# Bayes Lines Tool (BLT) - A SQL-script for analyzing diagnostic test results with an application to SARS-CoV-2-testing

Wouter Aukema [1*], Ulrike Kämmerer [2], Pieter Borger [3], Simon Goddek [4], Bobby Rajesh Malhotra [5],
Kevin McKernan [6], Rainer J. Klement [7*]

[1] Independent Data and Pattern Scientist, Brinkenbergweg 1, 7351BD, Hoenderloo, The Netherlands; Email: wouter@aukema.org;
Phone: +31626546406

[2] Department OB/Gyn, University Hospital of Würzburg, Josef-Schneider-Str.4, D-97080 Würzburg, Germany; Email: frak057@mail.uni-wuerzburg.de. ORCID ID: https://orcid.org/0000-0002-2311-6984

[3] The Independent Research Initiative on Information & Origins, 79540 Loerrach, Germany; Email: peterborger@hotmail.com

[4] Independent Scientist, Elias Beeckmanlaan 242, 6711 VS, Ede, The Netherlands; Email: simon@goddek.nl

[5] Department for Digital Arts, University for Applied Arts Vienna, Expositur Hintere Zollamtsstraße 17, 1030, Vienna, Austria; Email:
bobby.rajesh.malhotra@gmail.com

[6] Chief Scientific Officer, Medicinal Genomics, 100 Cummings Center, 406L, Beverly MA 01915, USA Email:
Kevin.McKernan@medicinalgenomics.com . ORCID ID: https://orcid.org/0000-0002-3908-1122

[7] Department of Radiotherapy and Radiation Oncology, Leopoldina Hospital Schweinfurt, Robert-Koch-Straße 10, 97422, Schweinfurt,
Germany; Email: rainer_klement@gmx.de; Phone: +49 9721 7202761; ORCID ID: https://orcid.org/0000-0003-1401-4270

* Corresponding authors

## Abstract

The performance of diagnostic tests crucially depends on the disease prevalence, test sensitivity, and test specificity. However, these quantities are often not well known when tests are performed outside clinical practice which makes the rating of the test results somewhat problematic. A current example is the mass testing taking place within the context of the world-wide SARS-CoV-2 crisis. Here, for the first time in history, the test results have a dramatic impact on political decisions. Therefore, transparent, comprehensible, and reliable data is mandatory. It is in the nature of wet lab tests that their quality and outcome are influenced by multiple factors reducing their performance by handling procedures, underlying test protocols, and analytical reagents. These limitations in sensitivity and specificity have to be taken into account when calculating the real test results. As a resolution method, we have developed a seminal Bayesian calculator, the Bayes Lines Tool (BLT), for back-solving disease prevalence, test sensitivity, test specificity, and, therefore, true positive, false positive, true negative and false negative numbers, from official test outcome reports. The calculator performs a simple

SQL query and can easily be implemented on any system supporting SQL. We provide three examples of SARS-CoV-2 test results from official government reports from the Netherlands, Germany, and the United Kingdom to illustrate the possible parameter space of prevalence, sensitivity, and specificity consistent with the observed data. Finally, we discuss this tool's multiple applications, including its putative importance for informing policy decisions.

**Keywords:** Bayes; COVID19; PCR Test; SARS-CoV-2; SQL

## 1. Introduction

In December 2019, a cluster of patients with pneumonia of unknown origin was associated with the emergence of a novel beta-coronavirus of bat origin (1), first named 2019-nCoV (2) and later specified as severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2) (3). This outbreak led to the rapid development of reverse transcriptase – quantitative polymerase chain reaction (RT-qPCR) tests to identify SARS-CoV-2 RNA in specimens obtained from patients (2,4).

After sporadic SARS-CoV-2 positive cases in January (5,6), from the end of February 2020 worldwide cases of the SARS-CoV-2-associated disease 'COVID-19' began to accumulate, causing policymakers in many countries to introduce countermeasures. These Non-Pharmaceutical Interventions (NPIs) predominantly started worldwide around March 2020 while the virus was characterized as a pandemic on 11 March 2020 (6,7). As a result, for almost one year now, large parts of the world are in a COVID-19 crisis-mode with daily reporting of SARS-CoV-2 cases in dashboards worldwide (8). The definition of 'cases' and 'prevalence estimates' was based on RT-qPCR testing, independent of the clinical diagnosis. Thereby, a person is considered a case (i.e. infected), once a test turns out positive (9).

Like all laboratory tests, however, the SARS-CoV-2 RT-qPCR tests are not flawless. This is because sensitivity and specificity depend on a multiplicity of confounding factors. These factors cover the test design, the lab application, and possible contaminations with substances/nucleic acids interfering with the reaction (10,11). Consequently, both false-negative and false-positive results have been reported (12,13). Nevertheless, the test system's limitations are rarely discussed in scientific publications and public health systems despite their crucial role for making inferences about the possible infection status of a tested person (14). Many more or less defined commercial and laboratory 'in house' tests are now routinely being used (15), often without standardised guidelines, which leads to entirely unknown test performance specifications (16). The few studies aiming to estimate sensitivity and specificity of SARS-CoV-2 RT-qPCR tests have reported sensitivities and specificities in the ranges ≳30% and ≳80%, respectively - therefore, the communicated data seldom can offer precise distinctions (14).

Given the critical role that dashboards and graphs based on SARS-CoV-2 test results play for policymakers, health professionals and the general public (8), our objective was to develop a Bayesian calculator that could calculate test quantities and prevalence solely based on officially reported numbers of total and positive tests, i.e. without making any *a priori* assumptions. In this way, time trend estimates and country-to-country comparisons of these test performance measures as well as disease prevalence estimates become possible, producing in-depth insights, making projections/ simulations possible, and providing a more holistic understanding of the daily incoming data in general.

## 2. Materials and Methods

### 2.1 General Description of the Calculator

The Bayes Lines Tool (BLT) calculator is based on Bayes' theorem and estimates the true and false positive and true and false negative numbers at a given time point for which the total number of tests performed and the number of positive test results is known. These three data points are usually reported and published by official government bodies daily and/or weekly. Thus, the model uses the following information:

    I.    Publishing date or report identifier of the test data
    II.    Number of performed tests
    III.    Number of reported positive results

The model takes this information as a given fact and uses it to make inferences about the test performance parameters (sensitivity and specificity) as well as the prevalence (also known as the base rate) - these inferences are essential for estimating the number of true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) (14). It is assumed that there is no knowledge of either the prevalence or the sensitivity and specificity of the tests used. Instead, the model explores all possible combinations of these three parameters within reasonable ranges specified by the user, and selects all combinations that result in TP+FP estimates consistent with the known number of positive tests. For the implementation presented here, the three parameters are varied as follows:

- Prevalence from 0.005 to 0.50 with 0.005 increments;
- Sensitivity from 0.30 to 0.999 with 0.005 increments;
- Specificity from 0.75 to 0.999 with 0.005 increments.

For each possible combination of prevalence, sensitivity and specificity, a Bayesian confusion matrix (CM) is calculated and stored. A single CM contains TP, FP, TN and FN in absolute numbers (Table 1).

*Table 1.* A confusion matrix for a SARS-CoV-2 test containing absolute numbers of true (TP) and false (FP)  positives and true (TN) and false (FN) negatives derived from equations (2)-(5).

| Actual infection status | Test result positive | Test result negative |
|---|---|---|
| INFECTED | TP | FN |
| NOT INFECTED | FP | TN |

For a given prevalence, sensitivity and specificity these are derived from Bayes' theorem:

$$P(I|T) = \frac{P(T|I) \times P(I)}{P(T)}$$

(1)

Here, $T$ denotes the hypothesis that a test comes out positive ($\neg T$its denial) and $I$ the hypothesis that an individual is infected so that $P(I)$is the prevalence and $P(T|I)$ is the test sensitivity. $P(T)$is the marginal probability of a positive test, given as the frequency of positive test results, whereas $P(I|T)$is the probability of being infected given that the test came out positive. With the normalizing constant $P(T)$ given as $P(T) = \frac{\text{\# positive tests}}{\text{\# tests}}$and $P(I|T)$given as the proportion of infected individuals among those in which the test came out positive, equation (1) becomes:

$$TP = P(I|T) \times \text{\# positive tests} = \text{sensitivity} \times \text{prevalence} \times \text{\# tests}$$

(2)

Equation (2) thus shows that the number of TPs depends on the prevalence, test sensitivity and total number of tests performed. Using $P(\neg T|\neg I)$=specificity, $P(T|\neg I)$=1–specificity, and$P(\neg T|I)$=1–sensitivity, an analogous derivation leads to

$$FP = P(\neg I|T) \times \text{\# positive tests} = (1 - \text{specificity}) \times (1 - \text{prevalence}) \times \text{\# tests}$$

(3)

$$TN = P(\neg I|\neg T) \times \# \; negative \; tests = specificity \times (1 - prevalence) \times \# \; tests \tag{4}$$

$$FN = P(I|\neg T) \times \# \; negative \; tests = (1 - sensitivity) \times prevalence \times \# \; tests \tag{5}$$

## 2.2 Implementation

We developed an SQL (Structured Query Language) query that generates all possible Bayesian CMs for a (series of) diagnostic test results, without making assumptions about prevalence, sensitivity or specificity.

The code in standard SQL is given as follows (Code 1):

```sql
with tests as
(
select
        :report_id::text as report_id, -- Feel free to call with any string, e.g. date-string
        :tests as tests_performed, -- integer
        :cases as positives_reported --integer
),

permutations as
(
select
        (prevalence::numeric / 1000)::numeric as prevalence,
        (sensitivity::numeric / 1000)::numeric as sensitivity,
        (specificity::numeric / 1000)::numeric as specificity
from
        generate_series(1, 500, 1) as prevalence,
        generate_series(30, 999, 5) as sensitivity,
        generate_series(75, 999, 5) as specificity
),

matrices as
(
select
        t.report_id,
        t.tests_performed,
        t.positives_reported,
        round(prevalence, 2) as prevalence, --just for cosmetic purposes
        round(sensitivity, 3) as sensitivity,
        round(specificity, 3) as specificity,
        (t.tests_performed * prevalence)::int as has_disease, --calculation with full precision
for data type numeric, but
        (t.tests_performed * (1 - prevalence))::int as hasnot_disease, --casting to integer for
cosmetic purposes
        (t.tests_performed * prevalence * sensitivity)::int as true_positives,
        (t.tests_performed * (1 - prevalence) * specificity)::int as true_negatives
from
```

```
        tests t,
        permutations p
)

select
        *,
        hasnot_disease - true_negatives as false_positives,
        has_disease - true_positives as false_negatives
from
        matrices
where
        (true_positives + (hasnot_disease - true_negatives)) = positives_reported
order by
        report_id,
        prevalence,
        sensitivity,
        specificity
```

In plain language, this code says: Given the test results and given all possible permutations and consequently all possible CMs, only return those CMs that match the test results. We make no assumptions about any of the three variables in the permutations, therefore the query considers all of them. Only with the resulting CMs that match the input data, we can perhaps identify patterns that provide insights for further investigation.

In order to produce CMs for a series of reports, such as daily test result numbers, several approaches are possible. For performance purposes, we advise using a Batch/Script approach (Code 2):

```
psql -t -h localhost -d postgres -U postgres -A --set=report_id='2021-01-10'  --set=tests=28757 -
-set=cases=3829 -f generate_matrices.sql > 2021-01-10.out
psql -t -h localhost -d postgres -U postgres -A --set=report_id='2021-01-11'  --set=tests=43944 -
-set=cases=5392 -f generate_matrices.sql > 2021-01-11.out
psql -t -h localhost -d postgres -U postgres -A --set=report_id='2021-01-12'  --set=tests=40643 -
-set=cases=4923 -f generate_matrices.sql > 2021-01-12.out
psql -t -h localhost -d postgres -U postgres -A --set=report_id='2021-01-13'  --set=tests=47387 -
-set=cases=5098 -f generate_matrices.sql > 2021-01-13.out
psql -t -h localhost -d postgres -U postgres -A --set=report_id='2021-01-14'  --set=tests=48412 -
-set=cases=5105 -f generate_matrices.sql > 2021-01-14.out
psql -t -h localhost -d postgres -U postgres -A --set=report_id='2021-01-15'  --set=tests=46015 -
-set=cases=4736 -f generate_matrices.sql > 2021-01-15.out
psql -t -h localhost -d postgres -U postgres -A --set=report_id='2021-01-16'  --set=tests=36880 -
-set=cases=3695 -f generate_matrices.sql > 2021-01-16.out
```

This produces ".out" text files that the user can pick up and merge together according to her or his own preferences.

The performance of the query depends on the three statements that determine the ranges and step sizes by which prevalence, sensitivity and specificity are varied. For example, the following SQL query results in 17,945,000 permutations (Code 3):

```sql
select
        count(*)
from
        generate_series(1, 500, 1) as prevalence,
        generate_series(30, 999, 5) as sensitivity,
        generate_series(75, 999, 5) as specificity
```

If a user's computer cannot handle the number of iterations, she/he can narrow the range, for instance, by increasing step size from 1 to 5 or even higher. This will reduce resolution and the corresponding number of matrices that match the input.

When narrowing the step size in the permutations, we recommend to widen the scope in the WHERE clause at the end. The example below provides a 0.1% margin in the match. If you increase step-size, you should widen the margins too (Code 4).

```sql
where
        (true_positives + (hasnot_disease - true_negatives)) between
0.999*positives_reported::numeric and 1.001*positives_reported::numeric
```

The following example shows how to perform the above SQL query against a PostgreSQL instance from a MS-DOS Command Prompt and displays the result (Code 5):

```
psql -t -h localhost -d postgres -U postgres -A --set=report_id='2021-01-10'  --set=tests=28757 -
-set=cases=3829 -f generate_matrices.sql

report_id|tests_performed|positives_reported|prevalence|sensitivity|specificity|has_disease|hasnot
_disease|true_positives|true_negatives|false_positives|false_negatives
2021-01-10|28757|3829|0.010|0.980|0.875|275|28482|269|24922|3560|6
2021-01-10|28757|3829|0.019|0.040|0.865|554|28203|22|24396|3807|532
2021-01-10|28757|3829|0.025|0.060|0.865|711|28046|43|24260|3786|668
2021-01-10|28757|3829|0.025|0.850|0.885|711|28046|604|24821|3225|107
2021-01-10|28757|3829|0.027|0.780|0.885|786|27971|613|24755|3216|173
2021-01-10|28757|3829|0.029|0.070|0.865|826|27931|58|24160|3771|768
2021-01-10|28757|3829|0.033|0.080|0.865|960|27797|77|24045|3752|883
2021-01-10|28757|3829|0.033|0.370|0.875|960|27797|355|24323|3474|605
2021-01-10|28757|3829|0.033|0.660|0.885|960|27797|633|24601|3196|327
2021-01-10|28757|3829|0.041|0.090|0.865|1172|27585|105|23861|3724|1067
2021-01-10|28757|3829|0.047|0.700|0.895|1362|27395|953|24519|2876|409
2021-01-10|28757|3829|0.050|0.480|0.885|1432|27325|687|24183|3142|745
2021-01-10|28757|3829|0.055|0.960|0.915|1582|27175|1519|24865|2310|63
2021-01-10|28757|3829|0.074|0.110|0.865|2136|26621|235|23027|3594|1901
2021-01-10|28757|3829|0.082|0.560|0.905|2361|26396|1322|23889|2507|1039
2021-01-10|28757|3829|0.095|0.590|0.915|2743|26014|1618|23803|2211|1125
2021-01-10|28757|3829|0.095|0.780|0.935|2743|26014|2139|24324|1690|604
2021-01-10|28757|3829|0.111|0.680|0.935|3186|25571|2167|23909|1662|1019
2021-01-10|28757|3829|0.111|0.760|0.945|3186|25571|2422|24164|1407|764
2021-01-10|28757|3829|0.157|0.820|0.995|4522|24235|3708|24114|121|814
2021-01-10|28757|3829|0.333|0.390|0.995|9572|19185|3733|19089|96|5839
```

```
2021-01-10|28757|3829|0.368|0.130|0.865|10579|18178|1375|15724|2454|9204
2021-01-10|28757|3829|0.497|0.040|0.775|14280|14477|571|11219|3258|13709
(23 rows)
```

The above results show all possible Bayes CMs, given 28,757 tests and 3,829 positive outcomes without making any assumption for prevalence, sensitivity or specificity. These results can be used in any tool of choice to visualize or evaluate. Examples will be given in the Results section.

**2.3 Data**

In its current form, the SQL query takes three parameters for which only two are being used in the generation of the matrices. For the examples demonstrated in the Results section below, we extracted SARS-CoV-19 test data from the Dutch Corona Dashboard database (https://coronadashboard.rijksoverheid.nl/landelijk/positief-geteste-mensen), the German RKI database (https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/Jan_2021/2021-01-20-de.pdf?__blob=publicationFile) and two UK data sources (https://coronavirus.data.gov.uk/details/testing, https://coronavirus.data.gov.uk/details/cases).

# 3. Results

In the following section we provide three examples that demonstrate the application of our calculator for The Netherlands, Germany and The United Kingdom.

**3.1 Example 1: The Dutch Corona Dashboard data**

SARS-CoV-2 test data have been extracted for January 10th 2021. There were 28,757 tests with 3,829 positive results. The calculator yielded 23 possible solutions that matched the number of positive tests. These were simply pasted into an MS Excel sheet to produce the plot displayed in Figure 1. The results are also and tabulated in Table 2. Most solutions favored a prevalence below 10%, but in these cases more false than true positives would have been measured. There were two solutions, one for a prevalence of 15.7%, the other for 33.3%, where the majority of positive tests would have been TPs.

**Figure 1.** Example showing all possible confusion matrices for one report from the Dutch Corona dashboard at 10-jan-2021, true positives and false positives. Note that prevalence increases non-linearly because only the solutions fitting the observed data are plotted.
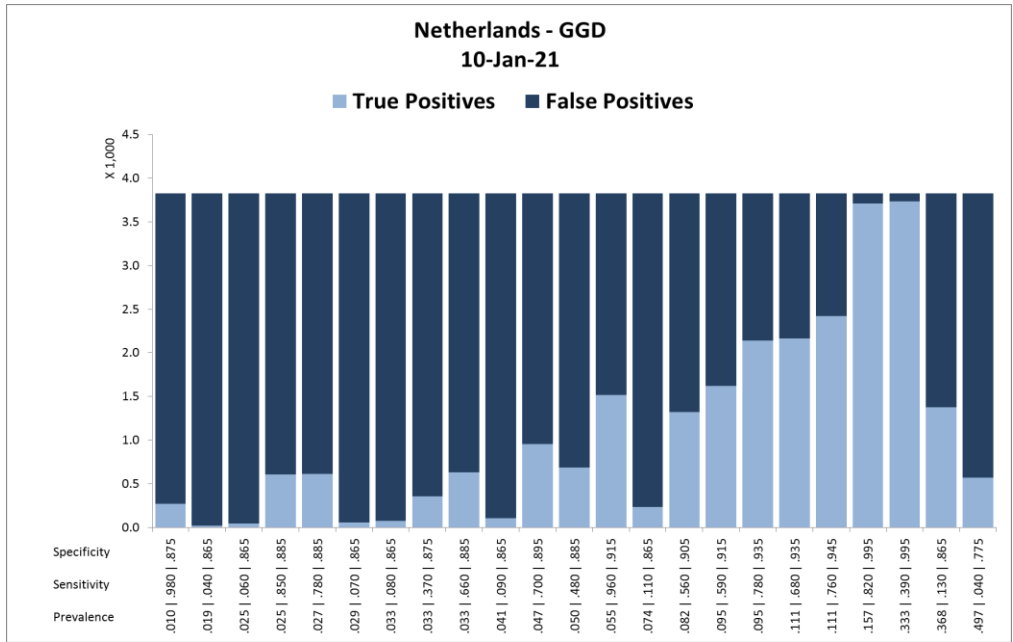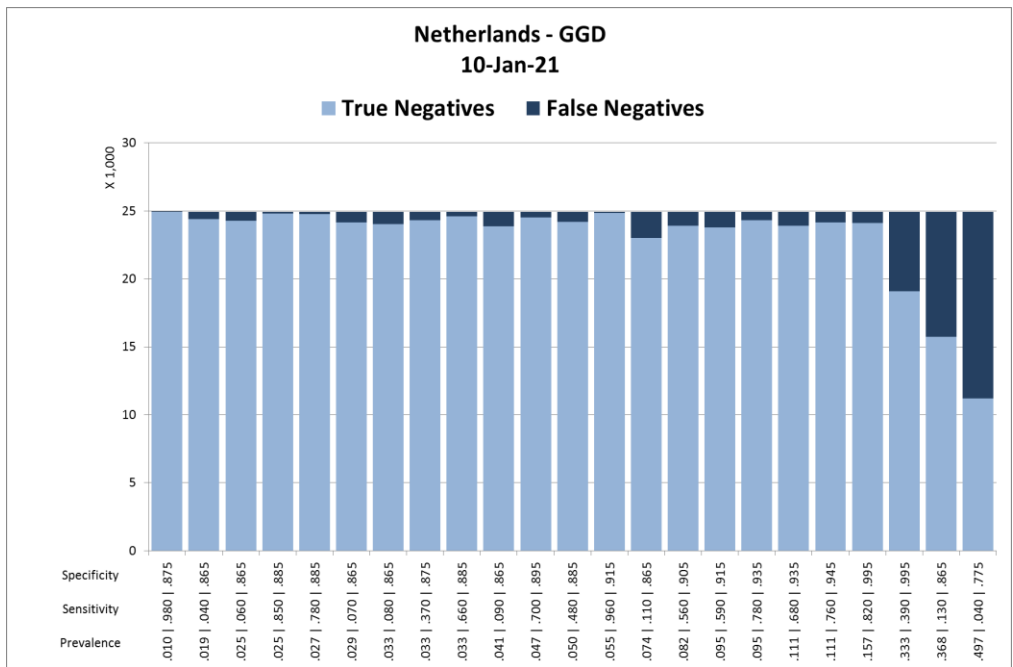
**Figure 2.** Example showing all possible confusion matrices for one report from the Dutch Corona dashboard at 10-jan-2021, true positives and false negatives. Note that prevalence increases non-linearly because only the solutions fitting the observed data are plotted.

**Table 2.** Dutch Corona Dashboard spreadsheet for January 10th 2021.

| report_date | tests_performed | positives_reported | prevalence | sensitivity | specificity | has_disease | hasnot_disease | true_positives | true_negatives | false_positives | false_negatives |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 10-Jan-21 | 28 757 | 3 829 | 1,0% | 98,0% | 87,5% | 275 | 28 482 | 269 | 24 922 | 3 560 | 6 |
| 10-Jan-21 | 28 757 | 3 829 | 1,9% | 4,0% | 86,5% | 554 | 28 203 | 22 | 24 396 | 3 807 | 532 |
| 10-Jan-21 | 28 757 | 3 829 | 2,5% | 6,0% | 86,5% | 711 | 28 046 | 43 | 24 260 | 3 786 | 668 |
| 10-Jan-21 | 28 757 | 3 829 | 2,5% | 85,0% | 88,5% | 711 | 28 046 | 604 | 24 821 | 3 225 | 107 |
| 10-Jan-21 | 28 757 | 3 829 | 2,7% | 78,0% | 88,5% | 786 | 27 971 | 613 | 24 755 | 3 216 | 173 |
| 10-Jan-21 | 28 757 | 3 829 | 2,9% | 7,0% | 86,5% | 826 | 27 931 | 58 | 24 160 | 3 771 | 768 |
| 10-Jan-21 | 28 757 | 3 829 | 3,3% | 8,0% | 86,5% | 960 | 27 797 | 77 | 24 045 | 3 752 | 883 |
| 10-Jan-21 | 28 757 | 3 829 | 3,3% | 37,0% | 87,5% | 960 | 27 797 | 355 | 24 323 | 3 474 | 605 |
| 10-Jan-21 | 28 757 | 3 829 | 3,3% | 66,0% | 88,5% | 960 | 27 797 | 633 | 24 601 | 3 196 | 327 |
| 10-Jan-21 | 28 757 | 3 829 | 4,1% | 9,0% | 86,5% | 1 172 | 27 585 | 105 | 23 861 | 3 724 | 1 067 |
| 10-Jan-21 | 28 757 | 3 829 | 4,7% | 70,0% | 89,5% | 1 362 | 27 395 | 953 | 24 519 | 2 876 | 409 |
| 10-Jan-21 | 28 757 | 3 829 | 5,0% | 48,0% | 88,5% | 1 432 | 27 325 | 687 | 24 183 | 3 142 | 745 |
| 10-Jan-21 | 28 757 | 3 829 | 5,5% | 96,0% | 91,5% | 1 582 | 27 175 | 1 519 | 24 865 | 2 310 | 63 |
| 10-Jan-21 | 28 757 | 3 829 | 7,4% | 11,0% | 86,5% | 2 136 | 26 621 | 235 | 23 027 | 3 594 | 1 901 |
| 10-Jan-21 | 28 757 | 3 829 | 8,2% | 56,0% | 90,5% | 2 361 | 26 396 | 1 322 | 23 889 | 2 507 | 1 039 |
| 10-Jan-21 | 28 757 | 3 829 | 9,5% | 59,0% | 91,5% | 2 743 | 26 014 | 1 618 | 23 803 | 2 211 | 1 125 |
| 10-Jan-21 | 28 757 | 3 829 | 9,5% | 78,0% | 93,5% | 2 743 | 26 014 | 2 139 | 24 324 | 1 690 | 604 |
| 10-Jan-21 | 28 757 | 3 829 | 11,1% | 68,0% | 93,5% | 3 186 | 25 571 | 2 167 | 23 909 | 1 662 | 1 019 |
| 10-Jan-21 | 28 757 | 3 829 | 11,1% | 76,0% | 94,5% | 3 186 | 25 571 | 2 422 | 24 164 | 1 407 | 764 |
| 10-Jan-21 | 28 757 | 3 829 | 15,7% | 82,0% | 99,5% | 4 522 | 24 235 | 3 708 | 24 114 | 121 | 814 |
| 10-Jan-21 | 28 757 | 3 829 | 33,3% | 39,0% | 99,5% | 9 572 | 19 185 | 3 733 | 19 089 | 96 | 5 839 |
| 10-Jan-21 | 28 757 | 3 829 | 36,8% | 13,0% | 86,5% | 10 579 | 18 178 | 1 375 | 15 724 | 2 454 | 9 204 |
| 10-Jan-21 | 28 757 | 3 829 | 49,7% | 4,0% | 77,5% | 14 280 | 14 477 | 571 | 11 219 | 3 258 | 13 709 |

## 3.2 Example 2: The German RKI data

This example, plotted in Figure 2 and tabulated in Table 3, shows the possible results obtained for week 2 in 2021 published by the German government organization Robert-Koch Institute (RKI). The time to obtain all CMs fitting the data was 1m42s on an IMac using PostgreSQL version 12 with no performance tuning, so the process was limited to 11MB ram usage. It is interesting to see that compared to the Dutch example (Figure 1), there is now a larger number of possible combinations consistent with the data, and FP numbers seem less dramatic in most

of the CMs generated. However, a general trend of increasing FP number with decreasing prevalence can also be seen in this example. Note that now prevalence, sensitivity and specificity are regarded as weekly averages, while in the Dutch example they are specific to one day (January 10th 2021).
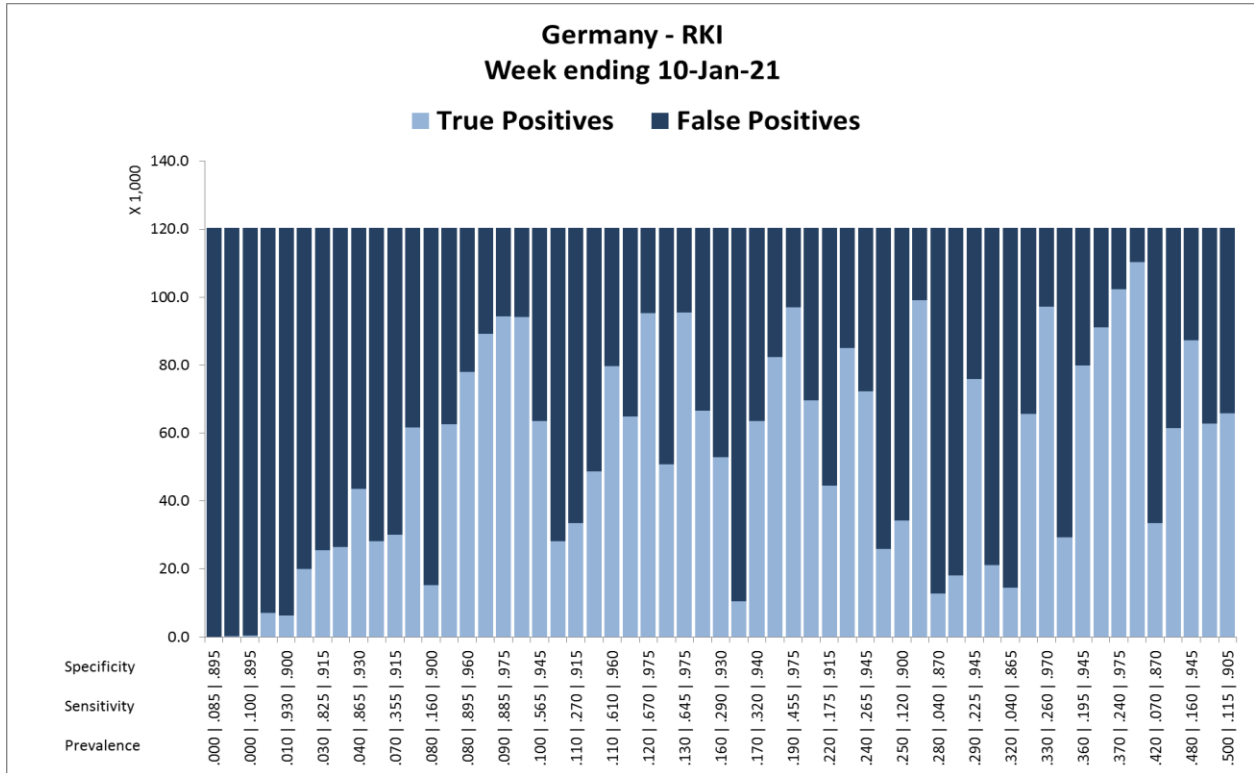


**Figure 3.** Results for German RKI data for week 2 in 2021, true and false positives. Note that prevalence increases non-linearly because only the solutions fitting the observed data are plotted.
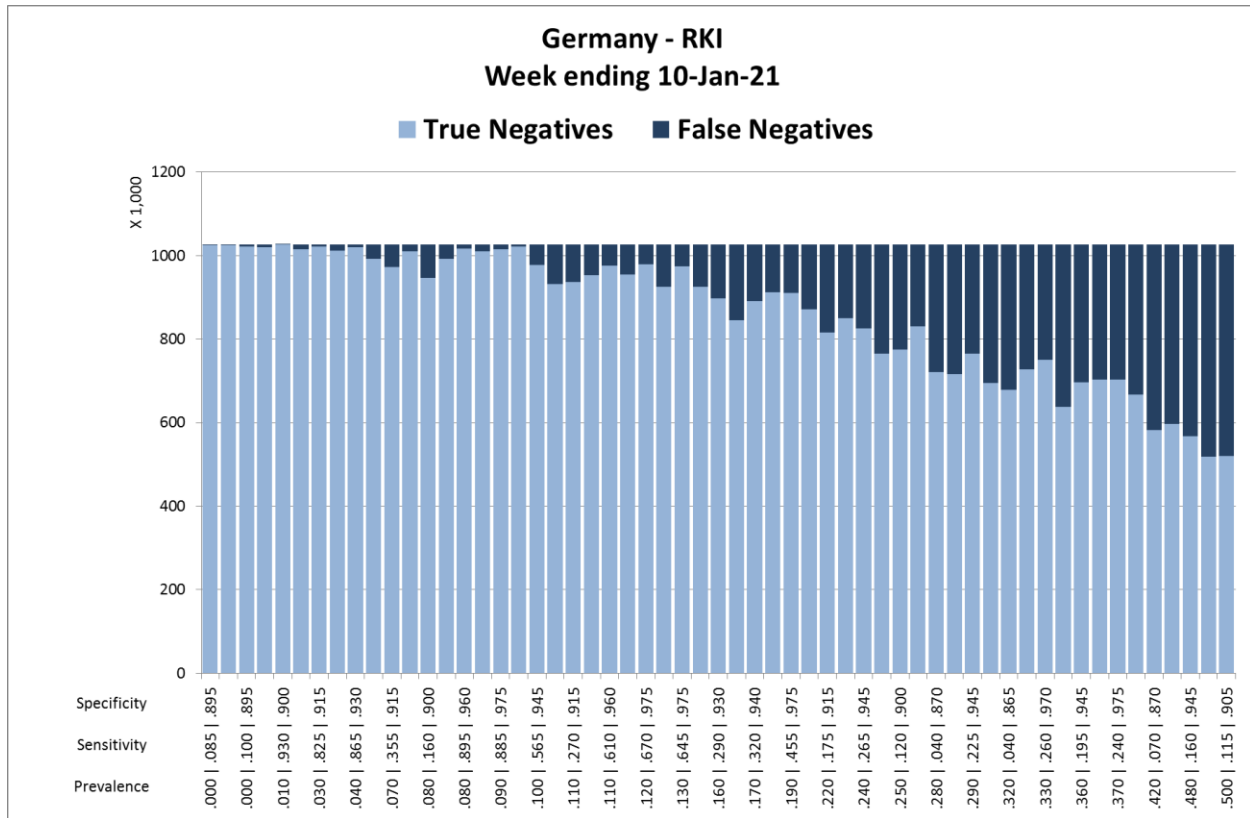
**Figure 4.** Results for German RKI data for week 2 in 2021, true and false negatives. Note that prevalence increases non-linearly because only the solutions fitting the observed data are plotted.

**Table 3.** Results for German RKI data for week 2 in 2021 (17).

| report_date | Tests_performed | positives_reported | prevalence | sensitivity | specificity | has_disease | hasnot_disease | true_positives | true_negatives | false_positives | false_negatives |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 10-Jan-21 | 1 146 983 | 120 410 | 0,0% | 8,5% | 89,5% | 1 147 | 1 145 836 | 97 | 1 025 523 | 120 313 | 1 050 |
| 10-Jan-21 | 1 146 983 | 120 410 | 0,0% | 9,5% | 89,5% | 2 294 | 1 144 689 | 218 | 1 024 497 | 120 192 | 2 076 |
| 10-Jan-21 | 1 146 983 | 120 410 | 0,0% | 10,0% | 89,5% | 4 588 | 1 142 395 | 459 | 1 022 444 | 119 951 | 4 129 |
| 10-Jan-21 | 1 146 983 | 120 410 | 1,0% | 51,5% | 90,0% | 13 764 | 1 133 219 | 7 088 | 1 019 897 | 113 322 | 6 676 |
| 10-Jan-21 | 1 146 983 | 120 410 | 1,0% | 93,0% | 90,0% | 6 882 | 1 140 101 | 6 400 | 1 026 091 | 114 010 | 482 |
| 10-Jan-21 | 1 146 983 | 120 410 | 3,0% | 62,5% | 91,0% | 32 116 | 1 114 867 | 20 072 | 1 014 529 | 100 338 | 12 044 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 10-Jan-21 | 1 146 983 | 120 410 | 3,0% | 82,5% | 91,5% | 30 969 | 1 116 014 | 25 549 | 1 021 153 | 94 861 | 5 420 |
| 10-Jan-21 | 1 146 983 | 120 410 | 4,0% | 64,0% | 91,5% | 41 291 | 1 105 692 | 26 426 | 1 011 708 | 93 984 | 14 865 |
| 10-Jan-21 | 1 146 983 | 120 410 | 4,0% | 86,5% | 93,0% | 50 467 | 1 096 516 | 43 654 | 1 019 760 | 76 756 | 6 813 |
| 10-Jan-21 | 1 146 983 | 120 410 | 5,0% | 45,5% | 91,5% | 61 937 | 1 085 046 | 28 181 | 992 817 | 92 229 | 33 756 |
| 10-Jan-21 | 1 146 983 | 120 410 | 7,0% | 35,5% | 91,5% | 84 877 | 1 062 106 | 30 131 | 971 827 | 90 279 | 54 746 |
| 10-Jan-21 | 1 146 983 | 120 410 | 7,0% | 79,0% | 94,5% | 77 995 | 1 068 988 | 61 616 | 1 010 194 | 58 794 | 16 379 |
| 10-Jan-21 | 1 146 983 | 120 410 | 8,0% | 16,0% | 90,0% | 95 200 | 1 051 783 | 15 232 | 946 605 | 105 178 | 79 968 |
| 10-Jan-21 | 1 146 983 | 120 410 | 8,0% | 65,0% | 94,5% | 96 347 | 1 050 636 | 62 625 | 992 851 | 57 785 | 33 722 |
| 10-Jan-21 | 1 146 983 | 120 410 | 8,0% | 89,5% | 96,0% | 87 171 | 1 059 812 | 78 018 | 1 017 420 | 42 392 | 9 153 |
| 10-Jan-21 | 1 146 983 | 120 410 | 9,0% | 84,5% | 97,0% | 105 522 | 1 041 461 | 89 166 | 1 010 217 | 31 244 | 16 356 |
| 10-Jan-21 | 1 146 983 | 120 410 | 9,0% | 88,5% | 97,5% | 106 669 | 1 040 314 | 94 402 | 1 014 306 | 26 008 | 12 267 |
| 10-Jan-21 | 1 146 983 | 120 410 | 9,0% | 95,5% | 97,5% | 98 641 | 1 048 342 | 94 202 | 1 022 134 | 26 208 | 4 439 |
| 10-Jan-21 | 1 146 983 | 120 410 | 10,0% | 56,5% | 94,5% | 112 404 | 1 034 579 | 63 508 | 977 677 | 56 902 | 48 896 |
| 10-Jan-21 | 1 146 983 | 120 410 | 11,0% | 23,0% | 91,0% | 122 727 | 1 024 256 | 28 227 | 932 073 | 92 183 | 94 500 |
| 10-Jan-21 | 1 146 983 | 120 410 | 11,0% | 27,0% | 91,5% | 123 874 | 1 023 109 | 33 446 | 936 145 | 86 964 | 90 428 |
| 10-Jan-21 | 1 146 983 | 120 410 | 11,0% | 40,0% | 93,0% | 121 580 | 1 025 403 | 48 632 | 953 625 | 71 778 | 72 948 |
| 10-Jan-21 | 1 146 983 | 120 410 | 11,0% | 61,0% | 96,0% | 130 756 | 1 016 227 | 79 761 | 975 578 | 40 649 | 50 995 |
| 10-Jan-21 | 1 146 983 | 120 410 | 12,0% | 47,5% | 94,5% | 136 491 | 1 010 492 | 64 833 | 954 915 | 55 577 | 71 658 |
| 10-Jan-21 | 1 146 983 | 120 410 | 12,0% | 67,0% | 97,5% | 142 226 | 1 004 757 | 95 291 | 979 638 | 25 119 | 46 935 |
| 10-Jan-21 | 1 146 983 | 120 410 | 13,0% | 33,5% | 93,0% | 151 402 | 995 581 | 50 720 | 925 891 | 69 690 | 100 682 |
| 10-Jan-21 | 1 146 983 | 120 410 | 13,0% | 64,5% | 97,5% | 147 961 | 999 022 | 95 435 | 974 047 | 24 975 | 52 526 |
| 10-Jan-21 | 1 146 983 | 120 410 | 15,0% | 39,5% | 94,5% | 168 607 | 978 376 | 66 600 | 924 566 | 53 810 | 102 007 |
| 10-Jan-21 | 1 146 983 | 120 410 | 16,0% | 29,0% | 93,0% | 182 370 | 964 613 | 52 887 | 897 090 | 67 523 | 129 483 |
| 10-Jan-21 | 1 146 983 | 120 410 | 17,0% | 5,5% | 88,5% | 191 546 | 955 437 | 10 535 | 845 562 | 109 875 | 181 011 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 10-Jan-21 | 1 146 983 | 120 410 | 17,0% | 32,0% | 94,0% | 198 428 | 948 555 | 63 497 | 891 642 | 56 913 | 134 931 |
| 10-Jan-21 | 1 146 983 | 120 410 | 17,0% | 42,0% | 96,0% | 196 134 | 950 849 | 82 376 | 912 815 | 38 034 | 113 758 |
| 10-Jan-21 | 1 146 983 | 120 410 | 19,0% | 45,5% | 97,5% | 213 339 | 933 644 | 97 069 | 910 303 | 23 341 | 116 270 |
| 10-Jan-21 | 1 146 983 | 120 410 | 20,0% | 31,0% | 94,5% | 224 809 | 922 174 | 69 691 | 871 455 | 50 719 | 155 118 |
| 10-Jan-21 | 1 146 983 | 120 410 | 22,0% | 17,5% | 91,5% | 254 630 | 892 353 | 44 560 | 816 503 | 75 850 | 210 070 |
| 10-Jan-21 | 1 146 983 | 120 410 | 23,0% | 32,5% | 96,0% | 261 512 | 885 471 | 84 991 | 850 052 | 35 419 | 176 521 |
| 10-Jan-21 | 1 146 983 | 120 410 | 24,0% | 26,5% | 94,5% | 272 982 | 874 001 | 72 340 | 825 931 | 48 070 | 200 642 |
| 10-Jan-21 | 1 146 983 | 120 410 | 25,0% | 9,0% | 89,0% | 287 893 | 859 090 | 25 910 | 764 590 | 94 500 | 261 983 |
| 10-Jan-21 | 1 146 983 | 120 410 | 25,0% | 12,0% | 90,0% | 285 599 | 861 384 | 34 272 | 775 246 | 86 138 | 251 327 |
| 10-Jan-21 | 1 146 983 | 120 410 | 26,0% | 33,5% | 97,5% | 295 922 | 851 061 | 99 134 | 829 785 | 21 276 | 196 788 |
| 10-Jan-21 | 1 146 983 | 120 410 | 28,0% | 4,0% | 87,0% | 318 861 | 828 122 | 12 754 | 720 466 | 107 656 | 306 107 |
| 10-Jan-21 | 1 146 983 | 120 410 | 29,0% | 5,5% | 87,5% | 328 037 | 818 946 | 18 042 | 716 578 | 102 368 | 309 995 |
| 10-Jan-21 | 1 146 983 | 120 410 | 29,0% | 22,5% | 94,5% | 337 213 | 809 770 | 75 873 | 765 233 | 44 537 | 261 340 |
| 10-Jan-21 | 1 146 983 | 120 410 | 31,0% | 6,0% | 87,5% | 353 271 | 793 712 | 21 196 | 694 498 | 99 214 | 332 075 |
| 10-Jan-21 | 1 146 983 | 120 410 | 32,0% | 4,0% | 86,5% | 362 447 | 784 536 | 14 498 | 678 624 | 105 912 | 347 949 |
| 10-Jan-21 | 1 146 983 | 120 410 | 32,0% | 18,0% | 93,0% | 364 741 | 782 242 | 65 653 | 727 485 | 54 757 | 299 088 |
| 10-Jan-21 | 1 146 983 | 120 410 | 33,0% | 26,0% | 97,0% | 373 916 | 773 067 | 97 218 | 749 875 | 23 192 | 276 698 |
| 10-Jan-21 | 1 146 983 | 120 410 | 36,0% | 7,0% | 87,5% | 417 502 | 729 481 | 29 225 | 638 296 | 91 185 | 388 277 |
| 10-Jan-21 | 1 146 983 | 120 410 | 36,0% | 19,5% | 94,5% | 409 473 | 737 510 | 79 847 | 696 947 | 40 563 | 329 626 |
| 10-Jan-21 | 1 146 983 | 120 410 | 36,0% | 22,0% | 96,0% | 414 061 | 732 922 | 91 093 | 703 605 | 29 317 | 322 968 |
| 10-Jan-21 | 1 146 983 | 120 410 | 37,0% | 24,0% | 97,5% | 426 678 | 720 305 | 102 403 | 702 298 | 18 007 | 324 275 |
| 10-Jan-21 | 1 146 983 | 120 410 | 41,0% | 23,5% | 98,5% | 469 116 | 677 867 | 110 242 | 667 699 | 10 168 | 358 874 |
| 10-Jan-21 | 1 146 983 | 120 410 | 42,0% | 7,0% | 87,0% | 478 292 | 668 691 | 33 480 | 581 761 | 86 930 | 444 812 |
| 10-Jan-21 | 1 146 983 | 120 410 | 43,0% | 12,5% | 91,0% | 490 909 | 656 074 | 61 364 | 597 028 | 59 046 | 429 545 |

| 10-Jan-21 | 1 146 983 | 120 410 | 48,0% | 16,0% | 94,5% | 545 964 | 601 019 | 87 354 | 567 963 | 33 056 | 458 610 |
| 10-Jan-21 | 1 146 983 | 120 410 | 50,0% | 11,0% | 90,0% | 571 198 | 575 785 | 62 832 | 518 207 | 57 578 | 508 366 |
| 10-Jan-21 | 1 146 983 | 120 410 | 50,0% | 11,5% | 90,5% | 572 345 | 574 638 | 65 820 | 520 048 | 54 590 | 506 525 |

### 3.3 Example 3: The UK data

On January 11th 2021, the UK government reported 536,947 tests with 56,733 positives. The query time to obtain all CMs fitting this data was 1m37s. The results are plotted in Figure 3 and displayed in Table 4. Again, most solutions indicate significant amounts of FP results, and only for high prevalence numbers in the range 16-29% is the FP rate below 50%.



**Figure 5.** Results for United Kingdom / ONS-data for week 2 in 2021, true and false positives. Note that prevalence increases non-linearly because only the solutions fitting the observed data are plotted.
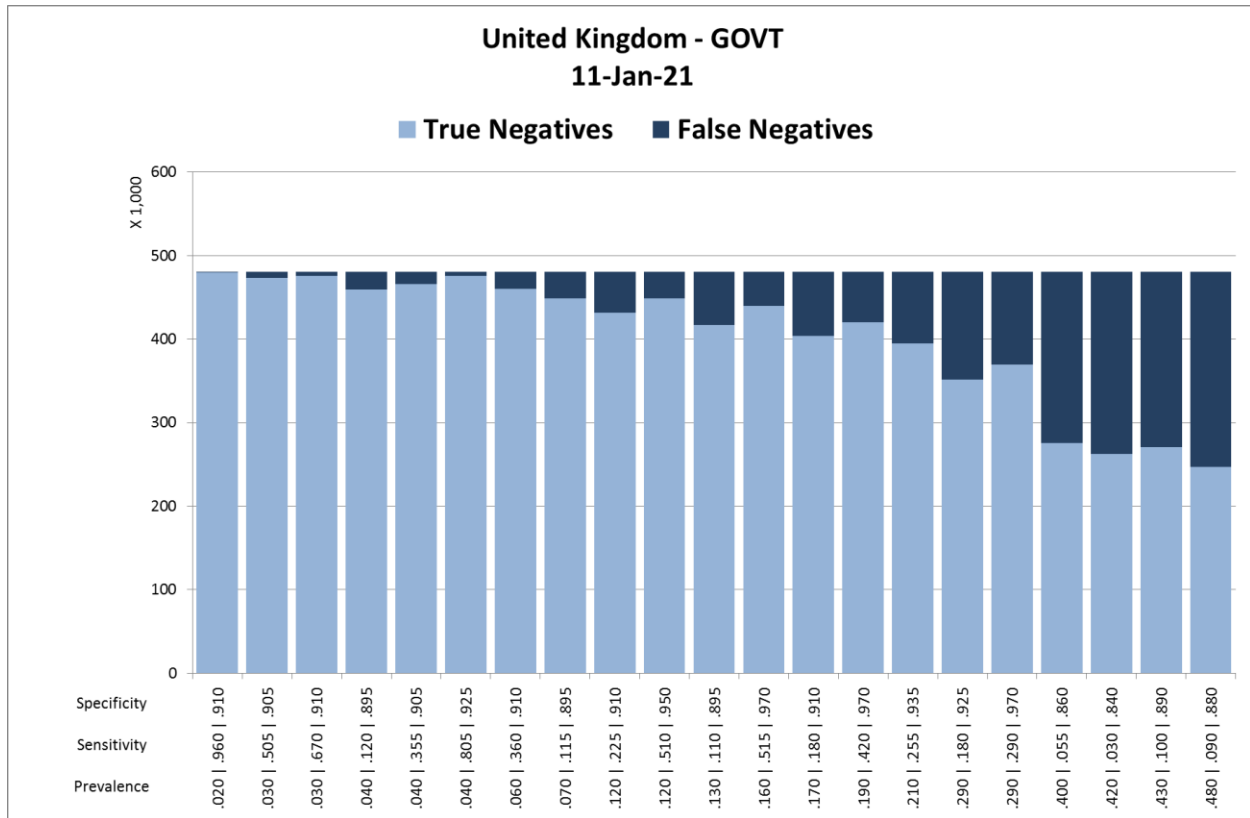
**Figure 6.** Results for United Kingdom / ONS-data for week 2 in 2021, true and false negatives. Note that prevalence increases non-linearly because only the solutions fitting the observed data are plotted.

**Table 4:** Results for UK Data for week 2 in 2021.

| report_date | tests_performed | positives_reported | prevalence | sensitivity | specificity | has_disease | hasnot_disease | true_positives | true_negatives | false_positives | false_negatives |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 11-Jan-21 | 536 947 | 56 733 | 2,0% | 96,0% | 91,0% | 9 665 | 527 282 | 9 278 | 479 827 | 47 455 | 387 |
| 11-Jan-21 | 536 947 | 56 733 | 3,0% | 50,5% | 90,5% | 13 961 | 522 986 | 7 050 | 473 303 | 49 683 | 6 911 |
| 11-Jan-21 | 536 947 | 56 733 | 3,0% | 67,0% | 91,0% | 14 498 | 522 449 | 9 713 | 475 429 | 47 020 | 4 785 |
| 11-Jan-21 | 536 947 | 56 733 | 4,0% | 12,0% | 89,5% | 23 626 | 513 321 | 2 835 | 459 423 | 53 898 | 20 791 |
| 11-Jan-21 | 536 947 | 56 733 | 4,0% | 35,5% | 90,5% | 22 015 | 514 932 | 7 815 | 466 014 | 48 918 | 14 200 |
| 11-Jan-21 | 536 947 | 56 733 | 4,0% | 80,5% | 92,5% | 22 552 | 514 395 | 18 154 | 475 816 | 38 579 | 4 398 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 11-Jan-21 | 536 947 | 56 733 | 6,0% | 36,0% | 91,0% | 31 143 | 505 804 | 11 211 | 460 282 | 45 522 | 19 932 |
| 11-Jan-21 | 536 947 | 56 733 | 7,0% | 11,5% | 89,5% | 35 439 | 501 508 | 4 075 | 448 850 | 52 658 | 31 364 |
| 11-Jan-21 | 536 947 | 56 733 | 12,0% | 22,5% | 91,0% | 62 286 | 474 661 | 14 014 | 431 942 | 42 719 | 48 272 |
| 11-Jan-21 | 536 947 | 56 733 | 12,0% | 51,0% | 95,0% | 64 971 | 471 976 | 33 135 | 448 378 | 23 598 | 31 836 |
| 11-Jan-21 | 536 947 | 56 733 | 13,0% | 11,0% | 89,5% | 70 877 | 466 070 | 7 796 | 417 133 | 48 937 | 63 081 |
| 11-Jan-21 | 536 947 | 56 733 | 16,0% | 51,5% | 97,0% | 83 764 | 453 183 | 43 138 | 439 588 | 13 595 | 40 626 |
| 11-Jan-21 | 536 947 | 56 733 | 17,0% | 18,0% | 91,0% | 93 429 | 443 518 | 16 817 | 403 602 | 39 916 | 76 612 |
| 11-Jan-21 | 536 947 | 56 733 | 19,0% | 42,0% | 97,0% | 104 168 | 432 779 | 43 750 | 419 796 | 12 983 | 60 418 |
| 11-Jan-21 | 536 947 | 56 733 | 21,0% | 25,5% | 93,5% | 114 907 | 422 040 | 29 301 | 394 608 | 27 432 | 85 606 |
| 11-Jan-21 | 536 947 | 56 733 | 29,0% | 18,0% | 92,5% | 156 789 | 380 158 | 28 222 | 351 647 | 28 511 | 128 567 |
| 11-Jan-21 | 536 947 | 56 733 | 29,0% | 29,0% | 97,0% | 156 252 | 380 695 | 45 313 | 369 275 | 11 420 | 110 939 |
| 11-Jan-21 | 536 947 | 56 733 | 40,0% | 5,5% | 86,0% | 216 927 | 320 020 | 11 931 | 275 218 | 44 802 | 204 996 |
| 11-Jan-21 | 536 947 | 56 733 | 42,0% | 3,0% | 84,0% | 224 444 | 312 503 | 6 733 | 262 503 | 50 000 | 217 711 |
| 11-Jan-21 | 536 947 | 56 733 | 43,0% | 10,0% | 89,0% | 233 035 | 303 912 | 23 303 | 270 482 | 33 430 | 209 732 |
| 11-Jan-21 | 536 947 | 56 733 | 48,0% | 9,0% | 88,0% | 256 661 | 280 286 | 23 099 | 246 652 | 33 634 | 233 562 |

## 4. Discussion

We have developed a Bayesian calculator tool allowing the estimation of possible values for the important variables prevalence, sensitivity and specificity on a daily or weekly basis (depending on the input data the user supplies). This in turn allows the calculation of real case numbers (divided into true positives and false negatives). The calculation is unbiased in that it uses all possible and sensible combinations of prevalence, sensitivity and specificity and letting Bayes' theorem decide which ones match the actually observed data. The result for a given matching combination of these three parameters is provided in the form of a CM which contains the TP, TN, FP and FN numbers. In the case where more than one combination is compatible with the given data, the user may start simulating different scenarios or use prior knowledge, e.g. about the prevalence on a given date, to further constrain the combinatorial possibilities of the output variables.

Prevalence is a crucial quantity for any inferences based on diagnostic tests, although in practise it is often not taken into account, resulting in the so-called base-rate fallacy (18). Our calculator may result in several possible prevalences compatible with the observed data; in this case, knowledge about the population having been tested may be used to constrain the possibilities. For example, in German hotspot regions in 2020, prevalences in the range 12-15% were estimated (19,20), while prevalence was zero in an asymptomatic German mother-and-child population tested in April 2020 (21). However, during the SARS-CoV-2 crisis an unprecedented mass testing not only of symptomatic, but also asymptomatic cases emerged as a strategy. Our results indicate that most of the positive test results may have been FP, if we assume that prevalence was below 5%. Such an assumption appears reasonable, as serological data from the Netherlands have estimated a prevalence of 2.7 % in early April 2020 (22) .

Our results therefore confirm the recent WHO statement "that disease prevalence alters the predictive value of test results; as disease prevalence decreases, the risk of false positive increases... This means that the probability that a person who has a positive result (SARS-CoV-2 detected) is truly infected with SARS-CoV-2 decreases as prevalence decreases, irrespective of the claimed specificity" (23). This statement may be more accurately described as the number of TPs decreasing relative to a constant FP rate so the 'risk of false positives' only increases relative to the TP numbers but the FP frequency is assumed to remain constant across a given number of tests. However, multiple modes of error may be in play. We should not assume FPs are independent of contamination from TP samples. There are higher risks of contamination in rapidly growing laboratories. Contamination of samples in the low disease prevalence seasons (summer) will go unnoticed as they do not produce a qPCR signal. Contamination prone methods may only become evident in the form of elevated and perhaps falsely assumed TPs once the disease prevalence increases in the winter.

In light of the WHO statement, the rationale for mass testing strategies implemented during periods of low prevalence (e.g. summer) appears questionable. Furthermore, mass testing increases the risk for poor sample handling and laboratory contamination which might partly explain the high FP numbers our calculator predicts. For example, Patrick *et al*. argued that besides intrinsic test performance, amplicon contamination due to high throughput processing of samples within a laboratory would be the best explanation for an increased rate of FP detections made during an outbreak of the human coronavirus HCoV-OC43 in a Canadian facility (24).

While much attention has been placed on population frequency of disease and its impact on false positives, it is critical to understand the role of false negatives and the impact these can have on track and trace systems. The nasal swabs are known to vary tremendously in RNaseP Ct values suggesting highly variable sampling or limited RNA stability in the testing reagent chain (25). Woloshin et al. demonstrate 27-40% FNs with nasopharyngeal and throat swabs

respectively and underscore the importance of understanding pretest probabilities when interpreting qPCR results(26).

With the script presented here, we can think of many variations when it comes to the size/amount of permutations, its step-size (granularity) and the 'where' clause as well as the strictness of matching TP+FP against the reported positives. For example, one could also increment prevalence on a log-scale to account for the fact that prevalence in many settings of diseases is very low (14).

We are aware that choices made in these areas have a significant impact on the number of matching CMs. An impact / sensitivity analysis was not performed, although we suspect that such analysis might reveal additional insights. However, we think that the amount of matching CMs per result that the above query produces, delivers sufficient material to make useful observations.

Future research would be very beneficial to identify a solid balance between precision (step-size in the permutations), number of matching CMs and overall query performance.


**5. Conclusions and Perspective**

We have developed a seminal, yet easy-to-use Bayesian calculator (Bayes Lines Tool, BLT) to estimate prevalence, sensitivity and specificity, and therefore TP, TN, FP and FN numbers, from official test outcome numbers. With typical reports - especially as produced for SARS-CoV-2 tests - revealing just the number of positives and number of tests performed, this paper describes an SQL implementation for this concept that generates all corresponding CMs, along with all matching permutations of prevalence, specificity and sensitivity of the tests covered by such simplified reports. Its implementation is thereby not limited to SQL but can be applied on any platform of choice.

The ability to assess posterior probability independent of the circumstances in which the diagnostic tests were performed, reveals a wide spectrum of opportunities for new applications both for the scientific community as well as for health professionals and policy makers around the globe. However, the tool may be especially relevant for the mass testing taking place within the containment strategies of worldwide governments against the SARS-CoV-2. The BLT SQL query for the first time allows one to display a real estimation of the SARS-CoV-2 situation against the background of testing volume and quality and thus will provide a valuable tool for decision makers to monitor the test strategy and the effect of interventional procedures.

This tool will not only allow official institutions to survey the test situation and obtain a better basis for planning their interventions, but also allows for individuals who got tested, to use the confusion matrices as an aid for interpreting their test results in view of the population they

were tested in. While much attention has been placed on population frequency of disease and its impact on false positives, it is critical to understand the role of false negatives and the impact these can have on track and trace systems. The nasal swabs are known to vary tremendously in RNaseP Ct values suggesting highly variable sampling or limited RNA stability in the testing reagent chain (25). Woloshin et al. demonstrate 27-40% FNs with nasopharyngeal and throat swabs respectively and underscore the importance of understanding pretest probabilities when interpreting qPCR results (26). These FN numbers are probably not due to the PCR itself, for which sensitivity is almost 100% (https://www.finddx.org/covid-19-old/sarscov2-eval-molecular/), but a matter of handling issues and the above discussed problems.

## 6.Data and software availability

The SQL-code and an example implementation in Excel can be obtained at https://bayeslines.org/.

## 7. Author contributions

<u>Wouter Aukema</u>

Roles: Formal Analysis, Investigation, Methodology, Resources, Software, Validation, Visualization Writing – Review & Editing

<u>Ulrike Kämmerer</u>

Roles: Conceptualization, Resources, Supervision, Validation, Original Draft Preparation

<u>Pieter Borger</u>

Roles: Conceptualization, Project Administration, Supervision, Validation, Writing –Review & Editing

<u>Simon Goddek</u>

Roles: Conceptualization, Project Administration, Supervision, Validation Writing – Review & Editing

<u>Bobby Rajesh Malhotra</u>

Roles: Conceptualization, Formal Analysis, Methodology, Resources, Validation, Visualization

Kevin McKernan

Roles: Investigation, Project Administration, Software, Validation, Writing – Review & Editing

Rainer J. Klement

Roles: Conceptualization, Formal Analysis, Investigation, Methodology, Resources, Validation, Writing – Original Draft Preparation

## 8. Competing Interests

## 9. Grant Information

## 10. Acknowledgements

## 11. References

1. Ren L-L, Wang Y-M, Wu Z-Q, Xiang Z-C, Guo L, Xu T, et al. Identification of a novel coronavirus causing severe pneumonia in human: a descriptive study. Chinese Medical Journal. 2020 May 5;133(9):1015–24.
2. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. New England Journal of Medicine. 2020 Feb 20;382(8):727–33.
3. Gorbalenya AE, Baker SC, Baric RS, de Groot RJ, Drosten C, Gulyaeva AA, et al. The species Severe acute respiratory syndrome-related coronavirus : classifying 2019-nCoV and naming it SARS-CoV-2. Nature Microbiology. 2020 Apr;5(4):536–44.
4. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Eurosurveillance. 2020 Jan 23;25(3):2000045.
5. Wölfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Müller MA, et al. Virological

assessment of hospitalized patients with COVID-2019. Nature. 2020 May;581(7809):465–9.

6. Hua J, Shaw R. Corona Virus (COVID-19) "Infodemic" and Emerging Issues through a Data Lens: The Case of China. Int J Environ Res Public Health. 2020 Mar 30;17(7).

7. WHO. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020 [Internet]. 2020 [cited 2021 Jan 13]. Available from: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

8. Everts J. The dashboard pandemic. Dialogues in Human Geography. 2020 Jul 1;10(2):260–4.

9. ECDC. Case definition for coronavirus disease 2019 (COVID-19), as of 3 December 2020 [Internet]. European Centre for Disease Prevention and Control. 2020 [cited 2021 Jan 13]. Available from: https://www.ecdc.europa.eu/en/covid-19/surveillance/case-definition

10. Zyl G van, Maritz J, Newman H, Preiser W. Lessons in diagnostic virology: expected and unexpected sources of error. Reviews in Medical Virology. 2019;29(4):e2052.

11. Younes N, Al-Sadeq DW, AL-Jighefee H, Younes S, Al-Jamal O, Daas HI, et al. Challenges in Laboratory Diagnosis of the Novel Coronavirus SARS-CoV-2. Viruses. 2020 Jun;12(6):582.

12. Wernike K, Keller M, Conraths FJ, Mettenleiter TC, Groschup MH, Beer M. Pitfalls in SARS-CoV-2 PCR diagnostics. Transboundary and Emerging Diseases [Internet]. [cited 2021 Jan 13];n/a(n/a). Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/tbed.13684

13. Arevalo-Rodriguez I, Buitrago-Garcia D, Simancas-Racines D, Zambrano-Achig P, Del Campo R, Ciapponi A, et al. False-negative results of initial RT-PCR assays for COVID-19: A systematic review. PLoS One. 2020;15(12):e0242958.

14. Klement RJ, Bandyopadhyay PS. The epistemology of a positive SARS-CoV-2 test. Acta Biotheor 2020. Available from: https://doi.org/10.1007/s10441-020-09393-w

15. Mascuch SJ, Fakhretaha-Aval S, Bowman JC, Ma MTH, Thomas G, Bommarius B, et al. A blueprint for academic labs to produce SARS-CoV-2 RT-qPCR test kits. medRxiv. 2020 Sep 1;2020.07.29.20163949.

16. Zhou H, Liu D, Ma L, Ma T, Xu T, Ren L, et al. A SARS-CoV-2 Reference Standard Quantified by Multiple Digital PCR Platforms for Quality Assessment of Molecular Tests. Anal Chem. 2021 Jan 19;93(2):715–21.

17. RKI. Täglicher Lagebericht des RKI zur Coronavirus-Krankheit-2019 (COVID-19) [Internet]. 2021 Jan [cited 2021 Jan 22]. Available from: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/Jan_2021/2021-01-20-de.pdf?__blob=publicationFile

18. Bar-Hillel M. The base-rate fallacy in probability judgments. Acta Psychologica. 1980 May 1;44(3):211–33.

19. Streeck H, Schulte B, Kümmerer BM, Richter E, Höller T, Fuhrmann C, et al. Infection fatality rate of SARS-CoV2 in a super-spreading event in Germany. Nature Communications. 2020 Nov 17;11(1):5829.

20. Santos-Hövener C, Neuhauser HK, Rosario AS, Busch M, Schlaud M, Hoffmann R, et al. Serology- and PCR-based cumulative incidence of SARS-CoV-2 infection in adults in a successfully contained early hotspot (CoMoLo study), Germany, May to June 2020. Eurosurveillance. 2020 Nov 26;25(47):2001752.

21. Reisinger EC, Possel R von, Warnke P, Geerdes-Fenge HF, Hemmer CJ, Pfefferle S, et al. Mütter-Screening in einem COVID-19-Niedrig-Pandemiegebiet: Bestimmung SARS-CoV-2-

spezifischer Antikörper bei 401 Rostocker Müttern mittels ELISA und Immunfluoreszenz-Bestätigungstest. Dtsch Med Wochenschr. 2020 Aug;145(17):e96–100.

22. Vos ERA, Hartog G den, Schepp RM, Kaaijk P, Vliet J van, Helm K, et al. Nationwide seroprevalence of SARS-CoV-2 and identification of risk factors in the general population of the Netherlands during the first epidemic wave. J Epidemiol Community Health [Internet]. 2020 Nov 30 [cited 2021 Jan 22]; Available from: https://jech.bmj.com/content/early/2020/11/28/jech-2020-215678

23. WHO. WHO Information Notice for IVD Users 2020/05 [Internet]. 2021 [cited 2021 Jan 22]. Available from: https://www.who.int/news/item/20-01-2021-who-information-notice-for-ivd-users-2020-05

24. Patrick DM, Petric M, Skowronski DM, Guasparini R, Booth TF, Krajden M, et al. An Outbreak of Human Coronavirus OC43 Infection and Serological Cross-reactivity with SARS Coronavirus. Can J Infect Dis Med Microbiol. 2006 Nov;17(6):330–6.

25. Dahdouh et al. https://pubmed.ncbi.nlm.nih.gov/33131699/

26. Woloshin et al. https://www.nejm.org/doi/full/10.1056/NEJMp2015897

## 12. Figures Legends

Figure 1 & 2: Example from the Dutch Corona Dashboard database for January 10th 2021. Note that prevalence increases non-linearly because only the solutions fitting the observed data are plotted. Figure 1 shows true & false positives, Figure 2 shows true & false negatives.

Figure 3 & 4: Results for German RKI data for week 2 in 2021 (17). Note that prevalence increases non-linearly because only the solutions fitting the observed data are plotted. Figure 3 shows true & false positives, Figure 4 shows true & false negatives.

Figure 5 & 6: Results for UK Data for week 2 in 2021. Note that prevalence increases non-linearly because only the solutions fitting the observed data are plotted. Figure 5 shows true & false positives, Figure 6 shows true & false negatives.