Combined panel test positivity is an uncertain predictor of disease probability in multiplex testing with implications for epidemiological estimates and clinical decision making.

Robert Challen[1,*], Anastasia Chatzilena[1], George Qian[1], Glenda Oben[1], Rachel Kwiatkowska[2,3], Catherine Hyams[1], Adam Finn[1], Leon Danon[1]

1 Bristol Vaccine Centre, University of Bristol, Bristol, UK.
2 Population Health Sciences, University of Bristol
3 NIHR Health Protection Unit in Behavioural Science and Evaluation, University of Bristol

* rob.challen@bristol.ac.uk

Abstract

When multiple tests are carried out simultaneously that identify distinct disease subtypes, it seems uncontroversial to combine the component results into a panel to detect the disease super-type in an individual, or quantify the burden of disease in the population. However, if the number of disease subtypes tested for is large then it is likely that the pre-test probability of each individual subtype is low. In this situation false positive results for each subtype test create a compound error that depends on sensitivity, specificity and prevalence of each subtype. This causes the combined test positivity to be a biased estimate of the post-test probability of disease, usually leading to significant over estimation. This complicates the interpretation of the combined panel result. In clinical applications, a higher that expected false positive rate for the super-type, could lead to inappropriate management if not correctly interpreted by the clinician. In epidemiological applications it can lead to significant uncertainty determining the burden of disease. At an epidemiological level correction of this bias is possible through Bayesian and frequentist methods, but requires a rigorous approach to propagate uncertainty, as estimates of sensitivity and specificity for multiplex tests are often not well characterised.

Introduction

Multiplex panel testing is a convenient and rapid diagnostic tool and is increasingly being used in clinical practice to differentiate between viral and bacterial causes of a range of disorders [1]. It has also been used in epidemiological studies to identify pneumococcal subtypes targetted by vaccines [2] or monitor disease spread [3]. Multiplex panel tests identify multiple subtypes of a wide range of disease, caused by different pathogens, or by different subtypes of the same pathogen [1], and may be based on immunological [**?**, **?**] or polymerase chain reaction platforms [4–9]. The number of targets tested for in each multiplex are increasing, but range from a handful, up to 48 different causative agents [3].

In the schematic in Fig 1, we distinguish between multiplex testing (subfigures A-D) and other types of multiple testing (subfigures E-G). Subfigures A-D show two component tests which identify each of two subtypes of disease. The disease subtypes

are present independent of each other and the disease super-type is present if any of the subtypes is present (B-C). In panel A we see that a false positive in one component, results in a false positive in the combined panel. In subfigure B one subtype is correctly detected, in C the other subtype, and in subfigure D the (usually) uncommon situation where both subtypes are present. In all subfigures A-D, the combined test result would be interpreted as positive. As described above, this design of test is usually extended to many more than two subtypes to make a multiplex panel.

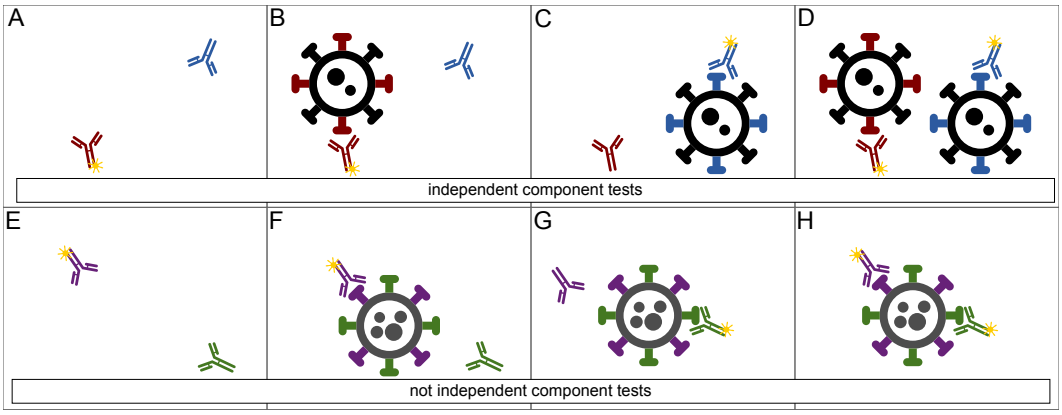

Fig 1. Two scenarios for multiple testing. This analysis is concerned with the situation where multiple tests are employed to detect multiple subtypes of disease, the results of which are combined to give an overall result, such that if any component test is positive, the combination is positive. This is a multiplex test and is depicted in panels A-D. An alternative, shown in panels E-H, and out of scope of this paper concerns the situation where multiple tests are used to identify a single condition, and in this case two interpretations of the multiple test results are possible, which either maximise test sensitivity or test specificity.

Subfigures E-H in Fig 1 show a different test design which is more related to multiple modalities of testing [10]. In this situation the multiple tests are looking for the same underlying cause of disease which does not have subtypes. In this scenario the interpretation of the 2 tests can either be that either test being positive infers disease, in which case all subfigures E-H show positive combined results, or that both tests must both be positive to identify the disease, in which care only subfigure H represents a positive result. These are not regarded as multiplex tests.

In more formal language we define a multiplex test as specifically consisting of a set of independent components which test different independent hypotheses, and the results of which are combined to give a panel result where any positive test result in a component implies a positive test result in the panel. From this point on we will only discuss multiplex panel tests.

If a condition is composed of many subtypes each individual subtype must be a fraction of the overall condition prevalence. The more subtypes in a multiplex panel, the smaller that fraction will be. If the pre-test probability for each component is low, then each component test is operating at a level where the positive predictive value of the test is also relatively low. This leads to inevitable false positives and some false negatives in each component.

The effect of this can be see in Fig 2 where a hypothetical disease subtype, present at 2%, 0.5% and 0% prevalence, is tested with a test with high specificity of 99.75% and moderate sensitivity of 80%. At 2% prevalence the false positives (cyan/blue) are more than balanced by the false negatives (orange/red), subfigure A, and the expected test positivity (magenta) is lower than the actual prevalence (black), subfigure B and C.

When the prevalence of the subtype is lower at 0.5% this pattern is reversed and the false positives outweigh the false negatives (D) leading to a higher test positivity than prevalence (E and F). In the 0% scenario in (G,H and I) all positives are by definition false positives, and distributed with high variance.

If a multiplex panel which consists of 20 subtypes, is applied to a disease which is present at a prevalence of 10%, then it is reasonable to expect that the three patterns in Fig 2 will be present in some combination. The components will have a mix of false positives and false negatives, in a manner dependant on the distribution of disease subtypes. In this particular scenario (20 highly specific tests at 10% prevalence) the balance of these will be towards false positives. Because any positive component results in a positive panel result, the component false positive errors compound in combination. In this example the error combines in such a way that the panel result will contain more false positives than false negatives, and the resulting test positivity rate will be an overestimate of true prevalence.

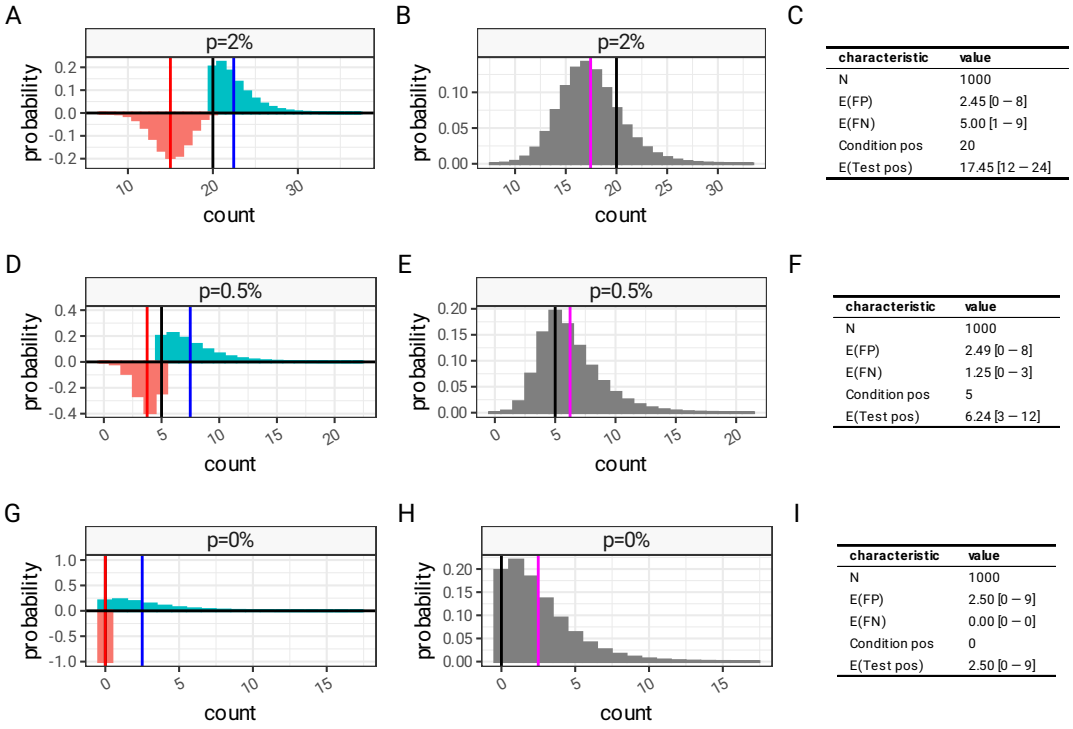

Fig 2. Error distributions of test results in low pre-test probability settings. Statistical distribution of false positives and false negatives of 1000 hypothetical test results with 0.9975 specificity and 0.8 sensitivity at different prevalence levels. Panel A, D and G show the disaggregated distribution of false positives and negatives and B, E and H the combined distribution of the error distribution of test positive observations compared to the true condition positives.

The compounding of error in numerous components has an analogy with the parallel testing of multiple statistical hypotheses, which requires careful interpretation to prevent over-interpretation of statistical significance tests, and addressed for example by Bonferroni correction [11]. This risk of over-interpretation of positive results is very similar in multiplex panel tests, as they are not adjusted for the risk of false positives from multiple components.

In the remainder of this paper we quantify this risk, and summarise the

mathematical properties of mulitplex tests. We demonstrate potential impact using a realistic simulation based on the example of pneumococcal serotypes, and study potential mitigation strategies. In S1 Appendix we provide the detail of the mathematical analysis, and validate our findings against a broad range of simulation scenarios. In S2 Appendix we provide specific detail on propagation of uncertainty associated with combined multiplex panel testing, and validate this against a set of realistic simulations. Supporting implementations of all methods described here are provided in S3 R package.

Materials and methods

Mathematical analysis and validation

Given a set of N multiplex panel component tests, the test positivity of the combined result is the combination of the components. For a specific patient k this is represented by the following expression, where I is an indicator function and O is observed test positivity.

$$I(O_{N,k}) = 1 - \prod_{n \in N} (1 - I(O_{n,k})) \tag{1}$$

The test positivity rate (or apparent prevalence: $\widehat{AP_N}$) for the panel result of N tests for a group of K patients is given by:

$$\widehat{AP_N} = \frac{1}{|K|} \sum_{k \in K} \left(1 - \prod_{n \in N} (1 - I(O_{n,k}))\right) \tag{2}$$

In S1 Appendix we determine that a true negative panel result can only be the result of a combination of true negative component results, and extension of this logic allows us to determine estimates of sensitivity and specificity expressions for combined panels. In Eq 3 and 4, $\widehat{AP_n}$ is the apparent prevalence (test positivity rate) for the component tests. $sens_n$ and $sens_N$ is the sensitivity of the components and combined panel, with $spec_N$ and $spec_n$ as the specificity.

$$\begin{aligned} spec_N &= \prod_{n \in N} spec_n \\ \widehat{sens_N} &\approx 1 - \frac{\prod_{n \in N} (1 - \widehat{AP_n}) - \prod_{n \in N} spec_n \times \frac{sens_n - \widehat{AP_n}}{spec_n + sens_n - 1}}{1 - \prod_{n \in N} \frac{sens_n - \widehat{AP_n}}{spec_n + sens_n - 1}} \end{aligned} \tag{3}$$

From this, we use the Rogan-Gladen estimator of true prevalence [12], to derive expressions for the true prevalence of a combined panel based on the test positivity, sensitivity and specificity of the components.

$$\widehat{prev_N} \approx \frac{\prod_{n \in N} spec_n - (1 - \widehat{AP_N})}{\prod_{n \in N} spec_n - \prod_{n \in N} (1 - \widehat{AP_n})} \left(1 - \prod_{n \in N} \frac{sens_n - \widehat{AP_n}}{spec_n + sens_n - 1}\right) \tag{4}$$

In S1 Appendix these estimators are demonstrated to perform well against a broad range of scenarios based on randomly generated synthetic multiplex panels, and the behaviour of these estimators is analysed in detail.

Application to realistic situations

To illustrate how this issue may affect epidemiological studies we focus on the specific example of pneumococcal serotypes, and vaccine effectiveness. We previously published the frequency of the 20 pneumococcal serotypes contained in PCV20, and identified in an invasive pneumococcal disease (IPD) cohort in Bristol between January 2021 and December 2022 [**?**].

This IPD distribution was scaled to give a realistic distribution of 20 subtypes in a hypothetical population with an overall PCV20 pneumococcal prevalence of 10%. Assuming all component tests have the same sensitivity (80%) and specificity (99.75%), and the estimators for panel sensitivity and specificity above we calculate the apparent prevalence of both components and panels.

To demonstrate the impact of component test sensitivity and specificity on panel test positivity, we use the IPD distribution and estimate the difference between apparent prevalence and true prevalence for the panel test result at 10% true prevalence assuming a broad range of component test specificities and sensitivities.

Uncertainty propagation

Our mathematical analysis assumes known values for specificity and sensitivity of component tests. However these quantities can only be estimated as a result of control group testing. Because of the low prevalence of individual subtypes is typically difficult to identify enough positive disease controls to make precise estimates of component test sensitivity. This makes interpretation of test positivity in both components and panels challenging.

There are approaches to estimating true prevalence from test positivity, in both a frequentist [13–15] and Bayesian framework [15, 16] for a single tests. In S2 Appendix we extended these frameworks to account for multiplex testing, and furthermore implemented a resampling procedure combined with the Rogan-Gladen estimator to propagate uncertainty, and tested this against a synthetic data set that is based on the IPD distribution scaled to an overall pneumococcal prevalence of 10% (further described in S2 Appendix). These methods are implemented as an R package "testerror" in S3 R package.

Results

In the IPD distribution the serotypes range from having no observed cases to making up 25.6% of the total [**?**]. When this is scaled to a synthetic population with 10% overall prevalence, the component prevalence ranges from 0% to 3.8% and as described in Fig 2. The majority of serotypes fall into the category where the apparent prevalence is higher that the true prevalence due to false positives, despite assuming a test with 99.75% specificity. However, the bias towards overestimation due to false positives is strongest for subtypes with low, or zero, prevalence, but the underestimation due to low sensitivity is strongest for subtypes with higher prevalence. This is represented in the linear relationship between true prevalence and apparent prevalence in Fig 3 (subfigure A). In S1 Appendix we identify that due to false positives in one test balancing out false positives in another test, panel test sensitivity is a complex quantity that counter-intuitively depends on disease prevalence, component distribution, sensitivity and specificity. Because of this the relationship between true panel prevalence and panel apparent prevalence (test positivity) is non linear (Fig 3 subfigure B), and in this particular scenario test positivity will be an over-estimate of true prevalence, until true prevalence exceeds 22%.

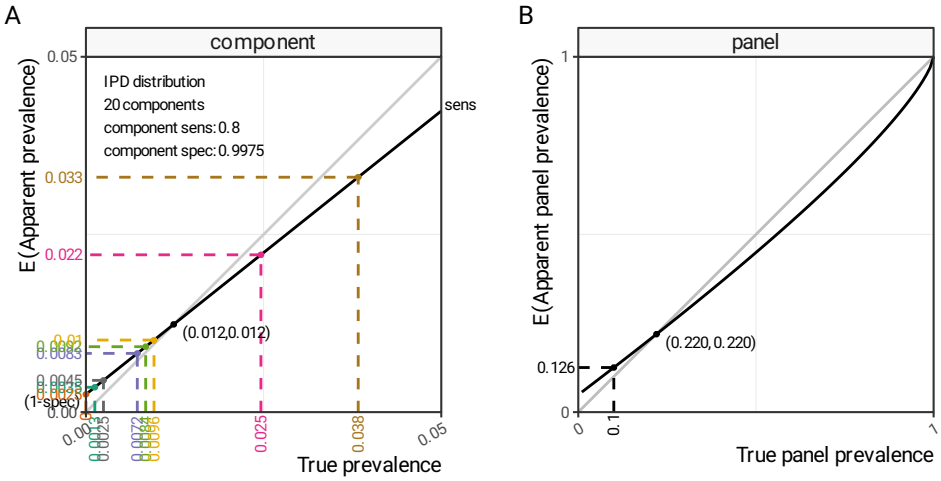

Fig 3. True versus apparent prevalence in multiplex test components and panel results. In a synthetic but realistic scenario, with excellent test specificity (99.75%) and moderate test sensitivity (80%), test positivity (apparent prevalence) is is expected to be higher that true prevalence is under a threshold of 1.2% (subfigure A). When a set of 20 components are combined, that together result in a panel prevalence of 10%, the combined error mean that the panel test positivity is higher than the true prevalence until a threshold of 22% is reached (subfigure B), and this relationship is not linear. Note that subfigure A and B are on very different scales.

The interdependence of component sensitivity and specificity in determining the difference between true and apparent prevalence is explored in Fig 4 for the scenario of 10% prevalence. The previous assumptions are marked as a blue cross in the figure, and at this high level of specificity (i.e. 99.75% - left dotted vertical line in Fig 4) the ratio between apparent and true prevalence depends on test sensitivity, and if we decrease sensitivity enough (to less than 50%) eventually the false negative rate exceeds the combined false positive rate and apparent prevalence is smaller than true prevalence. Blue represents apparent prevalence being lower than true prevalence (an underestimate) due to excess false negatives, and red higher due to excess false positives (an overestimate). The balance of error is most influenced by test specificity, with even marginally lower values of test specificity resulting in test positivity being a gross overestimate of panel prevalence. If the component test specificity is only 98% (left dotted line) the combined 2% false positive rate of 20 components is sufficient to drive the overall panel test positivity to 4 times the level of the true prevalence set in this simulation, and this is regardless of the test sensitivity.

We have described that even low false positive rates in component tests lead to overestimate of uncommon components. The converse is true for components with comparatively high prevalence. In the scenario we have been using as an example, despite the excellent specificity of the tests and 10% overall prevalence the balance of the component estimates is for test positivity to overestimate true prevalence. This is seen more clearly in Fig 5 (left subfigure) in which simulated true prevalence levels (blue) are lower than test positivity (red) for all by two of the components (serotypes 3 and 8). In the right subfigure we see the effect of combining these into vaccine groups of 7, 13, 15 and 20 components. As predicted these are larger overestimates, and the size of the overestimate depends both on the number of the components but the distribution

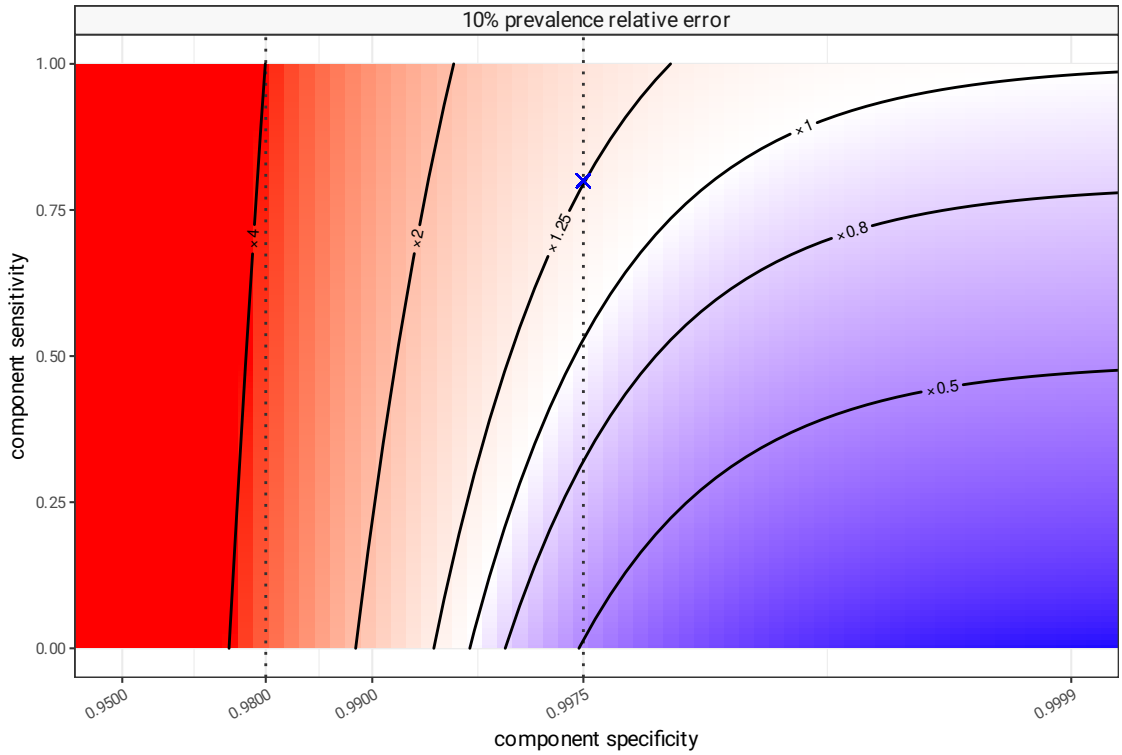

Fig 4. Bias in apparent prevalence as an estimator for true prevalence. A simulated scenario of 20 components realistically distributed following patterns seen in IPD, with a simulated true prevalence of 10%, and assuming a uniform sensitivity and a specificity for each of the component tests. Expected test positivity rates are calculated for all combinations of sensitivity and specificity, and compared to the true prevalence (10%) as a ratio. At sensitivity of 80% and specificity of 99.75% (the blue cross) the test positivity rate will be about 1.26 times that of the true prevalence. Blue areas represent parameter space where test positivity is an underestimate of true prevalence, and red where it is an overestimate.

of those components.

In S2 Appendix we describe methods for correcting this bias in both frequentist and Bayesian frameworks using the results from the mathematical analysis (S1 Appendix). In Fig 5 the Bayesian correction is applied and we are able to correctly predict the true prevalence (blue) allowing for uncertainty in our knowledge of test sensitivity and specificity. This is examined in a broader range of scenarios in S2 Appendix but in summary both Bayesian and Lang-Reiczigel (frequentist) approaches work well when we have good prior information about the test sensitivity and specificity, but if these assumptions are very wrong, then we cannot expect either method to produce accurate estimates.

Discussion

Combining multiplex test results into a panel result commonly results in test positivity that significantly overestimates true prevalence. Multiplex testing simultaneously tests many hypotheses, and by combining the result into a single panel result we compound error. This error can be significant because of the low positive predictive value of

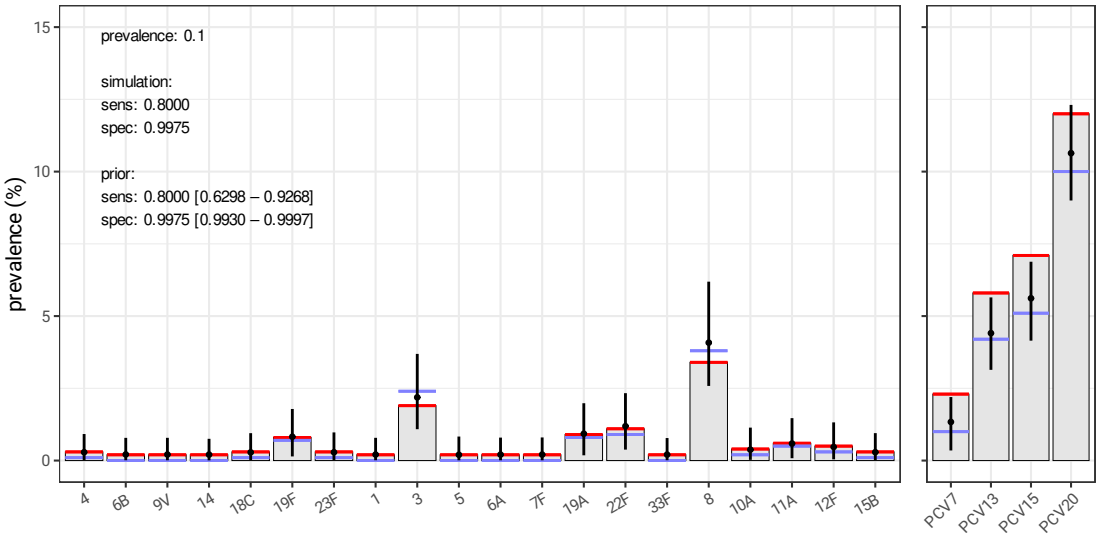

Fig 5. Correction of bias in a single IPD scenario. The relative frequency of the 20 pneumococcal serotypes contained in PCV20, and identified in Bristol within the last 2 years, were converted to a distribution of 20 subtypes to give an overall PCV20 pneumococcal prevalence of 10% (blue lines). Test positive samples were created assuming each serotype test had a sensitivity of 80% and a specificity of 99.75% (red lines) showing a mix of test positivity underestimating true prevalence (serotypes 3 and 8) and overestimating (the remainder). The simulated test result of the individual serotypes were aggregated into a PCV7 group (consisting of serotypes 4, 6B, 9V, 14, 18C, 19F, 23F), a PCV13 group (PCV7 groups plus 1, 3, 5, 6A, 7F, 19A), a PCV15 group (PCV13 plus 22F and 33F) and a PCV20 group (all serotypes). In the right subfigure combined test positivity for the groups (apparent prevalence - red lines) all overestimate true prevalence (blue line) for this scenario. Estimates of true prevalence from the test positivity, component sensitivity and specificity is possible using a Bayesian model described in S2 Appendix and these are shown as point estimates and 95% credible intervals (black).

individual component tests operating at low pre-test probability. This is critically dependent on component test specificity, and very high specificity is essential in tests which are designed to be interpreted as a combined result.

Panel test sensitivity is difficult to characterise. When multiplex tests are combined, components with a larger pre-test probability will have more false negatives. When combining false negative results they are cancelled out by true or false positives in other components. The specificity of the overall panel test is therefore a complicated function of component test sensitivity, specificity and pre-test probability (component prevalence), leading to higher panel sensitivity at higher prevalence. This is counter-intuitive as test sensitivity is usually regarded as independent of prevalence. This makes it very hard to draw conclusions from panel test positivity rates in populations with different prevalence.

It remains possible to estimate true prevalence from test positivity, despite the complexity of panel test specificity and sensitivity. Presenting test positivity rates implicitly present an apparent prevalence estimate which can be significantly biased. The expected value of test positivity is not a binomially distributed quantity and hence we cannot infer confidence intervals from an observation. Raw test positivity / apparent prevalence of panel test results is very hard to interpret. We recommend use of the

techniques described in this paper to produce modelled true prevalence estimates with confidence limits.

Sensitivity and specificity assumptions that incorporate uncertainty, are critical in producing modelled true prevalence estimates. Specificity estimates for multiplex testing usually rely on a disease free control group. This control group may also be used to determine cut points to achieve set specificity levels, and usually give us a reasonable estimate of component test specificity. Determining the sensitivity of the components of a multiplex test is much harder as it needs proven cases of disease with known subtype. These are difficult to find for rare disease subtypes, and gold standard identification of disease subtypes is not always available, or subject to its own error [17, 18]. This results in a great deal of uncertainty in estimates of component test sensitivity. In some situations panel test sensitivity is estimated directly, however as we saw above panel test sensitivity is dependant on a range of factors including overall prevalence, and component distribution. Any direct estimates of panel sensitivity are not generalisable outside of the specific population tested. The methods presented here for modelling true prevalence from multiplex tests do allow for the uncertainty in sensitivity and specificity to be propagated but as we saw in S2 Appendix, complete misspecification of either quantity prevent correct estimation of true prevalence.

The bias in panel test positivity is an inevitable consequence of combining multiple tests in environments with moderate to low prevalence. It can be prevented in a number of ways: either the specificity of the component tests must be increased, or second line confirmatory testing is performed, or the multiplex test can only be applied to populations with a very high prevalence. In the last case we may be able to use a multiplex test to determine which subtype of disease is causative if we already know the patient has the disease by using a different test, or using specific clinical diagnostic criteria that select patients with high probability of disease.

Conclusion

The Biofire FilmArray respiratory panel 2.1 is one a a number of multiplex panels directed at respiratory pathogens [1]. It detects 19 viruses [17, 19] and has been introduced in Bristol as a clinical test. A false positive for virus may result in conservative management of potentially treatable disease. Hence specificity of the viral components of the Biofire FilmArray multiplex test is important, as any false positives in the 19 viral components have to potential to influence therapy. There are multiple evaluations of the Biofire FilmArray panel [5, 17, 18, 20–22] most of which compare different multiplex tests and are conducted using unwell patients to determine positive percent agreement between different panels. There has been less focus on a large scale evaluation of test specificity using disease free controls, and hence the risk of misclassification of a patient as having any of the 19 viral diseases in the panel is not well described. With multiple comparisons in different populations it would be possible to use latent class approaches to infer component test parameters [**?**].

Uncertainty in test results due to lower sensitivity and specificity result in more noise at lower levels of prevalence [23, 24]. In test negative design vaccine effectiveness studies this acts to mask the effect of a vaccine in the lower prevalence vaccinated group. Hence test error always results in an underestimate of vaccine effectiveness [24]. The less sensitive the test is the greater this underestimate. For pneumococcal vaccination the serotype of pneumococcal disease is determined using two urine antigen detection (UAD) test panels [**?**, 2] which together can identify 24 different pneumococcal serotypes. This is designed to be highly specific with individual serotype tests being around 99.75%. Whilst not affecting conclusions on vaccine effectiveness, reuse of this panel test for determining disease burden is affected by the issues here. Despite excellent

specificity the large size of the panel creates uncertainty in prevalence estimates using UAD tests, and difficulty in comparing results to those of other similar studies.

Supporting information

S1 Appendix. Sensitivity and specificity of combined panel tests. Derivation of the performance metrics and true prevalence adjustments for combination tests.

S2 Appendix. Propagation of uncertainty of combined panel tests. Bayesian and frequentist approaches to estimating the uncertainty of panel test results.

S3 R package. testerror: Uncertainty in Multiplex Panel Testing. Provides methods to support the estimation of epidemiological parameters based on the results of multiplex panel tests, doi:10.5281/zenodo.7691196.
`https://bristol-vaccine-centre.github.io/testerror/`.

Acknowledgments

References

1. Ramanan P, Bryson AL, Binnicker MJ, Pritt BS, Patel R. Syndromic Panel-Based Testing in Clinical Microbiology. Clinical Microbiology Reviews. 2017;31(1). doi:10.1128/cmr.00024-17.

2. Bonten MJM, Huijts SM, Bolkenbaas M, Webber C, Patterson S, Gault S, et al. Polysaccharide Conjugate Vaccine against Pneumococcal Pneumonia in Adults. New England Journal of Medicine. 2015;372(12):1114–1125. doi:10.1056/NEJMoa1408544.

3. Henson SN, Elko EA, Swiderski PM, Liang Y, Engelbrektson AL, Piña A, et al. PepSeq: A Fully in Vitro Platform for Highly Multiplexed Serology Using Customizable DNA-barcoded Peptide Libraries. Nature Protocols. 2023;18(2):396–423. doi:10.1038/s41596-022-00766-8.

4. Mengelle C, Mansuy JM, Prere MF, Grouteau E, Claudet I, Kamar N, et al. Simultaneous Detection of Gastrointestinal Pathogens with a Multiplex Luminex-based Molecular Assay in Stool Samples from Diarrhoeic Patients. Clinical Microbiology and Infection. 2013;19(10):E458–E465. doi:10.1111/1469-0691.12255.

5. Murphy CN, Fowler R, Balada-Llasat JM, Carroll A, Stone H, Akerele O, et al. Multicenter Evaluation of the BioFire FilmArray Pneumonia/Pneumonia Plus Panel for Detection and Quantification of Agents of Lower Respiratory Tract Infection. Journal of Clinical Microbiology. 2020;58(7):10.1128/jcm.00128–20. doi:10.1128/jcm.00128-20.

6. Jääskeläinen AJ, Piiparinen H, Lappalainen M, Koskiniemi M, Vaheri A. Multiplex-PCR and Oligonucleotide Microarray for Detection of Eight Different Herpesviruses from Clinical Specimens. Journal of Clinical Virology. 2006;37(2):83–90. doi:10.1016/j.jcv.2006.05.010.

7. Jansen RR, Schinkel J, Koekkoek S, Pajkrt D, Beld M, de Jong MD, et al. Development and Evaluation of a Four-Tube Real Time Multiplex PCR Assay Covering Fourteen Respiratory Viruses, and Comparison to Its Corresponding Single Target Counterparts. Journal of Clinical Virology. 2011;51(3):179–185. doi:10.1016/j.jcv.2011.04.010.

8. Gröndahl B, Puppe W, Hoppe A, Kühne I, Weigl JAI, Schmitt HJ. Rapid Identification of Nine Microorganisms Causing Acute Respiratory Tract Infections by Single-Tube Multiplex Reverse Transcription-PCR: Feasibility Study. Journal of Clinical Microbiology. 1999;37(1):1–7. doi:10.1128/jcm.37.1.1-7.1999.

9. Hendolin PH, Markkanen A, Ylikoski J, Wahlfors JJ. Use of Multiplex PCR for Simultaneous Detection of Four Bacterial Species in Middle Ear Effusions. Journal of Clinical Microbiology. 1997;35(11):2854–2858. doi:10.1128/jcm.35.11.2854-2858.1997.

10. Weinstein S, Obuchowski NA, Lieber ML. Clinical Evaluation of Diagnostic Tests. American Journal of Roentgenology. 2005;184(1):14–19. doi:10.2214/ajr.184.1.01840014.

11. Shaffer JP. Multiple Hypothesis Testing. Annual Review of Psychology. 1995;46:561–585.

12. Rogan WJ, Gladen B. Estimating Prevalence from the Results of a Screening Test. American Journal of Epidemiology. 1978;107(1):71–76. doi:10.1093/oxfordjournals.aje.a112510.

13. Lang Z, Reiczigel J. Confidence Limits for Prevalence of Disease Adjusted for Estimated Sensitivity and Specificity. Preventive Veterinary Medicine. 2014;113(1):13–22. doi:10.1016/j.prevetmed.2013.09.015.

14. Thomas A, Shaheen NA, Hussein MA. An Efficient Confidence Interval Estimation for Prevalence Calculated from Misclassified Data. Biostatistics & Epidemiology. 2022;0(0):1–17. doi:10.1080/24709360.2022.2076530.

15. Flor M, Weiß M, Selhorst T, Müller-Graf C, Greiner M. Comparison of Bayesian and Frequentist Methods for Prevalence Estimation under Misclassification. BMC Public Health. 2020;20(1):1135. doi:10.1186/s12889-020-09177-4.

16. Gelman A, Carpenter B. Bayesian Analysis of Tests with Unknown Specificity and Sensitivity. Journal of the Royal Statistical Society Series C: Applied Statistics. 2020;69(5):1269–1283. doi:10.1111/rssc.12435.

17. Loeffelholz MJ, Pong DL, Pyles RB, Xiong Y, Miller AL, Bufton KK, et al. Comparison of the FilmArray Respiratory Panel and Prodesse Real-Time PCR Assays for Detection of Respiratory Pathogens. Journal of Clinical Microbiology. 2020;49(12):4083–4088. doi:10.1128/jcm.05010-11.

18. Leber AL, Everhart K, Daly JA, Hopper A, Harrington A, Schreckenberger P, et al. Multicenter Evaluation of BioFire FilmArray Respiratory Panel 2 for Detection of Viruses and Bacteria in Nasopharyngeal Swab Samples. Journal of Clinical Microbiology. 2018;56(6):10.1128/jcm.01945–17. doi:10.1128/jcm.01945-17.

19. Chang YC, Hsiao CT, Chen WL, Su YD, Hsueh PR. BioFire FilmArray Respiratory Panel RP2.1 for SARS-CoV-2 Detection: The Pitfalls. Journal of Infection. 2022;85(5):e149–e151. doi:10.1016/j.jinf.2022.07.030.

20. Popowitch EB, O'Neill SS, Miller MB. Comparison of the Biofire FilmArray RP, Genmark eSensor RVP, Luminex xTAG RVPv1, and Luminex xTAG RVP Fast Multiplex Assays for Detection of Respiratory Viruses. Journal of Clinical Microbiology. 2020;51(5):1528–1533. doi:10.1128/jcm.03368-12.

21. Babady NE. The FilmArray® Respiratory Panel: An Automated, Broadly Multiplexed Molecular Test for the Rapid and Accurate Detection of Respiratory Pathogens. Expert Review of Molecular Diagnostics. 2013;13(8):779–788. doi:10.1586/14737159.2013.848794.

22. Chan M, Koo SH, Jiang B, Lim PQ, Tan TY. Comparison of the Biofire FilmArray Respiratory Panel, Seegene AnyplexII RV16, and Argene for the Detection of Respiratory Viruses. Journal of Clinical Virology. 2018;106:13–17. doi:10.1016/j.jcv.2018.07.002.

23. Haile SR. Bias in (Sero)Prevalence Estimates; 2022.

24. Endo A, Funk S, Kucharski AJ. Bias Correction Methods for Test-Negative Designs in the Presence of Misclassification. Epidemiology and Infection. 2020;148:e216. doi:10.1017/S0950268820002058.