

# Gebruik van gepseudonimiseerd BSN bij gegevenskoppelingen: BBMRI Validatiekoppelingenstudie

Dr. Sandra van den Belt-Dusebout (NKI-AVL)

&

Dr. Rinus Voorham (PALGA)



# Inhoud

- Epidemiologisch onderzoek
- Koppelen van onderzoekspopulaties met registraties
- Validatiekoppelingenstudie
- Resultaten
- Mogelijke oplossing voor toekomstige koppelingen

# Epidemiologisch onderzoek: relatie tussen een blootstelling en latere ziekte



Onderzoekspopulatie: met blootstelling



Relatief  
risico = 2



Controlepopulatie: zonder blootstelling





# Opvragen van gegevens uit registraties m.b.v. koppelen

onderzoekspopulatie



persoonsidentificatie

Donald Duck M 15-07-1950 9612

Minnie Mouse V 15-07-1950 3055

Gijs Gans M 10-09-1982 6912

Mickey Mouse M 15-09-1934 2246

Katrien Duck V 30-07-2001 3089

Guus Geluk M 09-02-1991 4067

Klaartje Kip V 01-01-1988 2501

registraties



Registratie X



Registratie Y



Registratie Z



# Problemen bij koppelen met NAW

onderzoekspopulatie



persoonsidentificatie

D	Duck	M	15-07-1950	9612
---	------	---	------------	------

registratie



<b>K</b>	Duck	M	15-07-1950	9612
----------	------	---	------------	------

Zelfde?



D	Duck	M	15-07-1950	<b>3844</b>
---	------	---	------------	-------------

Zelfde?



D	<b>Dukc</b>	M	15-07-1950	9612
---	-------------	---	------------	------

Zelfde?



D	Duck	M	<b>01-01-1950</b>	9612
---	------	---	-------------------	------

Zelfde?

# Problemen bij koppelen met NAW

onderzoekspopulatie



persoonsidentificatie

D	Duck	M	15-07-1950	9612
---	------	---	------------	------

registratie



<b>K</b>	Duck	M	15-07-1950	9612
----------	------	---	------------	------

Zelfde?



D	Duck	M	15-07-1950	<b>3844</b>
---	------	---	------------	-------------

Zelfde?



D	<b>Dukc</b>	M	15-07-1950	9612
---	-------------	---	------------	------

Zelfde?



D	Duck	M	<b>01-01-1950</b>	9612
---	------	---	-------------------	------

Zelfde?

Wel/niet koppelen



hoger/lager risico

Extra gegevens nodig



voor zekere koppeling

# Problemen bij koppelen met pseudoniemen

onderzoekspopulatie



persoonsidentificatie

D	Duck	M	15-07-1950	9612
AsX5HH8	GTR65P00	hgTy7KL3	DfTQ23P0	BcV256Y0

registratie



<b>K</b> Duck	M	15-07-1950	9612	Zelfde?
---------------	---	------------	------	---------



D Duck	M	15-07-1950	<b>3844</b>	Zelfde?
--------	---	------------	-------------	---------



D <b>Dukc</b>	M	15-07-1950	9612	Zelfde?
---------------	---	------------	------	---------



D Duck	M	<b>01-01-1950</b>	9612	Zelfde?
--------	---	-------------------	------	---------



# Problemen bij koppelen met pseudoniemen

onderzoekspopulatie



persoonsidentificatie

<i>voorl</i>	<i>naam</i>	<i>geslacht</i>	<i>geb datum</i>	<i>PC4</i>
AsX5HH8	GTR65P00	hgTy7KL3	DfTQ23P0	BcV256Y0

registratie



<b>7uLFHtw6</b>	GTR65P00	hgTy7KL3	DfTQ23P0	BcV256Y0
-----------------	----------	----------	----------	----------

Zelfde?



AsX5HH8	GTR65P00	hgTy7KL3	DfTQ23P0	<b>PD599SSR</b>
---------	----------	----------	----------	-----------------

Zelfde?



AsX5HH8	<b>771AL1nb</b>	hgTy7KL3	DfTQ23P0	BcV256Y0
---------	-----------------	----------	----------	----------

Zelfde?



AsX5HH8	GTR65P00	hgTy7KL3	<b>PvS16AL8</b>	BcV256Y0
---------	----------	----------	-----------------	----------

Zelfde?

# Problemen bij koppelen met pseudoniemen

onderzoekspopulatie



persoonsidentificatie

*voorl*

*naam + geslacht + geb datum*

*PC4*

AsX5HH8

8UQWmKL7

BcV256Y0

registratie



7uLFHtw6

8UQWmKL7

BcV256Y0

Zelfde?



AsX5HH8

8UQWmKL7

PD599SSR

Zelfde?



AsX5HH8

L9SD55HJ

BcV256Y0

Zelfde?



AsX5HH8

VxC70PH2

BcV256Y0

Zelfde?

# BBMRI Science Voucher validatiekoppelingenstudie

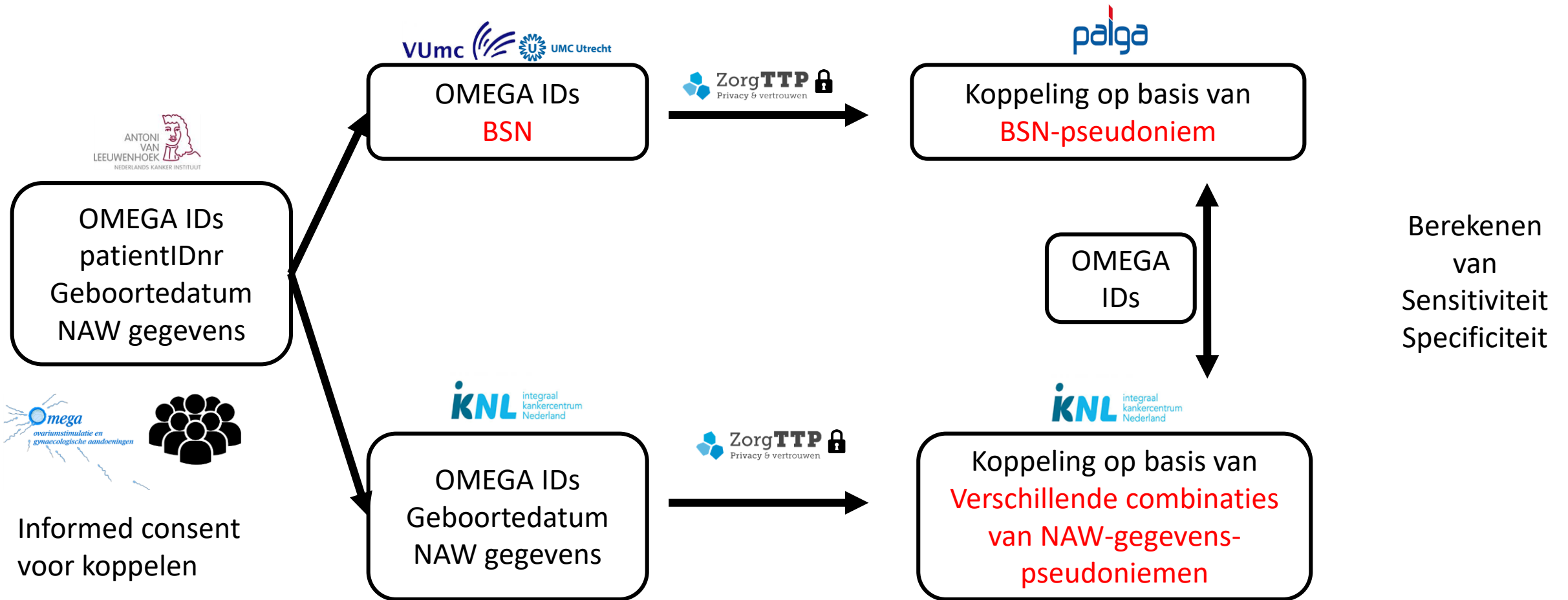


## Doel:

Het ontwikkelen van een efficiënte koppelprocedure met:

- correcte koppelingen
- optimale privacybescherming

# Studie opzet



# Sensitiviteit en specificiteit

		Pseudoniem BSN Koppeling (gouden standaard)		
		Ja	Nee	
NAW pseudoniemen Koppelingen	Ja	A: Terecht gekoppeld	B: Onterecht gekoppeld	A+B
	Nee	C: Onterecht niet gekoppeld	D: Terecht niet gekoppeld	C+D
		A+C: Gekoppeld in gouden standaard	B+D: Niet gekoppeld in gouden standaard	A+B+C+D

*Sensitiviteit =  $(A/(A+C))$  percentage terecht gekoppeld t.o.v. Pseudoniem BSN koppeling*

*Specificiteit =  $(D/(B+D))$  percentage terecht niet-gekoppeld t.o.v. Pseudoniem BSN koppeling*

# Resultaten

Koppel scenario	NAW pseudoniemen							Sensitiviteit	Specificiteit	Administratieve tweelingen
	geslacht	naam	geb datum	geb jaar	voorl	PC4	PC6	%	%	%
1	X	X	X		X			93,8	99,7	1,1
2	X		X				X	81,2	99,7	0
3	X	X		X				96,2	79,9	70,9
4	X	X		X	X	X		85,3	99,6	0
5	X		X		X	X		86,0	99,4	0,8
6	X	X			X	X		82,9	97,1	12,7
7	X	X	X		X	X		85,3	99,7	0
8	X	X	X					97,4	97,8	5,6

# Resultaten

Koppel scenario	NAW pseudoniemen							Sensitiviteit	Specificiteit	Administratieve tweelingen
	geslacht	naam	geb datum	geb jaar	voorl	PC4	PC6	%	%	%
1	X	X	X		X			93,8	99,7	1,1
2	X		X				X	81,2	99,7	0
3	X	X		X				96,2	79,9	70,9
4	X	X		X	X	X		85,3	99,6	0
5	X		X		X	X		86,0	99,4	0,8
6	X	X			X	X		82,9	97,1	12,7
7	X	X	X		X	X		85,3	99,7	0
8	X	X	X					97,4	97,8	5,6

# Resultaten

Koppel scenario	NAW pseudoniemen							Sensitiviteit	Specificiteit	Administratieve tweelingen
	geslacht	naam	geb datum	geb jaar	voori	PC4	PC6	%	%	%
1	X	X	X		X			93,8	99,7	1,1
2	X		X				X	81,2	99,7	0
3	X	X		X				96,2	79,9	70,9
4	X	X		X	X	X		85,3	99,6	0
5	X		X		X	X		86,0	99,4	0,8
6	X	X			X	X		82,9	97,1	12,7
7	X	X	X		X	X		85,3	99,7	0
8	X	X	X					97,4	97,8	5,6



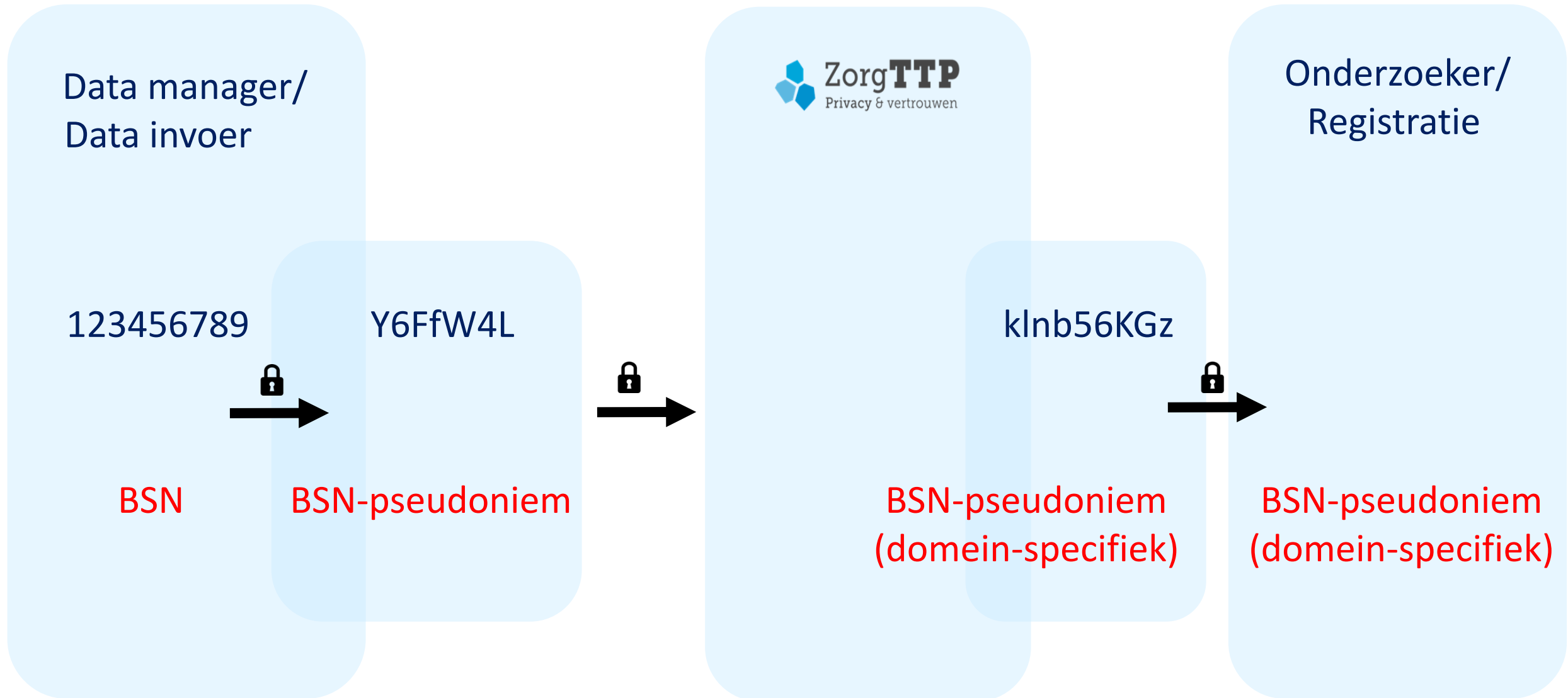
# Conclusies

- Koppelen met NAW-pseudoniemen levert slechtere koppelingen dan met BSN-pseudoniem
- Door gebruik van BSN-pseudoniem is patiënt “minder herleidbaar”
- Koppelen met BSN-pseudoniem is minder tijdrovend en minder uitzoekwerk



Meer met Minder

# Toekomst - stap 1 - invoer

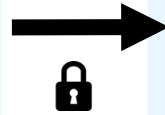


# Toekomst - stap 2 - koppelen

Onderzoeker/  
Registratie

klnb56KGz

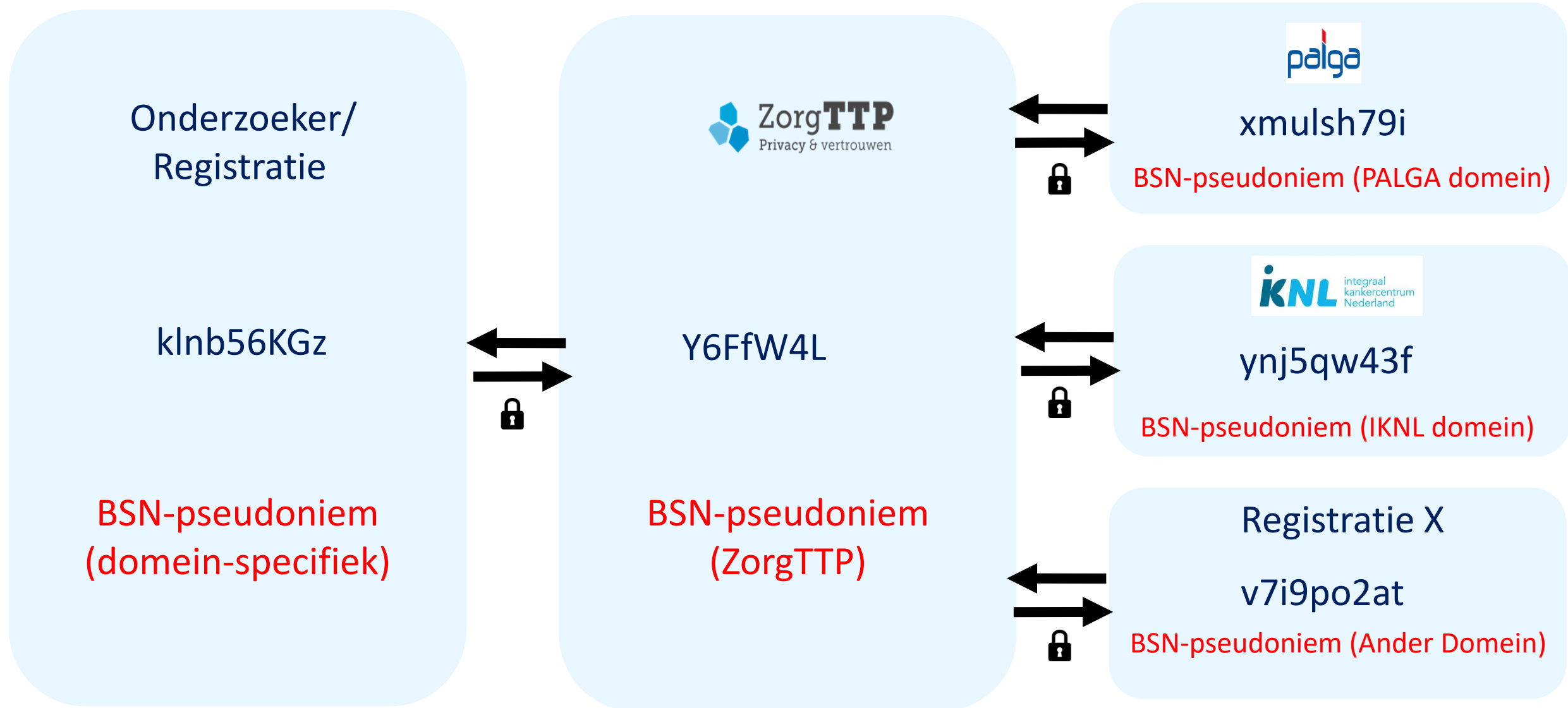
BSN-pseudoniem  
(domein-specifiek)



Y6FfW4L

BSN-pseudoniem  
(ZorgTTP)

# Toekomst - stap 2 - koppelen



# Hiervoor is er nodig?

- Aanpassing wet: toestaan verwerking van BSN voor onderzoek
- Opname van gepseudonimiseerd-BSN in onderzoeksdata en registraties
- (Zorg)TTP voor pseudonimisatie



VRAGEN?





# Gebruik van gepseudonimiseerd BSN bij gegevenskoppelingen: BBMRI Validatiekoppelingenstudie

