COVID Pregnancy-Prevalence-NL

Project

Real time monitoring of prevalence of SARS-CoV-2 infections using residual second trimester blood samples from the Dutch antenatal screening program: a repeated cross-sectional study

Project working group:

Amsterdam UMC

5.1.2e

RIVM

5.1.2e

Sanquin

5.1.2e

Project Steerning-Committee

Epidemiology and Statistics

5.1.2e

Nederlandse Samenvatting

Doel

Het doel van ons onderzoek is om de immuniteit voor het Coronavirus te bepalen bij een ongeselecteerde groep zwangere vrouwen woonachtig in heel Nederland.

Achtergrond

Sinds februari 2020 is Nederland in de ban van het Coronavirus (SARS-CoV-2 virus) dat leidt tot het ziektebeeld COVID-19. Daar niet iedereen even ziek wordt van dit virus en alleen de zieke mensen worden getest is onduidelijk hoeveel mensen daadwerkelijk COVID-19 hebben doorgemaakt. Na een doorgemaakte COVID-19 infectie zijn er antistoffen in het bloed aantoonbaar. Mensen die antistoffen hebben zijn waarschijnlijk voor enkele tijd immuun voor het virus en mensen die nog geen antistoffen hebben zijn nog bevattelijk voor het virus.

Uit angst voor grote aantallen besmette mensen heeft de Nederlandse overheid verregaande maatregelen getroffen, de zogenaamde 'intelligente' of 'targeted' lockdown, om verdere verspreiding van dit virus en piekbelasting in de ziekenhuizen te voorkomen. Deze maatregelen hebben grote economische gevolgen en drijven Nederland en Europa in een economische crisis. Tot 1 september zijn evenementen verboden.
Om deze maatregelen op een gegeven moment te kunnen versoepelen en het Nederlandse beleid, dat liberaler is dan dat van landen met een totale lockdown, te vergelijken met andere landen in Europa ook met het oog op toekomstige uitbraken van virussen, is het van belang dat er betrouwbare data zijn hoe de immuniteit in de samenleving is opgebouwd.
Op dit moment wordt bij verschillende bevolkingsgroepen de immuniteit bepaald. Zo is de immuniteit bij bloeddonoren 3.4%. Echter dit zijn in de regel mensen die geselecteerd worden op hun gezonde leefstijl en zijn met name blanke Nederlanders. Daarnaast is de immuniteit ook bepaald in een steekproef van de Nederlandse bevolking (het Pienter onderzoek). Deze steekproef bestaat uit vrijwilligers en is daarmee mogelijk ook gevoelig voor selectie (bijvoorbeeld mensen die meer angst hebben om het virus te dragen kunnen zich eerder melden als vrijwilliger), en kunnen daarmee de werkelijke immuniteit cijfers onder of overschatten.

In Nederland bevallen per jaar ongeveer 170.000 vrouwen. 14% van de zwangere vrouwen is Rhesus D-negatief en 19% van de zwangere vrouwen is Rhesus c negatief, kortom eenderde van alle zwangeren. Dit is onafhankelijk van afkomst of regio. Bij al deze vrouwen wordt bij een zwangerschapsduur van 27 weken bloed afgenomen om te kijken of zij Rhesus-antistoffen ontwikkelen ten gevolge van de zwangerschap. Dit bloed wordt centraal bij Sanquin geanalyseerd en het restmateriaal wordt bewaard voor de duur van de zwangerschap. Per jaar gaat dit om 56.000 bloedafnames verspreid over het gehele jaar ofwel om ongeveer 1000 bloedmonsters per week.

Methode

Wij stellen met ons onderzoek voor om de immuniteit te bepalen op reeds opgeslagen bloedmonsters van zwangere vrouwen. Het voordeel van deze groep is niet alleen dat zij woonachtig zijn door heel Nederland heen, maar ook dat deze representatief is voor alle culturele en sociaal-economische lagen van de bevolking. Door de immuniteit tegen SARS-CoV-2 in dit restmateriaal te meten kan per week de oplopende immuniteit onder zwangeren tegen dit virus worden gemeten.

Resultaten

De uitkomsten betreffen alleen immuniteitsgegevens tegen SARS-CoV-2 en zijn niet tot de persoon herleidbaar. Zwangeren worden zelf niet geïnformeerd over hun immuunstatus. Op deze manier kan met respect voor de privacywetgeving op eenvoudige wijze door het gebruik van restmateriaal "realtime" informatie worden verkregen. Vervolgens kunnen deze getallen gemodelleerd en geëxtrapoleerd worden naar de rest van de Nederlandse bevolking om zo de Nederlandse overheid behulpzaam te zijn bij het verantwoorden en aanpassen van de lockdown maatregelen. Ook kan het effect van overheidsmaatregelen als het heropenen van de basisscholen per 11 mei en hervatten van sportactiviteiten vanaf 28 april op de immuniteitsontwikkeling worden gemonitord.

Toegevoegde waarde

1) **Alle lagen van de bevolking.** Zwangeren zijn woonachtig door heel Nederland en zijn representatief voor alle culturele en sociaal-economische lagen van de bevolking. Tevens zijn vrouwen uit de leeftijdscategorie (15-45) ondervertegenwoordigd bij de bloeddonoren.
2) **Continue stroom.** Snel inzicht in veranderingen in immuniteit over de tijd: we zijn van plan om 1000 monsters per week te testen en de resultaten ook per week te presenteren. De samples liggen al centraal opgeslagen bij Sanquin en kunnen per robot meegenomen worden met de donoren samples.
3) **Rol van kinderen/kinderopvang/scholen**. Met de doelgroep zwangeren kan mogelijk beter het effect van de opening van de kinderopvang en de scholen op de besmettingsgraad worden gemonitord.

Problem

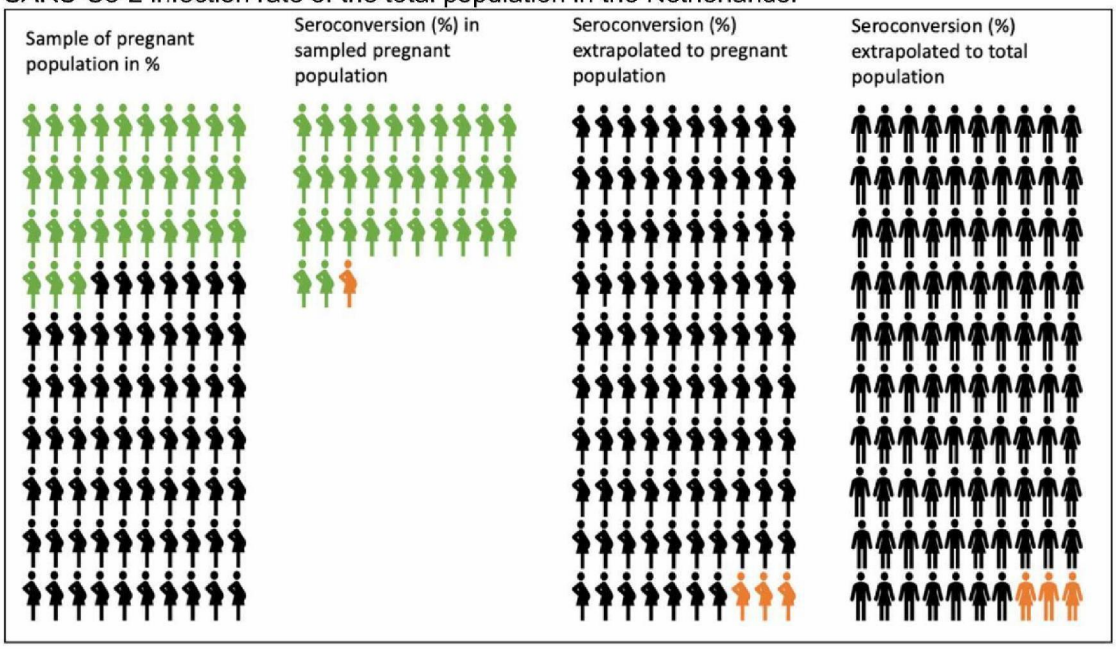
The SARS-CoV-2 pandemic has a major impact on health facilities and economies. Globally, it affects all layers of society with a maximum burden of disease among the elderly population and people with underlying health conditions.[1] At this moment, the official case fatality rate (CFR) is 3.4% (WHO March 3 2020).[2] The CFR data of different countries are widely available (e.g. www.woldometer.info.coronavirus), but also discussed for being unreliable due to its denominator that is highly depended on the number of patients tested (see additional information on the denominator issues in current available mortality data in Appendix 1).[3] The problem at this moment is that the CRF data are being used in mathematical modeling to estimate the impact of the SARS-CoV-2 pandemic (e.g. forecasting SARS-CoV-2 cases, predicting hospital capacity needs) and therefore used to guide critical decisions in healthcare, resource allocations and other political actions. A more reliable estimate of the number of cases and the number of cases that already developed immunity against the novel SARS-CoV-2 (IgG/IgM seroprevalence) can therefore be extremely valuable. These numbers are not known for the Netherlands as for many other countries worldwide. A recent analysis estimated that probably 86% (95%CI 85-90%) of all infections are undocumented.[4] Real-time monitoring of changes in the immunity of the Dutch population in relation to the lockdown measures taken in the Netherlands is of great importance. The Netherlands is one of the few countries that adopted an 'intelligent' or 'targeted' lockdown strategy based on the idea of herd immunity. As this intelligent lockdown differs from the stricter approaches in most other countries that are struggling with SARS-CoV-2, and by some countries even described as 'contentious', a thorough evaluation of the Dutch approach is of highest priority.

In this repeated cross-sectional study, we aim to provide an estimate on the number of people with immunity (IgG/IgM) against the novel SARS-CoV-2 virus in the Netherlands and monitor its changes over time in relation with the Dutch intelligent lockdown measures by using second trimester residual blood samples of pregnant women. With this information, the COVID-19 antibody seroprevalence and the infection fatality rate (IFR) can be estimated on a weekly basis over a 5 months period of time (February 2020 up to June 2020, or longer if necessary). This seroprevalence information will be extremely informative and can change the course of the political actions in the Netherlands (and potentially in other countries). The reliable and 'real-time' seroprevalence information can be used to inform clinical, economic and political decision-making with a weekly update.

Rationale

Pregnant women can be used as a proxy sample of the general population (Figure 1).

Figure 1. Blood samples analyzed in pregnant population (green) providing information on the SARS-CoV-2 infection/seroconversion (orange) that can be extrapolated to the SARS-Co-2 infection rate of the total population in the Netherlands.



Using samples and data from the antenatal screening program (coverage > 99%) in the Netherlands has several advantages:

1) national geographic coverage, this is important as the spread of SARS-CoV-2 is not evenly distributed in the Netherlands.

2) weekly updated data, a dynamic cohort of (500 to 1000/week) pregnant women in the second trimester of pregnancy will be used. The percentage of the population with serologic conversion in response to a (symptomatic or asymptomatic) SARS-CoV-2 infection can be accurately reported on a weekly basis. This 'real-time' data can have a crucial role in clinical and political decision making.

3) stored residual samples, as the (serology) blood samples in the second trimester of pregnancy is routinely collected, we can run the analysis on residual samples (33% of the pregnant population), so no extra blood sampling is needed. Also, part of the blood (14% of the pregnant population) is stored for a period of 6 months, so these samples allow us to 'look back' on the SARS-CoV-2 seroprevalence over the months when SARS-CoV-2 was first introduced into the Netherlands. This will help to understand the rapidity of the spread of the disease from the beginning.

4) extrapolation to the general population, as pregnant women represent all different social economic backgrounds and different comorbidities, this data can be fairly extrapolated to give an estimation on the number of infected people in the Netherlands.

5) Insight in the role of children in the transmission, by using information on gravidity/parity we can compare pregnant women without children in the house (first

pregnancy) with women with children in the house. Parity will be used as a proxy for the number of children in the house.
6) ethical and legal. We will act in accordance with the regulation 'Additional use of anonymous or coded human tissue' of the PSIE screening program for pregnant women. Contracts with Sanquin on the blood samples already incorporate the 'geen bezwaar/no objection' policy (https://draaiboekpsie.nl/over-psie/nader-gebruik-lichaamsmateriaal). This is permitted under the condition that the research protocol has been assessed by a Medical Ethics Committee (METC). The proposal has been declared 'niet-WMO plichtig' by the METC of Amsterdam UMC on 23-04-2020.(METC-number: WO20_177#20.205)

Due to these advantages, we will be able to provide a quick and 'real time' estimate of the SARS-CoV-2 seroprevalence in this population, to track changes in seroprevalence over time, and to extrapolate the outcomes to the total general population in the Netherlands (see statistical analysis plan below). We will also compare our seroprevalence with other seroprevalence studies like the Dutch blood donors study and the Pienter-Corona study.[5,6] Data can also be combined. For instance to determine the age profile of the seroprevalence and to follow this longitudinally. Longitudinal age profiles are now determined based on blood donors. However, women from the age category with small children are underrepresented in the sample of blood donors (reference: person communication 5.1.2e).
The COVID-Pregnancy-Prevalence project on its own (or combined with other projects) can also provide a fairly reliable infection fatality rate (IFR) estimate that is crucial to inform and evaluate clinical, economic and political decision making. Time is crucial, the faster we can get this information, the greater its value for decision making or initiation of other SARS-CoV-2 research projects.

Strategy

In the Netherlands there are approximately 170,000 pregnancies a year. In the pregnant population, all women receive standard serological examination in the first trimester (at 10-12 weeks' gestation) to assess blood type and determination of the fetal Rhesus D and c factor. This examination takes place locally (with 67 laboratories involved). Approximately 14% of all pregnant women are Rhesus D negative and 19% are Rhesus c negative. These women are all retested in the second trimester at 27 weeks' of gestation to assess fetal Rhesus factor for Rhesus D negative women and the presence of irregular antibodies for Rhesus c negative women [7]. All 27 weeks' blood samples are analyzed centrally at the Dutch Blood banking organization (Sanquin). Subsequently the blood samples of Rhesus D negative women are stored up to 6 months after delivery.[7] Since April 2020 the blood samples of Rhesus c women are also stored centrally. Thus, Sanquin yearly receives blood samples from 33% of the pregnant population amounting to approximately 56,100 samples a year or >4,500 samples a month and >1,000 samples a week.

We propose to conduct SARS-CoV-2 IgG and IgM serology on approximately 12.000 of those samples to reach a 97% to 99% confidence level on the prevalence of SARS-

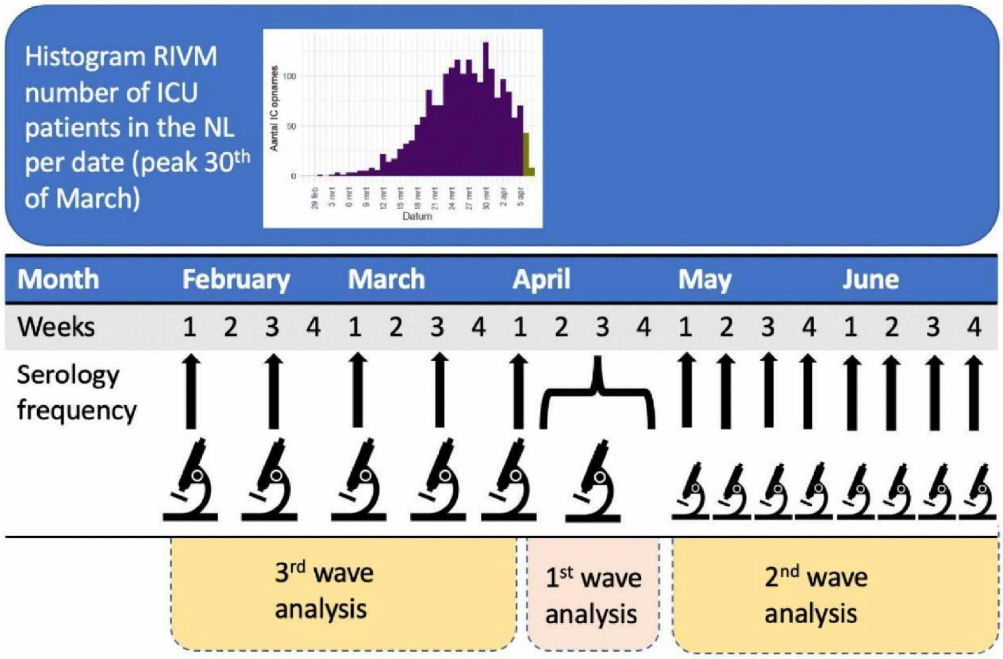
CoV-2 immunity status on the population (more details can be seen in Table 1 and Figure 3).

In order to be able to come with the most crucial information first we propose to do analysis in 2 waves:

1) Analysis of the samples that represent the peak of the admission rate of ICU cases in the Netherlands (see figure 2). These will be the 'first wave' analysis of 1000 samples in the month of April (2nd week till end of 4th week). This can be done in 1-3 days time in case there are enough serology tests available.
2) Analysis of the prospective samples per week in the month of May/June. To give up-to-date information on actual seroconversion status also in relation to political decisions (e.g. opening schools and other public facilities).
3) Analysis of the retrospective samples of February/March to evaluate the effect of several interventions (example: closing schools on March 16) and give a complete overview of the pandemic in the Netherlands. This analysis will be only done if resources are available.

More detailed information about the number of samples we are planning can be seen in section 'planned sampling'

Figure 2. Overview on strategy of analysis 1st till 3rd wave based on priority.



Legend: 1st wave (week 2 to 4) April samples providing actual seroconversion rates closely after the peak of the pandemic (see in histogram) in the Netherlands.[8] The 2nd wave: weekly samples in May/June to give 'real time' guidance. The 3rd wave: retrospective samples Feb/March

providing the whole time line of the pandemic in the Netherlands. One reference measurement will be done for the month of December (not shown in figure).

IgG & IgM serology

We will conduct SARS-CoV-2 antibody serology on second trimester serum blood residual material. Primary SARS-CoV-2 infection is characterized by the presence of detectable IgM antibodies with a median time of 12 days after the onset of infection. For IgG the median time is reported to be 14 days.[9]

At this moment (mid-April) the CE-marked serology tests is almost available. Due to its limited availability we are collaborating with Sanquin to use their 'in house' antibody screening serology test. The in-house antibody screening assay is a sensitive bridging assay that will detect antibodies of any isotype due to its design, thus enhancing the possibilities of detecting seroconverted samples. This test has been compared with the CE-marked serology test in almost 400 specimens already and has very similar (or even better) performance (see table 1). For this project, we will confirm every positive sample in the 'in house' test with the CE-marked serology test until the 'in house' test is more extensively validated (validation is ongoing).

Table 1. Sensitivity and Specificity of the Sanquin 'in house' COVID-19 antibody screen test (IgG/IgM).

		PCR confirmed		
		pos	neg	total
Antibody screen	pos	68	2	70
	neg	0	305	305
total		68	307	375
sensitivity:		**1**		
specificity:			**0.993485**	
diagnostic accuracy:				**0.994667**
*) one case borderline positive				

PCR confirmed cases are patients (8-19 days after onset symptoms) and convalescent sera.

In California, USA, for example, a test Cassette from the company 'Premier biotech' is used called 'COVID-19 IgG/IgM rapid Test Cassette on whole blood/serum/plasma'. The manufacturer reports a Sensitivity of IgM and IgG test of 91.8% (95% Confidence Intervals (CI): 83.8-96.6%) and 100.0% (95%CI 96.1-100.0%) respectively and a Specificity of 99.2% (95%CI 97.7-99.8%) and 99.5% (95%CI 98.1-99.9%) on a diagnostic study with over 400 specimens. Serology tests that we are going to use seem to have even better performance than the ones used in the USA seroprevalence studies.

As the test is not yet available in the Netherlands, the exact costs are not known. The test cassettes reported above cost $7 per piece. Inside information on the CE approved tests in the Netherlands are estimated at €20. This is a relatively inexpensive test compared to serology test available in the Netherlands, like CMV and Parvo IgG/IgM that have a cost of €30 (https://www.labmicta.nl/algemeen/tarieven/).

Power calculation samples needed

Estimate of expected prevalence (including assumptions that can change):
- 1,993 ICU patients have been admitted cumulatively as of 8 April 2020.[8]
- About 1: 5 of hospital admissions ends up at the ICU = 9,965 hospital admissions.
- Of the symptomatic patients, roughly 1:20 ends up in the hospital = 199,300.
- Chinese research (BMJ April 2020)[9] shows that 20%-50% of all COVID19 cases are symptomatic.[10]

Back calculation then gives approximately 400,000-1,000,000 infected people in the Netherlands (an estimated prevalence of ~ 3-6%). Not taking into account gender and age (i.e. more ICU admissions and hospital admissions in elderly people). A recent mathematical model of the RIVM shows a prevalence of 3.3% in women of the age 20-40 years old for the Netherlands (personal communication 5.1.2e, RIVM).
N.B. Nursing home patients are less likely to be admitted to hospitals, so the hospital admission data are an underestimate (not corrected for yet).
The first results from the serological study among Dutch blood donors showed a prevalence of 3.2%.[11]

Table 2. Sample Size for Frequency in a Population

Population size(for finite population correction factor or fpc)(*N*):	3500
Hypothesized % frequency of outcome factor in the population (*p*):	3%+/-1
Confidence limits as % of 100(absolute +/- %)(*d*):	1%
Design effect (for cluster surveys-*DEFF*):	1

Sample Size(*n*) for Various Confidence Levels

ConfidenceLevel(%)	Sample Size
95%	848
80%	421
90%	643
97%	986
99%	1245
99.9%	1659
99.99%	1951

Equation

Sample size $n = [DEFF*Np(1-p)]/ [(d^2/Z^2_{1-\alpha/2}*(N-1)+p*(1-p)]$

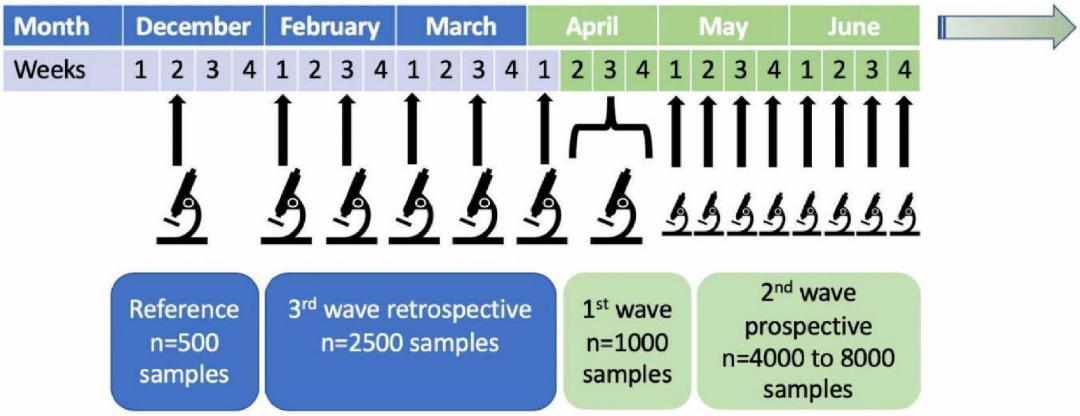
However, the above is based on assumptions that could change over time.
The needed number of samples per week is estimated to be 800-1000 samples in case of a prevalence of approximately 3% (with 95% CI of +/- 1%, see Table 2). If the prevalence is >3%, a sample size of 400-500 will be sufficient.[12]

Planned sampling

Sanquin has stored all blood samples of RhD negative women for a period of 6 months. We will be able to analyze approximately 90% of these samples for the period of February and March. For the month of April (2nd to 4th week) onwards (May and June) we will be able to use samples of both RhD negative as Rhc negative women as Sanquin will be able to store more blood of the RhD samples, and also store the Rhc samples (that are normally not stored) for the purpose of this study.

Volume needed for serology: Currently, ± 290 µl of plasma is being stored. This is sufficient to be able to repeat the serology test if necessary for any reason. Expected volumes needed: 30 ul in-house test, + possibly 100ul CE marked test + dead volume + extra to repeat → 290 µl = sufficient

Figure 3. Prospective (green) and retrospective (blue) data sample frequency over period December-June (or extended if necessary)



Prospective data:

1st wave

- April data (week 2 till 4) first to provide most needed up-to-date information
- April data: serology of Rhc negative and RhD negative women (33% of pregnant population= 170,000*0,33/52*3= 3,000 samples can be analyzed). Plan: we will analyze 1,000 random samples (prioritizing to analyze all available Rhd c negative samples first as these are more representative for the total population of pregnant women compared to the RhD negative samples that have an overrepresentation of Caucasian women).

- December will provide data from a reference population long before the first identified case in the Netherlands (27th of February). Serology of RhD negative women (14% of pregnant population= 170,000*0,14/12=1,900 samples/month and 450/week). We will analyze approximately 500 samples from December. This data/analysis is needed for validation of the test for this population.

2nd wave

- May and June 2
- Weekly samples
- Serology of Rhc negative and RhD negative women (33% of pregnant population= 170,000*0,33/52= 1,000 samples a week can be analyzed.

This dynamic structure will allow us to provide a unique 'real time' weekly estimate of the SARS-CoV-2 infection/seroconversion rate in the Netherlands. This repeated cross-sectional study will allow us to make more reliable estimates of the prevalence of SARS-CoV-2 IgG/IgM in the upcoming weeks, but more importantly, weekly changes in prevalence over time.

Retrospective data:

3rd wave

- February/March: serology of RhD negative women (14% of pregnant population= 170,000*0,14/12=1,900 samples/month). We will analyze 1,000 available samples for each month of which 500 samples taken from the first week of February and 500 samples taken in the third week of February. The same procedure will be done for March. This analysis can be run at the end of June-July due to the established period these samples should be stored for healthcare purposes (20 weeks after sampling).

This repeated cross-sectional study population will give us insight of the prevalence of SARS-CoV-2 IgG/IgM before and during the social distancing policy and at the peak of ICU admissions.

Statistical Analysis Plan

Data of IgG and/or IgM positive samples will be presented as descriptive numbers, providing a mean (with standard deviations) or median (with interquartile range) as appropriate. The denominator will be the total samples tested. This will provide a percentage of positive and negative tests results amongst the samples tested.

Extrapolation to estimate the number of infected people in the Netherlands.

A minimum dataset of 2 variables will be requested from 'RIVM-Centrum voor Bevolkingsonderzoek': 1) date of birth and 2) PC4 postal codes to guarantee anonymity of the pregnant women included in the sample. Postal codes are needed to provide insight in regional variation in seroprevalence (Brabant/Limburg and South-Holland are more seriously affected by the COVID-19 epidemic). Furthermore, the postal codes enable us to compare seroprevalence by municipality with numbers of COVID-19 cases

by municipality from the Osiris notification database. The birthyear is essential as this will enable to compare our project with other prevalence projects by age strata. According to the known prevalence of RhD negative (14%) and Rhc negative (19%) women[8], we can make a fair estimation of the test positives in the pregnant population. Additionally, known data from the Perinatal Registry Netherlands (PRN) will provide an overview of the age and ethnical distribution of the whole pregnant population and the national geographical coverage. Comparing this data with the Dutch age and ethnical distribution of the general population will allow us to give a weighted estimate on the prevalence of SARS-CoV-2 infections. We will present unadjusted and adjusted estimates with their 95% Confidence Intervals (95% CI).

Analysis on the insight in the role of children in the transmission. By knowing the pariety or indirectly the earlier blood sampling data on previous pregnancy (possible from 2011 onwards) we will be able to use this information on estimating the women that are pregnant on their first pregnancy (-nuli para- therefore no children at home) or a subsequent pregnancy (-multi para- and therefore very likely to have children living in the same household). A comparison between both groups can be made using linear regression analysis estimating the relation between nuli and multi parity on seroprevalence.

Sensitivity analysis will be performed to explore the impact of different factors on the results. For example, a potential higher false negative test results in pregnant population compared to non-pregnant population (something that is not known yet). We will use the Software package R for data analysis.

Strength and limitations

Most of the strength of using samples (and data) from the pregnant population are explained above. Our proposed approach also has several advantages compared to other initiatives that aim to estimate the prevalence of populations. For example the approach of the population of Iceland, in which people who accepted an offer of free testing were screened (n=5,500).[13] Or the cross-sectional PIENTER study in the Netherlands with at least two (possibly more) sampling points.[6] These approaches are based on potential biases sample of healthy people that volunteered to get an invasive blood sample test for this purpose only and may not have the ability to 'look back' on stored data nor provide weekly updated data. We will therefore be able to provide more and potentially at a faster speed the seroprevalence information than the above methods.

Another project that is proposed by Sanquin involves samples of individuals donating blood to the blood bank.[4] One possible drawback of that project is that blood donors include an overrepresentation of healthy people with a selective socio-economic status. These (and other factors) can influence their risk of getting exposed to SARS-CoV-2. The pregnant population is more diverse in relation to socio economic status and comorbidities. One example that illustrates this difference is the seroprevalence of Hepatitis B in blood donor (0.02) compared to pregnant population (0.28). This ten-fold difference in seroprevalence estimations using these populations illustrate the role of

potential selection bias in blood donors compared to pregnant population.[14] However, it should be noted that prevalence differences between the two populations can be smaller for respiratory infections compared to blood borne infections.

There are also some limitations to consider by using a pregnant population sample. First, it is a gender and age-range (15-45) specific population. These factors are only relevant in case gender and age has an impact on the risk of getting exposed to SARS-CoV-2. That is something we don't know. Second, it is also important to note that RhD negative blood type is more common amongst Caucasian population (15% Caucasian, 8% Black, 1% Asian populations).[15] This is less the case for Rhc negative blood types (20% Caucasian, 2% Black and 53% Asian), but it means that for both samples there is a relative underrepresentation of the Black population. This is an important thing to know in case the risk of exposure to SARS-CoV-2 is related to ethnical background. Third, one could argue that pregnant women practice more social distancing. However, pregnant women are not a special vulnerable population to SARS-CoV-2 infection, and we therefore don't think this will have an effect on the serology status.
Fourth, it is not known if pregnant women have a different immunologic response towards CoV-2 compared to a non-pregnant population. This may have an impact on especially a potential increase in false negative results which will lead to an underrepresentation of the overall infection rate. We will be able to correct for this when more basic understanding of CoV-2 infections in pregnant women are known. Until that time, we will explore different scenarios of different rates of false negative results in sensitivity analyses.

In summary, or main added value compared to the other prevalence initiatives in the Netherlands are:

1) Representation of all socio-economic-cultural layers of the society.
2) Continuous sampling that provides a 'real time' weekly information
3) Closely monitoring and analysis on the potential impact of children and measures like reopening of schools and childcare.

Cost-benefits

Costs

The costs of the COVID-19 serology test screening are yet to be defined. But when calculating the potential costs based on what we have heard about the serology tests that are going to be used soon in the Netherlands we estimate a total cost of 20 euro for the IgG and IgM test.

As seen in figure 2/3, there is difference in sample frequencies between the retrospective and prospective data samples. These periods are chosen to balance the costs of the samples with the benefit of having information over different time points.

First wave analysis:

- 1,000 samples month of April
- 20*1,000=20,000 euro

Second wave analysis:

- 2,000 samples month February/March = 40,000 euro
- 600 weekly samples month May/June =12,000 euro/week = 96,000 euro whole period

Total cost sampling: 20,000+40,000+96,000=**116,000 euro**

Benefits
The impact of COVID-19 has many sides, amongst them social, psychological, health related and economical. We will translate the potential economic impact in this analysis. A recent estimation. In March the 'Centraal Planbureau' published 9 potential scenarios' on COVID-19 impacting the economy. The range of loss in GDP reported in this file (https://www.cpb.nl/sites/default/files/omnidownload/CPB-Scenarios-maart-2020-Scenarios-economische-gevolgen-coronacrisis.pdf) varies from 1,2 to 7.7% (mean of 5%). Meaning, with a current GDP over 900 billion USD in the Netherlands (https://tradingeconomics.com/netherlands/gdp) there is a potential loss of 900*0.05=45 billion/year and 4 billion/month.

This research has the potential to give a fairly reliable infection fatality rate. In case the IFR is much lower than expected (there is some substantial proof from other studies and countries like Iceland), this could potentially impact decision making. If this decision making means that the economy can 'safely' open up again (meaning, opening restaurants and other trades that are currently on hold), and this study allow us to potential 'open' it 1 month earlier, this can potentially safe 4 billion.

This project also provides a structure to subsequently analyze other important hypothesis related to pregnancy and SARS-CoV-2 infection (e.g. vertical SARS-CoV-2 transmission linking maternal seroconversion and fetal blood samples, postnatal transmission, the rate of symptomatic/asymptomatic SARS-CoV-2 infected)

Time planning

End March-Beginning April:	Formation of Project group
First 2 weeks April:	Planning of logistics and application for funding
End April:	Start sampling prospective tests wave 1. Presentation and update results on website (RIVM)
May	Weekly updated of prospective samples May Results website RIVM. Publication MedrXiv
June	Weekly updated of prospective samples May Results website RIVM Publication MedrXiv
July	End sampling prospective and retrospective tests. Extension possible in case the 'real time' data is acknowledged to be crucial for 'real time' decision making.
September	Final publication of results

Budget

Manpower:	20 working weeks
Material:	COVID-19 IgG/IgM serology testing. See calculations above. Approximately 116,000 euro's.

-a more detailed buget description will follow-

References

1. CDC COVID-19 Response Team. Preliminary Estimates of the Prevalence of Selected Underlying Health Conditions Among Patients With Coronavirus Disease 2019 -United States, February 12-March 28, 2020. [https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7119513/pdf/mm6913e2.pdf]. Access date: April 13th 2020.
2. WHO Director-General's opening remarks at the media briefing on COVID-19 - 3 March 2020 - World Health Organization, March 3, 2020
3. Ioannidis J. A fiasco in the making? As the coronavirus pandemic takes hold, we are making decisions without reliable data. STAT. March 17, 2020.
4. Lit et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). Science 16 March 2020.
5. Sanquin test bloed op antistoffen tegen corona voor onderzoek naar groepsimmuniteit. Website [https://www.sanquin.org/nl/nieuws/2020/mrt/sanquin-test-bloed-op-antistoffen-tegen-corona?index=0] Access date: March 30th 2020.
6. PIENTER Corona study. Website [https://www.rivm.nl/en/pienter-corona-study] Access date: April 4 2020.
7. RIVM document 'rhesus bloedgroep tijdens zwangerschap' Website [https://www.rivm.nl/documenten/rhesus-bloedgroep-tijdens-zwangerschap]. Access date: March 30th.
8. 5.1.2e . COVID-19 Technische briefing Tweede Kamer. 8 april 2020. RIVM.[https://www.tweedekamer.nl/debat_en_vergadering/commissievergaderingen/details?id=2020A01496] Access date: April 10.
9. Anne-Genevieve Marcelin. Testing for SARS-CoV2 infection. Online educational Program. Paris, France.
10. Day M. Covid-19: four fifths of cases are asymptomatic, China figures indicate. BMJ 2 April 2020; 365.
11. 5.1.2e . Technische briefing Tweede Kamer 16 April 2020, https://www.tweedekamer.nl/debat_en_vergadering/commissievergaderingen/details?id=2020A01621. Access date: April 19 2020.
12. WHO. Population-based age-stratified seroepidemiological investigation protocol for COVID-19 virus infection. Version: 1.1 Date: 17 March 2020.
13. Gudbjartsson et al. Early Spread of SARS-Cov-2 in the Icelandic Population. MetRxiv preprint 26
14. RIVM document 'Sexual transmitted infections in the Netherlands in 2008.' [https://www.rivm.nl/bibliotheek/rapporten/2019-0007.pdf]. Access date: April 14 2020.
15. Reid, Lomas-Francis and Olson. Blood group antigen facts book, Reid 3rd edition. ISBN 9780240821306.

Appendix 1. Background information on the denominator problem.

In December 2019, Wuhan, a city in China became the center of an outbreak of pneumonia of unknown cause. In January, a SARS-CoV-2 virus was identified and the number of cases increased rapidly spreading globally as a pandemic. This pandemic faces many challenges on the level of clinical decision making facing resource problems (i.e. saturated ICU units) and also political measures (i.e. different strategies being applied from quarantine to social distancing in order to 'flatten the curve').

However, all this is done without reliable evidence on how many people have been infected with SARS-CoV-2. Now that we are in the middle of this pandemic, better information is needed to guide future decisions. One of this crucial information is the case fatality rate (risk of dying from SARS-CoV-2, also CFR). At this moment, the official CFR is 3.4%, reported early March by the World Health Organization.

This CRF is however much discussed as this is a crude estimate and it largely depends on our ability to capture infections that are in the denominator. To explain this, some more detailed information on how the CRF is measured is provided below.

The CRF is based on the number of people who had died on the nominator and the number of people who had documented infection as the denominator. This estimation can have several problems on the nominator side as well as the denominator side.

Problems on the nominator side cab be 1) capturing number of deaths *by* COVID-19 – (causal relation) or *with* COVID-19 (association); 2) missing numbers of death that occur outside the hospital; 3) some infected people now may die later.

However, the biggest uncertainty is the denominator. This number is based on the number of patients tested for the presence of COVID-19, irrespective of their symptoms. At this moment, we know that many countries are only testing people that 1) present themselves to a health facility or 2) present themselves with severe symptoms justifying hospital admittance.

Based on what we know now, many people infected by COVID-19 present with mild or no symptoms that are very difficult to distinguish from the common flue. So many of them will not present themselves at a health care facility or even being refused to test due to lack of testing resources. So the number of documented infections we know is just the tip of the iceberg.

Information from setting where we have more complete information about that denominator (for example the Princes Diamond cruise ship, or data from Iceland where a huge random sample is being tested) suggest that the infection fatality rate (IFR) is much lower than 3.4%.

Having a more accurate denominator in a population is crucial for guidance on modeling the measures hospitals and politics should be taken going forward in this pandemic.