Prosser *et al. Population Health Metrics* (2024) 22:39
https://doi.org/10.1186/s12963-024-00360-y

Population Health Metrics

COMMENT Open Access

Number needed to isolate - a new population health metric to quantify transmission reductions from isolation interventions for infectious diseases

Aaron Prosser[1*], Bartosz Helfer[2,3] and David L. Streiner[1]

Abstract

Background We have previously developed and reported on a procedure for estimating the purported benefits of immunity mandates using a novel variant of the number needed to treat (NNT) which we called the number needed to isolate (NNI). Here we demonstrate its broader properties as a useful population health metric.

Main body The NNI is analogous to the number needed to treat (NNT = 1/ARR), except the absolute risk reduction (ARR) is the absolute transmission risk in a specific population. The NNI is the number of susceptible hosts in a population who need to be isolated to prevent one transmission event from them. The properties and utility of the NNI were modeled using simulated data and its model predictions were validated using real world data. The properties of the NNI are described for three categories of data from a previous study on transmissibility of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2): (1) in different settings, (2) after a specific exposure and (3) depending on symptomaticity status of susceptible hosts.

Conclusions We provide a demonstration of the utility of the NNI as a valuable population health metric to quantify the transmission reductions from isolation interventions.

Keywords Isolation, Number needed to isolate, Number needed to treat, Reproduction number, SARS-CoV-2, Transmission, Infectious disease

Background

Isolating infectious or potentially infectious cases is a key risk-mitigating public health intervention for infectious diseases. These interventions seek to reduce transmissions to lower the burden of a disease within a population. They range in intensity from targeted quarantining to social distancing and lockdowns. The reproduction number (R) is a useful epidemiological metric of transmissibility relevant to isolation interventions. R_0 is the basic reproduction number. It expresses the average number of secondary cases generated by an average infected individual throughout the infectious period,

*Correspondence:
Aaron Prosser
aaron.prosser@medportal.ca
[1]Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, ON, Canada
[2]Meta Research Centre, University of Wrocław, Wrocław, Poland
[3]Institute of Psychology, University of Wrocław, Wrocław, Poland

assuming no immunity in the population and no utilization of risk-mitigating interventions. Over time more people develop immunity (e.g., vaccination, infection) and risk-mitigating interventions may be used, which reduces the proportion of susceptible hosts (e.g., unvaccinated individuals). As a result, the effective reproduction number (R_t) expresses R at time *t* in a population of susceptible and non-susceptible hosts. R_t will always be less than R_0 since the latter assumes an entirely susceptible population. R is an abstraction about the transmissibility of an infectious disease and must be estimated using mathematical models. R is a helpful metric for public health because it expresses disease transmissibility and helps assess the impact of interventions (e.g., vaccination, isolation). When $R > 1$, the number of cases is increasing, such that the infectious agent can establish itself within the population. If $R < 1$, then it cannot and dies out. Accordingly, interventions can be evaluated based on the extent to which R is reduced.

While R provides valuable information about transmissibility and the impact of public health interventions, it does not indicate how many susceptible hosts need to be isolated to reduce transmissions in a population. For this purpose, the number needed to treat (NNT) may be a helpful statistical analog. The NNT is 1 divided by the absolute risk reduction (ARR) of an intervention. The NNT is the number of patients who need to receive a treatment (e.g., statins) to prevent one outcome (e.g., myocardial infarction) in a given time frame. In this comment, we discuss a population health metric analogous to the NNT, which we termed the number needed to isolate (NNI). We have previously used this metric to estimate the purported benefits of immunity mandates [1]. Here we would like to generalize it and demonstrate its broader applicability as a population health metric using a case example.

The number needed to isolate (NNI)

NNI quantifies how many susceptible hosts need to be isolated to reduce transmissions in a population. In other words, it is based on the potential transmission risk which susceptible hosts (e.g., unvaccinated individuals) pose to others in order to shed light on the benefits of isolation measures may have which specifically target this population (e.g., immunity mandates). Its difference from the NNT is that the ARR of the NNI is the absolute risk (AR) of a transmission (AR_{tr}) in the population for a given type of situation (e.g., setting, exposure, status, etc.). The rationale for the NNI is that isolation interventions isolate susceptible hosts from a situation, such that the ARR is the AR_{tr} for that category:

$$NNI = \frac{1}{ARR} = \frac{1}{AR_{tr}}$$

The $AR_{tr}$is the AR within a given timeframe of a transmission event for each category of interest (e.g., specific setting, exposure, status, etc.) from a susceptible host in the general population. The reciprocal of this probability is the NNI. The NNI is the number of susceptible hosts in a population who need to be isolated on a given day to prevent one transmission event.

The NNI pertains to the transmission risk of hosts who are susceptible to infection because they lack immunity through prior infection or vaccination. Given this, the NNI pertains to a population of individuals who, by definition, are uninfected. This means that the NNI is estimating a *potential risk* rather than an *actual risk*. That is, susceptible hosts are individuals who could *potentially* become infected and then transmit an infection rather than *actual* hosts or their contacts. The reason why the NNI focuses on potential risks is that this is what isolation measures such as immunity mandates are focused on (i.e., they isolate *potentially* risky individuals from certain settings based on their immunity or vaccine status rather than their *actual* infectious status). The NNI attempts to quantify the risk reduction gained from isolation methods that isolate uninfected, susceptible individual from various settings in order to reduce transmission risks in a region. The $AR_{tr}$is estimated by taking the combined probability of the infection risk (IR) in a population and the secondary attack rate (SAR) observed in a situation. The combined probability is needed to estimate the AR_{tr}because SARs are the transmission risks amongst infected individuals, not the population and a person must be infected first before they can transmit an agent. This combined probability essentially estimates the potential risk that a susceptible host gets infected, goes into a given type of setting, and then transmits the infection. In other words, the AR_{tr} is the absolute risk of transmission *if* the susceptible host got infected and *if* that individual also went into a given type of setting.

The $AR_{tr}$is the risk of one transmission *event*, which may include one or more secondary infections. This is because the SAR is the proportion of infections amongst the contacts of an index case, such that the total number of secondary infections depends on the total number of contacts. It also concerns one generation of transmission caused by the index case related to a specific category (such as exposure, setting or symptomaticity status).

The IR is the point-prevalence of infectious cases in the general population. It is the estimated risk that a susceptible host is infected. The $AR_{tr}$is the risk on a given day (i.e., the day of the IR) because point-prevalence data are typically measured over one day. The contact duration for many SARs is also often less than one day, especially for non-household settings. Time is also incorporated into the NNI when one calculates the NNIs over time using changing point-prevalence data. An advantage of using

point-prevalence data to quantify the infection risk is that this number contains two critical pieces of information: the number of recovered individuals and the number of susceptible individuals. This data is contained within the prevalence of infectious cases on a given day because point-prevalence is determined by how many individuals recovered and how many susceptible individuals remain in the population.

The AR_{tr} is similar to the concept of the force of infection (FOI) which is calculated from the prevalence of infection over a given time period and an estimate of transmissibility. The difference is that the FOI attempts to parameterize information about how people interact with each other (e.g., rate of contact). The NNI does not attempt to parameterize the complexities of human contacts within a given setting (e.g., the average number of contacts, how that changes over time, etc.). We have opted for a simpler model for two reasons. First, our goal is to show how an intuitive population health metric can be easily calculated from readily available data (e.g., point prevalence of infectious cases, typical SARs in particular types of settings). Second, while it is laudable to try to explicitly model human interactions, this approach to modelling the absolute risk of transmission suffers from a major problem: human interactions are deeply multi-factorial and dynamic, such that they are almost impossible to accurately parameterize. This means there will be significant inaccuracies and uncertainties in the calculation of the NNI if one were to use the FOI. In our opinion, modelling human behaviour of the sort the NNI is concerned with is a form of conjecture which can create the illusion of certainty rather than calculating risks based on actually measurable data (e.g., point-prevalence of infectious cases, SARs). As a result, the NNI uses a simpler model using calculations from actually measurable data.

Now that we have summarized the nuances and considerations regarding the calculation and interpretation of the NNI, one can see how the NNI has utility as a population health metric for infectious diseases. First, it is scalable as a function of the point-prevalence data which are used to estimate the AR_{tr}, whether it be the local environment of a care home to larger jurisdictions. Second, it is generalizable in that one can estimate the NNI for any infectious agent, insofar as there are reliable data on the SARs and IRs of the infectious disease.

To demonstrate the properties and utility of the NNI, we use severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) as a case example. This is because it is a well-characterized disease with reliable data on its SARs. Moreover, the simulations of the NNI may be helpful for public health officials fighting the pandemic as they make isolation decisions. Therefore, the purpose of this paper is to formalize the NNI and model the NNI using simulated IR data in conjunction with the observed SARs of SARS-CoV-2. The NNI was modeled using the SARs data from a real-world study [2], across three categories including households and community settings, congregation exposure and symptomaticity status. It is important to emphasize that because the NNI is scalable, one can theoretically use any type of setting, exposure or status, depending on the population and the infectious agent one is studying. For example, if one is calculating the NNI for a hospital ward using the IR of that ward, the setting might be shared rooms vs. common areas. As will be demonstrated, all that is required to estimate the NNIs for the hospital ward are reliable estimates of the SARs for specific categories and the IR.

NNI formalized

For any given transmissible agent, the NNI in a population is modeled as follows:

$$NNI = \frac{1}{ARR} = \frac{1}{AR} = \frac{1}{IR \times \ SAR}$$

where the absolute risk reduction of isolation (ARR) is estimated by multiplying the IR by the SAR from susceptible hosts infected with the agent for the specific category of interest. The SARs must be multiplied by the IR because SARs are the transmission risk posed by infected individuals, not the population. Estimating the absolute transmission risk in a population requires calculating the combined probability of infection and transmission because a host must be infected first before they can transmit an agent. When the IR is 100%, the NNI is a mathematical quantity rather than a practical possibility. This is because the AR reduces to the SAR when IR=1, but technically everyone in the population would be currently infectious and thus there would be no one to transmit a virus to.

Case study

With the NNI formalized, we can use SARS-CoV-2 as a case example to demonstrate the properties and utility of the NNI. The SARs of susceptible hosts infected with the wild-type form of SARS-CoV-2 were extracted from a retrospective cohort study conducted in Tamil Nadu, a southern state of India, in March–May 2020 across three categories including setting (households or community), exposure (congregation or no congregation) and symptomaticity status (symptomatic or asymptomatic) [2]. Wild-type SARs from that period were used because they ensure results of the model are not confounded by vaccine immunity. SARS-CoV-2 vaccines were introduced in India in January 2021 [3]. Additionally, data from early 2020 provide more reliable estimates of the SARs among susceptible hosts because these individuals likely had very limited exposure to the virus and no form

Table 1 Estimates of the secondary attack rates to model the NNI for each category

Category	SAR (95% CI)
Setting	
Households	13.36 (12.24 to 14.55)
Community	1.3 (1.09 to 1.53)
Symptomaticity	
Symptomatic	6.1 (4.33 to 8.3)
Asymptomatic	3.64 (3.33 to 3.98)
Exposure	
Congregation	9.71 (8.79 to 10.68)
No congregation	2.19 (1.92 to 2.5)

95% CI=95% confidence interval. Brackets are the 95% confidence intervals (lower limit to upper limit). SAR=secondary attack rate. NNI=number needed to isolate. Exact data calculated from Karumanagoundar et al. 2021 [2]

of immunity. Specifically, Tamil Nadu reported the first case of COVID-19 on 18 March 2020 [4]. A property of the NNI which makes it similar to R_0 is that it assumes entirely susceptible index cases and susceptible contacts because the SARs were derived from studies when there was little-to-no immunity.

To model how the NNIs change as a function of the current point-prevalence of infectious cases, values of IR were simulated ranging from 0.10 to 100% and the results were plotted. To provide a real-world IR data as a reference point for the simulated data, we report the population-weighted seroprevalence of SARS-CoV-2 infection in India for the period overlapping with the SARs data (measured from May to June 2020), based on a national serosurvey data [5]. While seroprevalence data is not ideal for estimating infection risk as point prevalence data, we used it in this case study because, firstly, point-prevalence data was lacking and, secondly, in its absence, period prevalence provides an overall impression of the level of infectious cases during the case study period. As such, period prevalence data can be used to estimate IR when point-prevalence data is lacking. IR and SARs are presented with corresponding 95% confidence intervals (CI).

Table 1 displays the wild-type SARs for each of the three categories of setting, exposure and symptomaticity, as described above.

Figure 1 plots the NNIs for those three categories and each of their two sub-categories: household and community for setting, symptomatic and asymptomatic cases for symptomaticity, and congregation and no-congregation for exposure. The 95% CI of the NNIs are calculated using the 95% CI of the SARs.

For comparison, the real-world SARS-CoV-2 IR in India for the period overlapping with the SARs data was reported as 0.73 (95% CI: 0.34–1.13) [5]. Based on Fig. 1, an IR value of 0.73 would yield very large NNIs for all three categories. Specifically, even for the sub-categories which facilitate transmission (rendered in blue on Fig. 1, i.e., household settings, symptomatic cases and congregation exposure), the NNI values would be well above 1,000. For the sub-categories rendered in red (i.e., community setting, asymptomatic index cases and no congregation exposure, for which the SARs are lower; see Table 1), the IR of 0.76 would correspond to NNI-values going into several thousand. This can be thought to represent a low reduction in transmission from isolation of susceptible hosts. One possible benchmark to make the interpretation of the NNI more straightforward comes from the NNT.

The NNTs of other interventions in medicine can be used as working benchmarks to interpret the NNIs, because they share the same scale, properties, and have a similar conceptual basis. For example, the NNTs of influenza vaccines for preventing infection range from 5 (children), 29 (older adults), and 71 (healthy adults) [6]. The NNT of the human papillomavirus (HPV) vaccine for

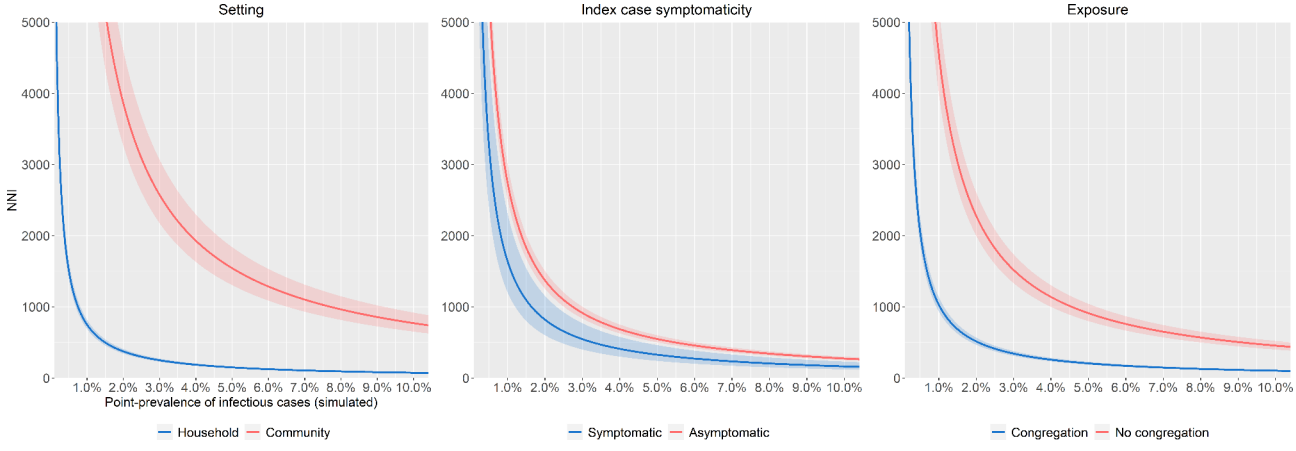

Fig. 1 Modeling the number of susceptible hosts needed to isolate (NNI) depending on specific setting, exposure and symptomaticity. The NNIs were calculated using data on SARS-CoV-2 wild-type variant in a southern region of India in early 2020 [2] and are plotted as a function of simulated infection risks (0.10–100%)

preventing any cervical pre-cancer is 60 [7]. The NNTs for antihypertensives to prevent 1 death over 5 years range from 1157 in healthy young women to 17 in high-risk older men [8].

As with the NNT, there is a nonlinear relationship between the ARR and the NNI such that at low IRs, the NNIs are high because the AR_{tr} are low. As seen in Fig. 1, isolation strategies are effective interventions for reducing SARS-CoV-2 transmissions from susceptible hosts when the IR rises above 5–10% in a given population.

A valuable population health metric

The NNI provides an intuitive population health metric which complements R to quantify the impact of isolation and to monitor isolation interventions during epidemics and pandemics. This is because, while R provides helpful information, it does not tell us how many susceptible hosts need to be isolated to reduce transmissions in a population, whereas the NNI provides this information using readily available data on measurable risks. Infection risk data at different scales (e.g., hospital wards, municipal districts, jurisdictions) can be used to calculate the NNI to quantify the transmission reductions at different scales. The NNI can generalize across different infectious diseases, insofar as there are reliable data on the SARs and IRs of the infectious agent. Given that the NNI can be calculated over time using changing point-prevalence data, it can be employed as a useful monitoring tool during epidemics.

Limitations

The main limitation of this study is that the benchmarks for interpreting the NNI have not been fully established. Various thresholds for interpreting NNTs have been proposed, acknowledging its context-dependency and multifactorial nature, but in general even for severe conditions and heavily time-dependent outcomes, it is rare that a medical treatment would be recommended if its NNTs go into hundreds [9, 10]. While the NNT provides helpful guidance, the time frames and outcomes are not the same. The NNT is primarily concerned with within-individual outcomes (e.g., myocardial infarction), whereas the NNI for transmission concerns a between-individual outcome where one or more other individuals may be impacted (i.e., transmission event). The NNI is also based on the risks over a given day (i.e., the day of the point-prevalence of infectious cases), and not, for example, months or years which are the risk periods over which many NNTs are measured. Point-prevalence is a more appropriate metric of the IR than incidence, period prevalence, or forecasted risk for four reasons. First, the risk of infection depends not just on new infectious cases, but existing ones too. Second, while risk over time is important, public health officials and communities are primarily concerned about the *current* risk of infection, not, for example, the risk over the past 3 months. Third, incidence, period prevalence, and forecasted risk depend on the time at risk. In general, shorter time windows will lower these metrics than longer time windows. If a time window is long enough, a cumulative risk can be very high even if the risk each day is low. However, there is no non-arbitrary way to set the time window to define the correct time at risk. Time at risk is not an issue for point-prevalence because it is a cross-section in time, usually over one day. Fourth, forecasting future risks is very difficult to do accurately because of the multifactorial nature and uncertainty in estimating all the relevant variables driving viral dynamics, whereas point-prevalence is an actually measurable risk. Relatedly, using period prevalence over a retrospective time window to predict what the IR will be over the next months or years is challenging since it assumes the future will correspond to the past. Therefore, point-prevalence is used to calculate the NNI. The advantage is that point-prevalence is a measurable risk (i.e., the *current* risk). The disadvantage is that it does not quantify future risks or risks over larger periods of time. Notwithstanding these limitations, the NNTs seen in other fields of medicine could provide a working benchmark for public health officials in their decision-making.

Another limitation with the NNI is that it assumes a constant level of transmission because the SAR is a fixed parameter. Since the SAR is a fixed parameter in the calculation, one way to look at the NNI is that it is calculated from an estimate of the *overall* risk of transmission in a given type of setting. Homogeneous transmissibility is obviously not true for infectious diseases such as SARS-CoV-2 where there can be "super-spreading" events which are transmission events with a high SAR. This limitation is mitigated by the fact that one can calculate CIs around the NNI based on the CIs of the SARs. This is because, by definition, CIs model the extent of variability (heterogeneity). Therefore, the NNI can model the variability in transmission, including super-spreading events, by calculating the CIs around the NNI using the CIs of the SARs, which is what we did in this case study. The NNI is thus a very flexible measure that can model both the overall transmission risk and the variability in transmission risk in order to quantity the impact of isolation measures which focus on isolating non-immune susceptible hosts to reduce transmission such as immunity mandates.

Strengths and future research

These limitations are balanced by strengths of our study, including the use of high-quality real-world data fulfilling the key theoretical assumptions of model. The NNI itself

is underpinned by a fundamental simplicity and transparency of our model, inspired by a widely-used health metric, which is easy to calculate, reproduce and understand.

Future research should focus on applying the NNI to real-world contexts of epidemics and pandemics using a range of different pathogens. It would also be helpful to have databases of the NNI to assist public health officials in their decision-making. Furthermore, it would be helpful to explore the statistical associations within populations between the NNI and R, as well as other population health metrics for infectious agents (e.g., hospital/ICU occupancy, deaths, levels of immunity). Relatedly, time trends and predictors of the NNI can be explored to link this metric to other constructs. We anticipate that as the NNI is studied in different infectious diseases, we will learn more about the thresholds for interpreting the magnitude of this effect size, much like medicine has learned about the NNT as more diseases and treatments have been studied. Finally, future work can expand the model to include parameters which take into account the complexities of human interactions within a setting.

Conclusions

This paper comments on a broad applicability of a new population health metric called the number needed to isolate (NNI), which we have previously developed in a narrower context of estimating the purported benefits of immunity mandates [1]. NNI expresses the number of susceptible hosts in a population who need to be isolated from a category to prevent one transmission event from susceptible hosts in that category (e.g. across different settings, after a specific exposure or depending on symptomaticity status). Our simulations revealed that the NNI has similar properties as the NNT and that the NNT could be used as a working benchmark for interpreting the NNI. By using the Tamil Nadu SARS-CoV-2 data as a case example, we have provided a useful demonstration of the properties and validity of the NNI, which provides public health officials with a flexible statistical tool because it is scalable and generalizable and which allows to efficiently monitor isolation interventions in real-time as epidemics and pandemics progress.

Abbreviations

ARR	Absolute Risk Reduction
ATR	Absolute Transmission Risk
CI	Confidence Interval
IR	Infection Risk
NNI	Number Needed to Isolate
NNT	Number Needed to Treat
R	Reproduction Number
R_0	Basic Reproduction Number
R_t	Effective Reproduction Number
SAR	Secondary Attack Rate
SARS	Cov 2 Severe Acute Respiratory Syndrome Coronavirus 2

Acknowledgements

We would like to thank Dr. Nathan Bakker and Dr. Mahesh Shenai for their valuable feedback on the ideas in this manuscript.

Author contributions

AP, BH, and DS contributed equally to this work. AP, BH, and DS conceived the idea. AP and BH performed the formal analysis. AP performed the data visualization. AP and BH, wrote the first draft of the manuscript. All authors gave critical feedback on the revised report and approved the final version of the manuscript. The corresponding author attests all listed authors meet authorship criteria and that no others meeting criteria have been omitted.

Funding

BH is supported by the University of Wroclaw within the "Excellence Initiative – Research University" framework (grant number 0320/2020/20). The funder was not involved in the design of the study and collection, analysis and interpretation of data and in writing of the manuscript. AP and DLS declare no additional funding.

Data availability

The datasets supporting this analysis are available in the GitHub repository at https://github.com/TheNNIforViralTransmission/SARS-CoV-2. All methods were carried out in accordance with relevant guidelines and regulations.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 1 February 2023 / Accepted: 12 December 2024
Published online: 23 December 2024

References

1. Prosser A, Helfer B, Streiner D. Estimating the risk reduction of isolation on COVID-19 non-household transmission and severe/critical illness in non-immune individuals: September to November 2021. J Eval Clin Pract. 2023;in press.
2. Karumanagoundar K, Raju M, Ponnaiah M, Kaur P, Rubeshkumar P, Sakthivel M, et al. Secondary attack rate of COVID-19 among contacts and risk factors, Tamil Nadu, March–May 2020: a retrospective cohort study. BMJ open. 2021;11(11):e051491.
3. Chakraborty C, Sharma AR, Bhattacharya M, Agoramoorthy G, Lee S-S. Asian-Origin Approved COVID-19 Vaccines and Current Status of COVID-19 Vaccination Program in Asia: A Critical Analysis. Vaccines. 2021;9(6):600.
4. State Control Room DoPHaPMHaFWD, Government of Tamil Nadu. Available from: https://stopcorona.tn.gov.in/wp-content/uploads/files/MediaBulletin180320COVID19.pdf.
5. Murhekar MV, Bhatnagar T, Selvaraju S, Rade K, Saravanakumar V, Vivian Thangaraj JW, et al. Prevalence of SARS-CoV-2 infection in India: Findings from the national serosurvey, May-June 2020. Indian J Med Res. 2020;152(1–2):48–60.
6. Liang S. Vaccines for Preventing Influenza in Healthy Individuals 2018 [https://www.thennt.com/nnt/vaccines-preventing-influenza-healthy-individuals/
7. Li JJ, Stetz J. HPV vaccines for prevention of cervical pre-cancer in adolescent girls and women 2018 [Available from: https://www.thennt.com/nnt/hpv-vaccines-prevention-cervical-pre-cancer-adolescent-girls-women/
8. Gueyffier F, Boutitie F, Boissel J-P, Pocock S, Coope J, Cutler J, et al. Effect of antihypertensive drug treatment on cardiovascular outcomes in women and men: a meta-analysis of individual patient data from randomized, controlled trials. Ann Intern Med. 1997;126(10):761–7.
9. Sinclair JC, Cook RJ, Guyatt GH, Pauker SG, Cook DJ. When should an effective treatment be used? Derivation of the threshold number needed to treat and the minimum event rate for treatment. J Clin Epidemiol. 2001;54(3):253–62.

10. Nguyen C, Naunton M, Thomas J, Todd L, McEwen J, Bushell M. Availability and use of number needed to treat (NNT) based decision aids for pharmaceutical interventions. Exploratory Res Clin Social Pharm. 2021;2:100039.

Publisher's note