Assessment of the efficacy of SARS-CoV-2 vaccines in non-human primate studies - a systematic review

Michel Jacques **Counotte**^{1*}, Mariana **Avelino de Souza Santos**¹, Koert J **Stittelaar**¹, Wim H M **van der Poel**¹, Jose L **Gonzales**¹

¹Wageningen Bioveterinary Research, Wageningen University and Research, Lelystad, the Netherlands

*Corresponding author: michel.counotte@wur.nl

8 Abstract

9 The outbreak of Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) triggered the rapid and
10 successful development of vaccines to help mitigate the effect of the coronavirus disease 2019 (COVID-
11 19) and circulation of the virus. Preclinical vaccine trials provide a wealth of information about the
12 presence and persistence of virus in different anatomical sites. We systematically reviewed all available
13 preclinical SARS-CoV-2 candidate vaccine studies where non-human primates were challenged after
14 vaccination. We found marked heterogeneity in experimental design between the studies. Most of the
15 tested vaccines, only triggered a low or moderate reduction of viral RNA in the upper respiratory tract;
16 We need to consider that most SARS-CoV-2 vaccines that protect against disease might not fully protect
17 against infectiousness and vaccinated individuals might still contribute to SARS-CoV-2 transmission.
18 Careful assessment of secondary attack rates from vaccinated individuals is warranted. Standardization
19 in design and reporting of preclinical trials is necessary.

Introduction

The pandemic of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) that causes coronavirus disease 2019 (COVID-19) prompted an outbreak response without precedent. To mitigate the devastating effect of the disease, vaccine development has been rapid and successful. In under a year after the World Health Organization declared the outbreak a public health emergency of international concern¹, the first SARS-CoV-2 vaccine (Pfizer-BioNTech) had received emergency use authorization by the United States Food and Drug Administration². In China and Russia emergency access to the first vaccines was already granted mid-2020; the 'CoronaVac' vaccine received approval for emergency use in July 2020³, and the Russian Sputnik V vaccine received emergency authorization in early August 2020⁴. As of March 26, 2021, there were 92 vaccine candidates in trials, of which 13 had received approval in one or more countries⁵.

An important metric to quantify the performance of a vaccine is the 'vaccine efficacy', which, in a clinical setting, is often defined as capacity of a vaccine to prevent (severe) disease⁶. However, vaccine efficacy is not only reflected by the prevention of symptomatic disease after infection, but also the effect on the susceptibility and infectiousness^{7,8}, e.g., the ability of a virus to establish infection in a host and the ability of a host to infect others. As Lipsitch and Dean (2020) comprehensively explain, there is a necessity to look beyond the effect of a vaccine on an individual, and also examine the efficacy in the context of preventing transmission⁹. Where frontline health workers are among the first to be vaccinated, Lipsitch and Dean note that "A worst-case scenario is a vaccine that reduces disease while permitting viral shedding; this could fail to reduce transmission or conceivably even increase transmission if it suppressed symptoms.". Even though vaccination might reduce the severity of disease, the vaccination effect on the risk of transmission also needs to be considered. Individuals, despite having an asymptomatic SARS-CoV-2 infection, can still transmit virus¹⁰. Similarly, this could mean that infected vaccinated people, despite not showing any symptoms or even being aware of their infection status, pose a risk to others. Ongoing transmission under this scenario could lead to emergence of new virus variants.

The typical development cycle of any vaccine includes an assessment of safety, immunogenicity and efficacy in preclinical animal models¹¹. Successful candidates are then trialled in humans during different phases. Preclinical studies already provide a wealth of information regarding the vaccine efficacy and transmission dynamics. Several animal models have been established in both small animals like mice, hamsters, and ferrets, and larger animals such as non-human primates (NHP)¹².

Early in the pandemic, NHP models were established. Experimental infection of *Cynomolgus* macaques (*Macaca fascicularis*), with SARS-CoV-2, caused COVID-19-like disease and viral shedding in the nose and throat¹³. Similarly, a rhesus macaque (*Macaca mulatta*) model was established, that mimicked human infection characteristics, however, with less marked clinical signs¹⁴. Both models allowed assessment of viral shedding kinetics in the upper and lower respiratory tract. Especially for vaccine development, challenge studies in these animal models provide insight in the efficacy of the vaccines to prevent virus replication and illness, but most importantly on their effect on infectiousness by measuring viral shedding over time; a measurement that is much harder to obtain in human field trials.

Here, we present a systematic assessment and comparison of the preclinical SARS-CoV-2 vaccine evaluation studies in NHPs to provide insight in the vaccine efficacy of the candidate vaccines.

Results

Description of the included studies

We performed a systematic review of the literature indexed in PubMed, EMBASE, BioRxiv and MedRxiv until January 20, 2021. We identified and screened 742 studies of which we included 147 based on title and abstract (Supplementary File: PRISMA: flowchart & checklist). After the screening of the full text, we retained 21 studies¹⁵⁻³⁵.

In the retained studies, that describe preclinical challenge experiment in non-human primates, a total of 16 studies used rhesus macaques^{15,17-20,22-25,28,30-35} five used cynomolgus macaques^{16,21,26,27,29} The age of the animals ranged between 2-22 years; the majority of animals were young adults (Supplementary Figure 1). On average two vaccinations (range: 1-7) per study were tested. These were either different formulations or different doses (Supplementary Table 1). The animals were challenged after vaccination at a median duration of 28 days (range: 8-91 days) after administration of the last vaccine. Challenges were administered at different doses, in different volumes, via different routes: In the majority of studies, the animals (13 out of 21) were challenged both intranasally and intratracheally. In five studies, animals were challenged only either intranasally or intratracheally. In Table 1 we provide an overview of the experimental methods of the included studies. Additional properties of the vaccines are provided in Supplementary Table 2.

Author [ref.]	Manufacturer	Type of vaccine	Species (macaque)	Age (year)	Challenge dose	Timing of challenge*	Challenge route	Challenge strain
Bewley ¹⁵	unknown	Inactivated	rhesus	2-4	5x10 ⁶ PFU	d14	IT, IN	SARS-CoV-2 Victoria/01/2020
Brouwer ¹⁶	unknown	Subunit (nanoparticle-based)	cynomolgus	5-6	10 ⁶ PFU	d14	IT, IN	hCoV-19/France/IDF0372/2020
Corbett ¹⁷	Moderna	mRNA	rhesus	3-6	7.6x10 ⁵ PFU	d28	IT, IN	SARS-CoV-2 USA-WA1/2020
Feng ¹⁸	CanSino	Vector: Adeno	rhesus	6-14	2x10 ⁴ or 400 TCID ₅₀	d30, d56	IT	2019-nCoV-WIV04
Gabitzsch ¹⁹	ImmunityBio	Vector: Adeno	rhesus	>2.5	10 ⁶ TCID ₅₀	d28	IT, IN	not reported
Gao ²⁰	Sinovac	Inactivated	rhesus	3-4	10 ⁶ TCID ₅₀	d8	IT	SARS-CoV-2 CN1
Guebre-Xabier ²¹	Novavax	Subunit	cynomolgus	>3	1.1x10 ⁴ PFU	d14	IT, IN	2019-nCoV/USA-WA1/2020107 (Guangdong Provincial CDC, Guangdong, China)
Liang ²²	Clover Biopharmaceuticals	Subunit	rhesus	3-6	2.6x10 ⁶ TCID ₅₀	d14	IT, IN	USA-WA1/2020
Mercado ²³	Janssen Vaccines & Prevention	Vector: Adeno	rhesus	6-12	1.0x10 ⁵ TCID ₅₀	d42	IT, IN	USA-WA1/2020
Patel ²⁴	Inovio	DNA	rhesus	unknown	1.1x10 ⁴ PFU (1.2x10 ⁸ vp)	d91	IT, IN	USA-WA1/2020
Rauch ²⁵	CureVac	mRNA	rhesus	3-6	5x10 ⁶ PFU	d28	IT, IN	VERO/hSLAM cell passage 3 (Victoria/1/2020)
Sanchez-Felipe ²⁶	KU Leuven	Vector: Yellow-fever	cynomolgus	unknown	1.5x10 ⁴ TCID ₅₀	d14	IT, IN	not reported

Author [ref.]	Manufacturer	Type of vaccine	Species (macaque)	Age (year)	Challenge dose	Timing of challenge*	Challenge route	Challenge strain
Seo ²⁷	Genexine/Vaxigen	DNA	cynomolgus	5-6	2.6x10 ⁷ TCID50	d35	IT, IN, PO, IV, conjunctival	SARS-CoV-2 virus, obtained from the National Culture Collection for Pathogens (accession number 43326)
Solforosi ²⁸	Janssen Vaccines & Prevention	Vector: Adeno	rhesus	13-22	1x10 ⁵ TCID50	d35, d91	IT, IN	G614 SARS-CoV-2
Sun ²⁹	unknown	Subunit	cynomolgus	>3	10 ⁶ TCID50/mL (3.2ml)	d28	IT, IN, ocular	BetaCoV(BJ01)
van Doremalen ³⁰	Astrazeneca/Oxford	Vector: Adeno	rhesus	2-4	4x10 ⁵ TCID50/ml (6.5ml total)	d28	IT, IN, PO, ocular	nCoV-WA1-2020
van Doremalen ³¹	Astrazeneca/Oxford	Vector: Adeno	rhesus	4-11	10 ⁶ TCID50	d28	IT, IN	SARS-CoV-2/human/USA/RML-7/2020
Vogel ³²	Pfizer/BioNTech	mRNA	rhesus	2-4	1.05x10 ⁶ PFU	d55	IT, IN	USA-WA1/2020
Wang ³³	Sinopharm	Inactivated	rhesus	2-4	10 ⁶ TCID50	d10	IT	not reported
Yang ³⁴	unknown	Subunit	rhesus	5-9	0.5 ml, 10 ⁶ PFU/ml (5x10 ⁵ PFU)	d21	IN	not reported
Yu ³⁵	Harvard/Janssen Vaccines & Prevention	DNA	rhesus	6-12	1.2x10 ⁸ vp (1.1x10 ⁴ PFU)	d21	IT, IN	not reported

Table 1. Methodological properties of the included studies. The manufacturer, name, type, dosage, timing, and route of the vaccines used. The timing, route and SARS-CoV-2 strain of the challenge used. *Timing of challenge is defined as the time between the last vaccination and challenge. Abbreviations: IM: intramuscular, IT: intratracheal, IN: intranasal, SC: subcutaneous, ID: intradermal, PO: *per os*/ oral, IV: intravenous, PFU: plaque-forming units, TCID50: 50% tissue culture infective dose, VP: viral particles, wk: week, d: day.

Reduction in excretion of viral RNA

For the quantification of viral RNA in the different anatomical sites after challenge, the data of individual animals was provided for seven studies^{17-19,23,26,30,33}. For the remaining 14 studies we extracted the aggregated results from the text or the figures^{15,16,20-22,24,25,27-29,31,32,34,35}. All studies reported measurements of the presence of viral RNA (genomic or subgenomic) at multiple timepoints. Animals were followed-up after challenge for a median period of seven days (range: 4-14 days). In most studies (17 out of 21) the presence of viral RNA in the upper respiratory airways was detected either by using nasal samples^{15-19,21-25,27-32,35} or by tracheal samples (10 out of 21)^{15,16,20,22,26-29,33,34}; To detect viral RNA in the lower airways, bronchoalveolar lavage (BAL) was used in 10 out of 21 studies^{17,19,21,23,24,28,30-32,35}. Seven studies assessed the viral excretion using anal swabs^{16,20,22,29,32-34}.

In the upper respiratory tract, as assessed with nasal samples, the best performing vaccine candidate per study caused only a low (6 out of 12 studies), or moderate (5 out of 12) reduction of viral genomic RNA, and low (5 out of 11 studies) or moderate (3 out of 11 studies) reduction of subgenomic RNA. Figure 1 provides an overview of the relative reduction of excretion of viral RNA in the different samples for the best performing vaccine candidates by study. A full overview of all tested vaccines is provided in Supplementary Figure 2. We observed a negative trend between the challenge dose and the reduction of viral genomic RNA (Figure 2). We did not see a trend for the age of the animals used and the time between last vaccination and challenge (Supplementary Figure 3 & 4).

Figure 1. Reduction in viral genomic and subgenomic RNA after vaccination in different anatomical sites, expressed as relative difference in area under the curve (AUC). The best performing vaccines are shown by study; control groups serve as reference group for the comparison and are thus omitted from the figure. Labels show the first author of the study and the vaccine

manufacturer in brackets. For visualization purposes we grouped the difference in AUC using 'Jenks natural breaks' into three groups: low (<18%), moderate (18-44%) and high (>44%) reduction of viral RNA excretion. Abbreviations: gRNA, genomic ribonucleic acid (RNA); sgRNA, subgenomic RNA; BAL, Bronchioalveolar lavage; NS, nasal or nasopharyngeal swab; TS, tracheal swab or tracheal brush.

Figure 2. The relationship between challenge dose and reduction in area under the curve (AUC) between the best-performing vaccine per study in different anatomical sites. Abbreviations: gRNA, genomic ribonucleic acid (RNA); sgRNA, subgenomic RNA; BAL, Bronchioalveolar lavage; NS, nasal or nasopharyngeal swab; TS, tracheal swab or tracheal brush.

Reduction in clinical and pathological outcomes

Reports on clinical and pathological outcomes were used as indicators of the potential efficacy of the vaccine candidates to reduce or prevent disease. Out of the 21 included studies, 14 reported histopathological findings^{15,17,18,20,21,25-30,32-34} and eight reported on the clinical status of the animals^{15,19,22,23,28,30,32,33}. The reporting was often qualitative (Figure 2), and clinical symptoms were seldom reported, which may indicate the difficulty of monitoring clinical signs in these animal models. Of the 14 studies reporting pathological findings, 12 indicated a reduction in severity (histopathology lesions) in the vaccinated group. Whilst clinical status was reported in eight studies only two reported a reduction in clinical manifestation. All studies that reported a reduction in viral load in the lung also reported (qualitatively or quantitatively) a reduction in histopathological lesions. The data did not allow a quantitative comparison of the results.

Histopathology	Immuno-histopath.	Viral load lung	Clinical outcomes	
reduced in vaccine group	NA	reduced in vaccine group	null, both groups	Bewley (unknown)
NA	NA	NA	NA	Brouwer (unknown)
reduced in vaccine group	reduced in vaccine group	NA	NA	Corbett (Moderna)
reduced in vaccine group	NA	reduced in vaccine group	NA	Feng (CanSino)
NA	NA	NA	null, both groups	Gabitzsch (ImmunityBio)
reduced in vaccine group	NA	reduced in vaccine group	NA	Gao (Sinovac)
reduced in vaccine group	NA	NA	NA	Guebre-Xabier (Novavax)
NA	NA	reduced in vaccine group	reduced in vaccine group	Liang (Clover Biopharmaceuticals)
NA	NA	NA	null, both groups	Mercado (Janssen Vaccines & Prevention)
NA	NA	NA	NA	Patel (Inovio)
reduced in vaccine group	reduced in vaccine group	reduced in vaccine group	NA	Rauch (CureVac)
null, both groups	NA	NA	NA	Sanchez-Felipe (KU Leuven)
reduced in vaccine group	NA	NA	NA	Seo (Genexine/VaxiGen)
reduced in vaccine group	reduced in vaccine group	reduced in vaccine group	null, both groups	Soforosi (Janssen Vaccines & Prevention)
reduced in vaccine group	NA	reduced in vaccine group	NA	Sun (unknown)
reduced in vaccine group	NA	reduced in vaccine group	reduced in vaccine group	van Doremalen (Astrazeneca/Oxford)
NA	NA	reduced in vaccine group	NA	van Doremalen2 (Astrazeneca/Oxford)
inconclusive	NA	NA	null, both groups	Vogel (Pfizer/BioNTech)
reduced in vaccine group	NA	reduced in vaccine group	null, both groups	Wang (Sinopharm)
reduced in vaccine group	NA	reduced in vaccine group	NA	Yang (unknown)
NA	NA	NA	NA	Yu (Harvard/Janssen Vaccines & Prevention)

Histopathology & clinical findings

- Positive, quantitative
- Positive, qualitative
- Null finding
- Not reported

Figure 3. Summary of pathological and clinical results from the included preclinical vaccine studies. Abbreviations: NA, not available.

Discussion

Here, we provided a systematic and standardized summary of the results of preclinical SARS-CoV-2 vaccine trials in NHPs, where animals were challenged after vaccination. There was a strong heterogeneity in the methodology and reporting between studies, hampering comparability of the candidate vaccines. However, we were able to summarize the efficacy of the vaccines to reduce viral RNA shedding after challenge by standardizing and stratifying the results by different anatomical sites and diagnostic method. We found that vaccination with many of the candidate vaccines only seem to result in a low or moderate reduction of the shedding of viral RNA in the upper respiratory tract as detected in nasal swabs. The effect on the clinical parameters and histological findings were often only reported qualitatively.

A strength of this review is that we standardized the results of preclinical NHP vaccine models where possible, and visualized the remaining heterogeneity. Compared to other published reviews on the same topic, we are the first to provide a systematic quantitative analysis of the data that facilitates the interpretation and future trial design, where in other reviews, the results of preclinical vaccine models have only been summarized narratively¹¹.

Our review also has limitations. First, we chose to assess the vaccine efficacy only in NHP models. Despite being closest to humans on many aspects, the NHP models do not mimic the COVID-19 clinical disease in humans³². However, the profile of viral shedding, which is the most relevant proxy measure of infectiousness, seems similar^{12,13,36}. Second, the way we calculated the area under the curve (AUC) was hampered by the reporting quality of the studies. The absence of viral RNA cannot be measured other than that the value falls below a certain limit of detection. This threshold varies between techniques and the way the data is visually presented differs between studies. For example, studies can plot points at the limit of detection, or some arbitrary value to 'aid' visualization: Yang et al impute artificial zero's³⁴ and Vogel et al. chose to plot values below the limit of detection at halve the value of that limit³². No access to the full data, hence relying on digitizing images, creates an artificial 'baseline' AUC at the limit of detection, which affects the theoretical reduction in AUC between vaccine and control groups. We did account for this by providing a semi-quantitative description of the reduction in shedding of RNA. Last, the results might have been biased towards positive findings; Out of the 13 vaccines that are on the market as of March 26, 2021⁵, we were only able to identify preclinical trials for eight in non-human primates. The absence of a complete overview could be an effect of a publication bias.

Our study implies that the lack of efficacy of many of the candidate vaccines to prevent infectiousness calls for caution. We interpret the limited reduction of viral RNA in the upper respiratory tract of some of the vaccine candidates as reason to consider individuals vaccinated with these vaccines to be potentially infectious after natural infection despite vaccination. However, interpreting the presence and reduction of viral RNA in the different samples as a measure of reduction of infectiousness depends on several assumptions that need to be validated or at least be made explicit. The evidence that comes from preclinical studies suffers from 'indirectness'³⁷. We do not measure the presence of infectious particles in vaccinated humans after natural infection, but measure viral RNA in artificially challenged NHPs. Despite that it is likely that viral RNA correlated with infectious virus, it remains unclear whether these RNA titres actually correspond with *de novo* produced infectious particles, or that a fraction of viral the RNA represents still input challenge virus³⁵. Some authors argued that subgenomic RNA might be more indicative of replication competent virus^{31,38}, however, due to its stability, this might be an overestimate³⁹. We need to consider that the level of shedding might be age-dependent as well, since higher levels of SARS-CoV-2 RNA were detected in nasal swabs of older animals compared with younger animals^{13,40}, thus the shedding might be more prominent in older individuals.

The results of the NHP challenge trials only allowed for a limited quantitative interpretation of the results in the context of the efficacy to prevent disease. The viral loads in lungs could be considered an indirect indicator of severity (lower viral loads may lead to milder lung pathology). In most studies, the vaccines did reduce either (histo)pathological changes or viral load in the lungs; the clinical changes, however,

were often poorly assessed. Although NHP models might not always show severe COVID-19 disease^{13,14}, Munster et al. presented a clinical score for NHP SARS-CoV-2 infection experiments¹⁴, which showed that assessing clinical changes in this animal model is possible. From the assessed studies, only Doremalen et al.³⁰ quantified, using Munster et al. clinical score, the clinical effect of the challenge after vaccination. Liang et al., despite referring to Munster et al., only present the body weight as measure of clinical effectiveness²². A meticulous assessment of clinical outcomes, as done for example by Doremalen et al., does provide additional information for assessing the potential efficacy of vaccine candidates.

Our study also has several implications for the design and reporting of preclinical studies: First, it highlights the need for standardization. We found differences in almost all aspects of the preclinical trial design: In the timing, the dose, the strain and the administration route of the challenge; the diagnostic methods used, the selected timepoints at which the animals were sampled, and the assays used. These parameters influence the outcomes: For example, we showed here the relationship between the relative reduction in viral RNA shedding (efficacy) and the challenge dose and in the literature a more prominent viral replication in aged macaques than young macaques is described⁴⁰. Besides the properties of the experimental designs that we assessed, plenty other factors might influence the results: including serological- and/or microbiological status of the animals, the specifications of the challenge virus (RNA sequence heterogeneity and matching with the vaccine), sampling strategy, etc.. This all highlights a clear need for standardization and registration of protocols. Preregistration of the study and its protocol would allow others to harmonize and compare methods. Human trial registration has been the standard for over a decade⁴¹. Second, it highlights the need to improve and standardize the quality of reporting: there is a heterogeneity in how the data is analysed and presented, adding another layer of difficulty to compare studies. Even within the same outcome, assay, and experiment, aggregating data and showing medians or means gives different results. The choice of the authors is often not transparent and seems ad hoc. Third, sharing research and data, preferably raw data, allows to make these comparisons more accurate. Especially during disease outbreaks, rapid open access sharing of manuscripts and its underlying data is of utmost importance⁴². It does not only increase transparency and reproducibility, but helps tremendously in the interpretation and aggregation of data.

The development of SARS-CoV-2 vaccines is far from finished. New virus variants will continue to emerge and vaccines might become less effective against these new strains as was shown for the ChAdOx1 vaccine and its marked reduced efficacy against the B.1.351 variant⁴³. NHP trials might hold a key in predicting efficacy, however, only with standardized methods will studies be comparable and results could be used to be translated into guidance. Transmission experiments similar to those in ferrets⁴⁴, will help correlate persistence of viral RNA to actual infectiousness. Furthermore, we need to continue to improve our understanding on how the results of NHP models translate to 'real world' vaccine efficacy, in humans. For that, we need data. Importantly, as discussed above, we need to understand how effective these vaccines are in preventing infectiousness. We need carefully designed trials and observational studies^{8,45}. We need to document the secondary attack rates originating from the vaccinated and unvaccinated groups; however, selective vaccination of risk groups complicates the interpretation of these comparisons. More advanced study designs, comparing populations with different vaccine coverages, could provide additional estimates of the vaccine efficacy on infectiousness⁴⁵. Early observations from Israel, where the vaccine roll-out has been rapid, reveals that individuals vaccinated with the BNT162b2 mRNA vaccine had markedly reduced RNA titres in the blood from 12 days post-vaccination. Consistent with these findings was an observed drop of the average CT values as vaccination roll-out progressed⁴⁶. These findings suggest a reduced infectiousness after vaccination.

In conclusion, preclinical challenge studies show that the majority of SARS-CoV-2 vaccines might not fully protect against infectiousness. Careful assessment of secondary attack rates from vaccinated individuals is warranted. Standardization of trial design and reporting will help contribute to the comparability and interpretation of preclinical trials.

Methods

Research question

We defined the elements of the research question by PICO (Population, Intervention, Control, Outcome) question⁴⁷: The population of interest was defined as 'any non-human primates'. The intervention of interest as 'vaccination against SARS-CoV-2, followed by a challenge experiment'; experiments needed to compare against a control group. The outcome of interest was 'vaccine efficacy'. We use the reduction in shedding of viral RNA as proxy for the vaccine efficacy to prevent infectiousness (Figure 4). The reduction in clinical and pathological findings is interpreted as the vaccine efficacy to prevent disease.

Figure 4. The relationship between the challenge experiments' outcomes (A) and the vaccine efficacy (B). (A) The timeline of a typical non-human primate vaccination-challenge experiment. After challenge, viral shedding and clinical parameters are collected until time of euthanasia, after which pathological examination is performed. (B) Vaccine efficacy (VE) can help: (i) Protect an individual (1) against infection (VE_S), (ii) Prevent infectiousness (VE_I) and prevent transmission to a second individual (2) for which viral shedding can be a proxy measure, (iii) Prevent illness (VE_P) for which clinical or pathological findings can serve as indicator.

Search strategy

We searched the electronic databases PubMed, EMBASE, BioRxiv, and MedRxiv as indexed by the 'COVID Open Access Project'⁴⁸. We filtered the dataset of COVID-19 related research based on keywords in the title and or abstract: "(animal OR mouse OR mice OR hamster OR macaque OR primate OR monkey OR preclinical OR pre-clinical) AND (vaccine OR immunization OR vaccination)". The full review protocol is available on PROSPERO (registration: CRD42021231199)⁴⁹.

We imported the deduplicated search results into the online screening tool 'Rayyan'⁵⁰. Additionally to the search-strategy, we checked the reference lists of the included studies and review articles on the same topic for relevant publications. We included studies published before January 20, 2021.

Eligibility criteria

We included studies that assessed COVID-19 vaccine candidates in preclinical NHP models, and that conducted a challenge study in both preprint and peer-reviewed literature. We included animal models in any of the following species: Rhesus macaques (*Macaca mulatta*), cynomolgus macaques (*Macaca fascicularis*), African green monkeys (*Chlorocebus aethiops*), and other non-human primates. We included animal models regardless of the sex any age of the animals. We only included studies that compared a control group with the vaccinated group or groups. Studies that did not provide extractable data on the presence of viral RNA after challenge were excluded.

Study selection and data extraction

Two reviewers screened titles and abstracts of the retrieved papers independently. If retained, we assessed the full text of the study. One reviewer extracted data into piloted extraction forms in Excel (version 16.0.12527). A second reviewer reviewed exclusion decisions and data entry. Disagreement was resolved by discussion between the reviewers.

We extracted the full citation information, and relevant information on the methods used in the different studies (the full data is available in Supplementary Text 1). We extracted the results either from the text, figures or the supplementary data files. For studies that did not provide the full data, we used WebPlotDigitizer⁵¹ to extract the data from the figures. We collected the type of RNA that was assessed (subgenomic or genomic), from which anatomical sites the samples came, and the raw values of number of viral RNA copies at all provided time-points. For the reported results of vaccination on reduction in pathological and clinical changes, we summarized the direction of the effect (increased, null, or reduced) and whether the reporting was qualitative or quantitative.

To allow the comparison between studies, we standardized the nomenclature of groups. Groups that were challenged, but that did not receive a vaccine are considered 'control' groups, we provide the corresponding name of the intervention groups in the supplementary Table 1.

We also standardized the anatomical sites where samples were taken. We grouped nasal swabs, nasal washes, and nasal brushes as "nasal samples (NS)", tracheal swab and tracheal brushes as "tracheal samples (TS)". Anal and rectal swabs were grouped as "anal samples". We assumed that the nasal and tracheal samples provided information on the upper respiratory tract; the BAL provided insight in the presence of virus in the lower respiratory tract. Anal samples provided proof of systemic infection. To standardize the challenge dose in plaque forming units (PFU), we assumed that the Median Tissue Culture Infectious Dose (TCID₅₀) was approximated by halve of the PFU value⁵².

Synthesis of the evidence

We provided a narrative synthesis of the methodology of the different challenge experiments. For the outcome 'vaccine efficacy to reduce infectiousness', we used the reduction of shedding of RNA as measured by RT-PCR in the different anatomical sites as a proxy measure. We assessed the relative reduction in viral RNA shedding by calculating the difference in AUC between the vaccinated group and the control group. We defined the relative difference as the AUC of the vaccinated group minus the AUC of the control group, divided by the AUC of the control group. For visualization purposes we grouped the difference in AUC using 'Jenks natural breaks' into three groups: low (<18%), moderate (18-44%) and high (>44%) reduction of viral RNA excretion⁵³. We explored the relationship between relative reduction in AUC and the challenge dose visually. All analyses were performed using R (version 4.0.3)⁵⁴. We assessed the efficacy in the prevention of disease, by summarizing the reported results on clinical outcomes and pathological changes. We tabulated the description of the outcome, and whether the outcome was quantitatively or qualitatively described. We report this review following the Preferred Reporting Items for Systematic review and Meta-Analysis (PRISMA)⁵⁵.

References

- 1 World Health Organization. *Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV)*, <[https://www.who.int/news/item/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-\(2019-ncov\)](https://www.who.int/news/item/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov))> (2020).
- 2 United States Food & Drug Administration. *United States Food & Drug Administration: Pfizer-BioNTech COVID-19 Vaccine*, <<https://www.fda.gov/emergency-preparedness-and-response/coronavirus-disease-2019-covid-19/pfizer-biontech-covid-19-vaccine>> (2021).
- 3 Reuters. *Sinovac's coronavirus vaccine candidate approved for emergency use in China*, <<https://www.reuters.com/article/us-health-coronavirus-china-vaccines-idUSKBN2500Z3>> (2020).
- 4 Burki, T. K. The Russian vaccine for COVID-19. *The Lancet Respiratory Medicine* **8**, e85-e86, doi:10.1016/S2213-2600(20)30402-1 (2020).
- 5 COVID19 Vaccine Tracker Team. *COVID19 Vaccine Tracker*, <<https://covid19.trackvaccines.org/>> (2021).
- 6 Orenstein, W. A. *et al.* Field evaluation of vaccine efficacy. *Bulletin of the World Health Organization* **63**, 1055-1068 (1985).
- 7 Basta, N. E., Halloran, M. E., Matrajt, L. & Longini, I. M., Jr. Estimating Influenza Vaccine Efficacy From Challenge and Community-based Study Data. *American Journal of Epidemiology* **168**, 1343-1352, doi:10.1093/aje/kwn259 (2008).
- 8 Halloran, M. E., Struchiner, C. J. & Longini, I. M., Jr. Study designs for evaluating different efficacy and effectiveness aspects of vaccines. *Am J Epidemiol* **146**, 789-803, doi:10.1093/oxfordjournals.aje.a009196 (1997).
- 9 Lipsitch, M. & Dean, N. E. Understanding COVID-19 vaccine efficacy. *Science* **370**, 763-765, doi:10.1126/science.abe5938 (2020).
- 10 Buitrago-Garcia, D. *et al.* Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis. *PLOS Medicine* **17**, e1003346, doi:10.1371/journal.pmed.1003346 (2020).
- 11 Klasse, P. J., Nixon, D. F. & Moore, J. P. Immunogenicity of clinically relevant SARS-CoV-2 vaccines in non-human primates and humans. *Science Advances*, eabe8065, doi:10.1126/sciadv.abe8065 (2021).
- 12 Munoz-Fontela, C. *et al.* Animal models for COVID-19. *Nature* **586**, 509-515, doi:10.1038/s41586-020-2787-6 (2020).
- 13 Rockx, B. *et al.* Comparative pathogenesis of COVID-19, MERS, and SARS in a nonhuman primate model. *Science* **368**, 1012-1015, doi:10.1126/science.abb7314 (2020).
- 14 Munster, V. J. *et al.* Respiratory disease in rhesus macaques inoculated with SARS-CoV-2. *Nature* **585**, 268-272, doi:10.1038/s41586-020-2324-7 (2020).
- 15 Bewley, K. R. *et al.* Immunological and pathological outcomes of SARS-CoV-2 challenge after formalin-inactivated vaccine immunisation of ferrets and rhesus macaques. *BioRxiv*, doi:10.1101/2020.12.21.423746 Preprint at (2020).
- 16 Brouwer, P. J. M. *et al.* Two-component spike nanoparticle vaccine protects macaques from SARS-CoV-2 infection. *Cell*, doi:10.1016/j.cell.2021.01.035 (2021).
- 17 Corbett, K. S. *et al.* Evaluation of the mRNA-1273 Vaccine against SARS-CoV-2 in Nonhuman Primates. *N Engl J Med* **383**, 1544-1555, doi:10.1056/NEJMoa2024671 (2020).
- 18 Feng, L. *et al.* An adenovirus-vectored COVID-19 vaccine confers protection from SARS-COV-2 challenge in rhesus macaques. *Nat Commun* **11**, 4207, doi:10.1038/s41467-020-18077-5 (2020).
- 19 Gabitzsch, E. *et al.* Complete Protection of Nasal and Lung Airways Against SARS-CoV-2 Challenge by Antibody Plus Th1 Dominant N- and S-Specific T-Cell Responses to Subcutaneous Prime and Thermally-Stable Oral Boost Bivalent hAd5 Vaccination in an NHP Study. *BioRxiv*, doi:10.1101/2020.12.08.416297 Preprint at <https://www.biorxiv.org/content/10.1101/2020.12.08.416297> (2020).
- 20 Gao, Q. *et al.* Development of an inactivated vaccine candidate for SARS-CoV-2. *Science* **369**, 77-81, doi:10.1126/science.abc1932 (2020).
- 21 Guebre-Xabier, M. *et al.* NVX-CoV2373 vaccine protects cynomolgus macaque upper and lower airways against SARS-CoV-2 challenge. *Vaccine* **38**, 7892-7896, doi:10.1016/j.vaccine.2020.10.064 (2020).
- 22 Liang, J. G. *et al.* S-Trimer, a COVID-19 subunit vaccine candidate, induces protective immunity in nonhuman primates. *Nat Commun* **12**, 1346, doi:10.1038/s41467-021-21 (2021).
- 23 Mercado, N. B. *et al.* Single-shot Ad26 vaccine protects against SARS-CoV-2 in rhesus macaques. *Nature* **586**, 583-588, doi:10.1038/s41586-020-2607-z (2020).
- 24 Patel, A. *et al.* Intradermal-delivered DNA vaccine provides anamnestic protection in a rhesus macaque SARS-CoV-2 challenge model. *BioRxiv*, doi:10.1101/2020.07.28.225649 Preprint at <https://www.biorxiv.org/content/10.1101/2020.07.28.225649> (2020).
- 25 Rauch, S. *et al.* mRNA vaccine CVnCoV protects non-human primates from SARS-CoV-2 challenge infection. *BioRxiv*, doi:10.1101/2020.12.23.424138 Preprint at <https://www.biorxiv.org/content/10.1101/2020.12.23.424138> (2020).
- 26 Sanchez-Felipe, L. *et al.* A single-dose live-attenuated YF17D-vectored SARS-CoV-2 vaccine candidate. *Nature* **590**, 320-325, doi:10.1038/s41586-020-3035-9 (2021).
- 27 Seo, Y. B. *et al.* Soluble Spike DNA Vaccine Provides Long-Term Protective Immunity against SARS-CoV-2 in Mice and Nonhuman Primates. *Vaccines* **9**, 307, doi:10.3390/vaccines9040307 (2020).

- 28 Solforosi, L. *et al.* Immunogenicity of one- and two-dose regimens of the Ad26.COVS.2 COVID-19 vaccine candidate in adult and aged rhesus macaques. *BioRxiv*, doi:10.1101/2020.11.17.368258 Preprint at <https://www.biorxiv.org/content/10.1101/2020.11.17.368258> (2020).
- 29 Sun, Y. *et al.* Recombinant vaccine containing an RBD-Fc fusion induced protection against SARS-CoV-2 in nonhuman primates and mice. *Cell Mol Immunol* **18**, 1070–1073, doi:10.1038/s41423-021-00658-z (2021).
- 30 van Doremalen, N. *et al.* ChAdOx1 nCoV-19 vaccine prevents SARS-CoV-2 pneumonia in rhesus macaques. *Nature* **586**, 578–582, doi:10.1038/s41586-020-2608-y (2020).
- 31 van Doremalen, N. *et al.* Intranasal ChAdOx1 nCoV-19/AZD1222 vaccination reduces shedding of SARS-CoV-2 D614G in rhesus macaques. *BioRxiv*, doi:10.1101/2021.01.09.426058 Preprint at <https://www.biorxiv.org/content/10.1101/2021.01.09.426058> (2021).
- 32 Vogel, A. B. *et al.* BNT162b vaccines protect rhesus macaques from SARS-CoV-2. *Nature*, doi:10.1038/s41586-021-03275-y (2021).
- 33 Wang, H. *et al.* Development of an Inactivated Vaccine Candidate, BBIBP-CorV, with Potent Protection against SARS-CoV-2. *Cell* **182**, 713–721.e719, doi:10.1016/j.cell.2020.06.008 (2020).
- 34 Yang, J. *et al.* A vaccine targeting the RBD of the S protein of SARS-CoV-2 induces protective immunity. *Nature* **586**, 572–577, doi:10.1038/s41586-020-2599-8 (2020).
- 35 Yu, J. *et al.* DNA vaccine protection against SARS-CoV-2 in rhesus macaques. *Science* **369**, 806–811, doi:10.1126/science.abc6284 (2020).
- 36 Böszörményi, K. P. *et al.* Comparison of SARS-CoV-2 infection in two non-human primate species: rhesus and cynomolgus macaques. *bioRxiv*, 2020.2011.2005.369413, doi:10.1101/2020.11.05.369413 Preprint at <https://www.biorxiv.org/content/10.1101/2020.11.05.369413> (2020).
- 37 Guyatt, G. H. *et al.* GRADE guidelines: 8. Rating the quality of evidence--indirectness. *J Clin Epidemiol* **64**, 1303–1310, doi:10.1016/j.jclinepi.2011.04.014 (2011).
- 38 Dagotto, G. *et al.* Comparison of Subgenomic and Total RNA in SARS-CoV-2 Challenged Rhesus Macaques. *Journal of Virology*, JVI.02370-02320, doi:10.1128/jvi.02370-20 (2021).
- 39 Alexandersen, S., Chamings, A. & Bhatta, T. R. SARS-CoV-2 genomic and subgenomic RNAs in diagnostic samples are not an indicator of active replication. *Nat Commun* **11**, 6059, doi:10.1038/s41467-020-19883-7 (2020).
- 40 Yu, P. *et al.* Age-related rhesus macaque models of COVID-19. *Animal Model Exp Med* **3**, 93–97, doi:10.1002/ame2.12108 (2020).
- 41 De Angelis, C. *et al.* Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *N Engl J Med* **351**, 1250–1251, doi:10.1056/NEJMe048225 (2004).
- 42 Moorthy, V., Henao Restrepo, A. M., Preziosi, M. P. & Swaminathan, S. Data sharing for novel coronavirus (COVID-19). *Bull World Health Organ* **98**, 150, doi:10.2471/BLT.20.251561 (2020).
- 43 Madhi, S. A. *et al.* Efficacy of the ChAdOx1 nCoV-19 Covid-19 Vaccine against the B.1.351 Variant. *New England Journal of Medicine*, doi:10.1056/NEJMoa2102214 (2021).
- 44 Richard, M. *et al.* SARS-CoV-2 is transmitted via contact and via the air between ferrets. *Nat Commun* **11**, 3496, doi:10.1038/s41467-020-17367-2 (2020).
- 45 Longini, I. M., Jr., Sagatelian, K., Rida, W. N. & Halloran, M. E. Optimal vaccine trial design when estimating vaccine efficacy for susceptibility and infectiousness from multiple populations. *Stat Med* **17**, 1121–1136, doi:10.1002/(sici)1097-0258(19980530)17:10<1121::aid-sim824>3.0.co;2-e (1998).
- 46 Petter, E. *et al.* Initial real world evidence for lower viral load of individuals who have been vaccinated by BNT162b2. *medRxiv*, 2021.2002.2008.21251329, doi:10.1101/2021.02.08.21251329 Preprint at <https://www.medrxiv.org/content/medrxiv/early/2021/02/08/2021.02.08.21251329.full.pdf> (2021).
- 47 McKenzie, J. *et al.* in *Cochrane Handbook for Systematic Reviews of Interventions* Vol. 6.2 Ch. Chapter 3.
- 48 COVID-19 Open Access Project. *Living Evidence on COVID-19*, <<https://ispmbern.github.io/covid-19/living-review/>> (2020).
- 49 PROSPERO. *Quantifying vaccine efficacy of SARS-CoV-2 candidate vaccines in animal models from preclinical studies – a systematic review - PROSPERO Protocol CRD42021231199*, <https://www.crd.york.ac.uk/prospere/display_record.php?ID=CRD42021231199> (2021).
- 50 Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev* **5**, 210, doi:10.1186/s13643-016-0384-4 (2016).
- 51 Rohatgi, A. *WebPlotDigitizer User Manual*, <<https://automeris.io/WebPlotDigitizer/userManual.pdf>> (2020).
- 52 Wulff, N. H., Tzatzaris, M. & Young, P. J. Monte Carlo simulation of the Spearman-Kärber TCID50. *J Clin Bioinforma* **2**, 5, doi:10.1186/2043-9113-2-5 (2012).
- 53 Jenks, G. F. The Data Model Concept in Statistical Mapping. *International Yearbook of Cartography* **7**, 186–190 (1967).
- 54 R Core Team. *R: A language and environment for statistical computing.*, <<https://www.R-project.org/>> (2020).
- 55 Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. & Prisma Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* **6**, e1000097, doi:10.1371/journal.pmed.1000097 (2009).

Acknowledgements

MJC, MAdSS, and JLG received salary support from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 773830. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

Protocol/methods: MJC, JLG; **Screening/Data extraction:** MJC, MAdSS; **Interpretation of results:** MJC, JLG, KS, WvdP; **Writing of the first draft:** MJC; **Writing/editing manuscript:** MJC, JLG, MAdSS, KS, WvdP.

Competing interests

None declared.