR:3,5% (1 year only)
Vichi, 2014	Acta diabetol	Italy	2005-2010	0-4	I	Yes (4 years back)	32	IR: 13,4-100000
Anello, 2015	Epidemiol Prev	Venice (Mestre) municipality	2002-2009	all ages	III	-	4	
Ballotari, 2015a	BMC Public Health	Reggio-Emilia province	2009	20-74	I	-	23	PR: 2,1%-7,5%Crude; 1,7%-6,5% Adj, divided for sex and citizenship
Ballotari, 2015b	International J of Endocrinology	Reggio-Emilia province	2009	20-84	II	-	23	
Fedeli, 2015	Nutrition, Metabolism & Cardiovascular Diseases	Veneto Region	2009-2012	20-59	I	-	4	PR: M=4%-2,4%, F=2,5%-1,4%, Italians and migrants
Franchini, 2015	Current Pharmaceutical Design	1 LHU of Tuscany Region	2010	75+	II	-	24	
Giorda, 2015	BMJ Open	Piedmont Region	2009-2013	56+	II	-	21	
Marigliano, 2015	Diabetes research and clinical practice	Veneto Region	2008-2013	0-18	I	not present in the initial list (not specified from which date)	33	PR: 1,26-1000 IR: 16,5-100,000

Author, year	Journal	Setting	Analyzed period	Selected age range	Objective of the algorithm used*	Incident case identification	Algorithm ID	RESULTS Prevalence and/or incidence rate estimates (only articles with objective I) *
Pagano, 2015	BMC Health services research	Turin	1989-2003	all ages	II	-	4	
Policardo, 2015	Journal of Diabetes and Its complications	Tuscany Region	2005-2011	16+	II	-	4	
Valent, 2015	Journal of Diabetes and Its complications	Friuli Venezia Giulia Region	2002-2014	20-94	II	-	25	
Barletta, 2016	EJPH	Tuscany Region	2009-2012	all ages	II	-	4	
Bianchi, 2016	European Journal of Gastroenterology & Hepatology	Lombardy Region	2006-2011	1-17	II	Yes (No DPD 1 year before the index event)	34	
Bruno, 2016	Nutrition, Metabolism & Cardiovascular Diseases	18 LHUs of the italian ARNO observatory	2002-2012	0-29	I	Yes (1 year back)	35	PR: 137-166,9-100.000 IR: 16.1-100.000
Canova, 2016	The Journal of Pediatrics	Friuli Venezia Giulia Region	1989-2012	0-18	III	Yes (birth cohorts: only DM diagnosis after the index date)	36	
Fortunato, 2016	Journal of Diabetes Research	Apulia Region	2009-2013	0-18	I	Yes (5 years back)	37	IR: 25,2-100.000
Gini A., 2016	Cancer Epidemiology	Friuli Venezia Giulia Region	2002-2009	40-84	II	-	25	
Marcellusi, 2016	BMJ Open Diabetes Research and Care	Marche Region	2008-2011	all ages	II	-	19	
Pagano, 2016	Nutrition, Metabolism & Cardiovascular Diseases	4 LHUs of the italian ARNO observatory	2007-2012	all ages	II	Yes (5 years back)	19	

Author, year	Journal	Setting	Analyzed period	Selected age range	Objective of the algorithm used*	Incident case identification	Algorithm ID	RESULTS Prevalence and/or incidence rate estimates (only articles with objective I) *
Seghieri, 2016	Journal of Diabetes and Its complications	Tuscany Region	2005-2011	all ages	III	Yes (NOT in the DIABETES REGISTRY (described in Gini2013) before index discharge (for 6 years))	26	
Valent, 2016	Acta diabetol	Friuli Venezia Giulia Region	2010-2013	0-18	I	Yes (1 year back)	38	PR: 15,1-10 000 IR: 15,8-100 000
Ballotari, 2017a	International J of Endocrinology	Reggio-Emilia province	2009-2011	30-84	II	-	18	
Ballotari, 2017b	BMC Cancer	Reggio-Emilia province	2009	20-84	II	-	23	
Baviera, 2017a	Acta diabetol	Lombardy Region	2002	65+	II	-	19	
Baviera, 2017b	Nutrition, Metabolism & Cardiovascular Diseases	Lombardy Region	2002 AND 2012	45-84	II	-	19	
D'Ovidio, 2017	European Journal of Neurology	Turin	1989-2014	14+	II	-	27, 39	
Gini, 2017	PLOS One	5 LHUs in Lombardy, Veneto, ER, Tuscany, Apulia	2008-2011	all ages	II	-	28	
Giorda, 2017	European Journal of Endocrinology	Piedmont Region	2007-2012	all ages	III	-	29, 40	
Policardo, 2017a	Acta diabetol	Tuscany Region	2010	all ages	II	-	4	
Policardo, 2017b	Journal of Diabetes and its complications	Tuscany Region	2008	16+	II	-	4	

Author, year	Journal	Setting	Analyzed period	Selected age range	Objective of the algorithm used*	Incident case identification	Algorithm ID	RESULTS Prevalence and/or incidence rate estimates (only articles with objective I) *
Profili, 2017	EJPH	Tuscany Region	2005	15+	II	-	4	
Valent, 2017a	Acta diabetol	Hospital of Udine	2002-2016	all ages	II	-	16	
Valent, 2017b	J Inj Violence Res.	Friuli Venezia Giulia Region	2002-2014	18+	II	-	16	
Valent, 2017c	Journal of Renal Care	Friuli Venezia Giulia Region	2002-2014	all ages	III	-	16	
Articles reporting algorithms with Objectives IV-VII								
Antonelli-Incalzi, 2007	BMC Public Health	Rome	1997-2000	75+	VI			
Gnavi, 2007	EJPH	Turin	2001-2002	30-85	IV			
Corrao, 2009	Journal of Human Hypertension	Milan	1999-2003	40-80	IV			
Agabiti, 2010	Eur Respir J	Rome	1997-30-9-2005	35+	IV			
Corrao, 2011	Journal of Hypertension	Lombardy Region	2000-2007	18+	IV			
Degli Esposti, 2011	Monaldi Arch Chest Dis	5 LHUs in Emilia-Romagna, Tuscany, and Umbria Region	2006-30-6-2008	18+	IV			
De Berardis, 2012b	JAMA	12 LHUs in Apulia Region	2002-2008	all ages	IV			
Macchia, 2012	European Heart Journal	22 Italian LHUs	2002-2007	all ages	IV			
Guaraldi, 2013	ClinicoEconomics and Outcomes	22 LHUs of the italian ARNO observatory	2009	18+	VII			(Buja2014: 5,43%M,5,05%F 16+)

Author, year	Journal	Setting	Analyzed period	Selected age range	Objective of the algorithm used*	Incident case identification	Algorithm ID	RESULTS Prevalence and/or incidence rate estimates (only articles with objective I) *
	Research							
Pereira de Sousa, 2013	Italian Federation of Cardiology	Hospitals of the "Alto friulano" Area	2004-2009	all ages	IV			(Scalone2014:3,5% (1 year considered))
Bonifazi, 2014	The Breast	Lombardy Region	2006-2009	all ages	IV			
Lombardo, 2014	PLOS One	Italy	2001-2010	0-100	VII			
Vinceti, 2014	Eur J Epidemiol	Emilia-Romagna Region	2003-2010	all ages	IV			
Fabiani, 2015	Vaccine	Piedmont, Friuli Venezia Giulia, Latium, Apulia Region	2008-2009	all ages	IV			
Monaldi, 2015	ClinicoEconomics and Outcomes Research	9 LHUs in Lombardy, Veneto, Emilia-Romagna, Tuscany, Abruzzi, Latium, Apulia Regions	2003-2007	18+	IV			
Pintaudi, 2015	Diabetologia	12 LHUs in Apulia Region	2002-2010	all ages	VII			
Maggioni, 2016	European Journal of Heart Failure	5 LHUs of the italian ARNO observatory	2007-2012	all ages	IV			
Notarnicola, 2016	J Biol Regul Homeos Agents	Apulia Region	2006-2010	males 65+ females 50+	VII			
Roberto, 2016	PLOS One	Tuscany Region+ international comparison	2005-2011	16+	VII			
Trifirò, 2016	Clinical Drug Investigation	Caserta LHU	2006-2011	all ages	V			

Author, year	Journal	Setting	Analyzed period	Selected age range	Objective of the algorithm used*	Incident case identification	Algorithm ID	RESULTS Prevalence and/or incidence rate estimates (only articles with objective I) *
Baldi, 2017	Int J Cardiol	Lombardy Region	2007-2010	all ages	IV			
Borsari, 2017	Diabetes Research and Clinical Practice	Emilia-Romagna Region	1997-2010	all ages	IV			
Terzano, 2017	European Review for Medical and Pharmacological Sciences	Rome-A LHU	2010-2012	all ages	IV			

* **I** to measure the occurrence of the disease; **II** to identify a population/cohort of subjects affected by the disease of interest; **III** to identify the disease as outcome; **IV** to identify the disease as comorbidity for statistical adjustments; **V** to identify the disease as exclusion criteria for other conditions; **VI** to calculate hospitalization rates or disease-specific drug prescription rates; **VII** other objectives

PR: prevalence rate; **SPR:** standardized prevalence rate; **IR:** incidence rate; **SIR:** standardized incidence rate

Table S2. Characteristics of pertinent papers for Thyroid disorders included in the review.

Author, year	Journal	Setting	Analyzed period	Selected age range	Objective of the algorithm used*	Incident case identification	Algorithm ID	RESULTS Prevalence and/or incidence rate estimates (only articles with objective I) *
Articles reporting algorithms with Objective I-III								
Chini, 2011	BMC Public Health	Lazio Region, Italy	2006	all years	I	-	1, 2	PR (all thyroid disorders): 43.9-1000
Maccagna, 2016	BioMed Research International	Apulia Region, Italy	2008	women 50+ years; men 65+ years	II	-	3, 4	
Ca-va, 2016	J Pediatr	Friuli Venezia Giulia Region, Italy	1989-2012	0-23 years	III	yes	5, 6	
Giorda, 2017	Eur J Endocri-l	Piedmont Region, Italy	2012-2013	all years	I	yes	7, 8	PR (hypothyroidism): 31.1-1000 (95% CI 30.9–31.3) IR (hypothyroidism): 6.99-1000 (95% CI 6.91–7.07)
Greggio, 2017	Endocrine Connections	Multicentre (Twelve Italian Local Health Units)	2001-2014	0-13 years	I	-	9	PR (subclinical hypothyroidism): 0.20-1000

* **I** to measure the occurrence of the disease; **II** to identify a population/cohort of subjects affected by the disease of interest; **III** to identify the disease as outcome; **IV** to identify the disease as comorbidity for statistical adjustments; **V** to identify the disease as exclusion criteria for other conditions; **VI** to calculate hospitalization rates or disease-specific drug prescription rates; **VII** other objectives

PR: prevalence rate; **SPR:** standardized prevalence rate; **IR:** incidence rate; **SIR:** standardized incidence rate