

別添 3

厚生労働行政推進調査事業費補助金（厚生労働科学特別研究事業） 総括研究報告書

生物統計学的な観点からのワクチン開発における治験計画の立案の迅速化のための研究

分担研究：選択し得る治験デザイン・計画の候補の検討と性能評価

研究分担者： 安藤 友紀

所属： 独立行政法人医薬品医療機器総合機構

研究要旨：

本研究では、全世界に様々な側面から影響を与えた新型コロナウイルス感染症と、そのワクチン開発に関する経験や現状を踏まえ、将来的な新興感染症に対するワクチンを開発する際に想定される状況、状況に応じて考慮すべき要素、有用と考えられる治験デザインについて検討した。また、有効性に関する早期判断が必要となる状況を想定し、その際に採用し得るデザインについて、シミュレーションを用いて性能評価を実施した。

特に初期のワクチン開発では感染症等に関する情報量も少なく、また、感染症をとりまく環境は時間的に大きく変化すると考えられる。治験実施途中での判断やデザインの変更を可能とする治験デザイン、最近議論されている治験デザインや解析手法を状況に応じて利用することが有用であると考えられた。また、迅速なワクチン開発のためには、多地域・多ワクチンが関わる可能性もある大規模な治験の計画や実施が可能な実施体制が構築可能であること、最終的に有効性の検証に至るエビデンスの構築のため、各種要因の変化を可能な限り事前に考慮した治験計画が可能であることは重要であると考えられる。感染症の状況によっては有効性に関する早期の判断を行ってワクチンを供給すべきとされる可能性も踏まえ、治験途中での意思決定に伴う不確実性等について、必要時に議論できるよう予め認識しておくことが重要である。平時にこのような想定に関する議論や検討を行うことにより、実際の感染症発生時のワクチン開発者及び審査担当者の双方による必要な対応が可能となると考える。

A. 研究目的

新型コロナウイルス感染症は全世界に様々な側面から甚大な影響を与え、その感染のインパクトの大きさから、各国における新規ワクチン開発は産官学並びに社会的な支持もあり平時の比ではない勢いで進められてきた。本邦においても迅速なワクチンの導入が求められ、海外で開発されたワクチンの薬事承認申請及び承認手続きが迅速に進められてきている。しかしながら2022年3月末時点で、本邦で開発された国産ワクチンは実用化に至っておらず、その開発体制の様々な課題が浮き彫りにされている。課題の一部として、想定外の状況下で迅速に治験デザイン・計画を立案する体制が不十分であること、すでに確立したワクチンが存在する状況下で新規ワクチンを開発する際の治験デザイン・計画のあり方についての検討が必要であることが挙げられる。

本分担研究では、新型コロナウイルス感染症発生以降の状況や既承認の新型コロナウイルスワクチンの開発の経緯等を参考に、将来的な新興感染症の発生時に迅速にワクチンを開発する際に考慮すべき要素、取り得る治験デザインについて検討した。また、新興感染症の発生初期には、ワクチンの有効性に関する迅速な判断、承認等の意思決定が必要とされることを想定し、考えられるシナリオに対する早期判断のためのデザインについて、その性能評価を行った。

B. 研究方法

1. 新興感染症のワクチンを開発する際に考慮すべき要素及び取り得る治験デザ

イン

新型コロナウイルス感染症発生以降のワクチン開発に関する、本邦において既承認のワクチンの公開情報、本邦を含む各地域及びWHO等による国際的な規制や推奨事項、関連する国際的なワークショップ等における議論の内容等を踏まえ、新興感染症に対するワクチンを開発する際に想定される状況、状況に応じて考慮すべき要素、有用と考えられるデザインについて検討した。

2. 有効性早期判断のためのデザインの性能評価

1での検討を踏まえ、新興感染症に対するワクチンの開発初期には、関連する情報が少ない中でも、開発されたワクチンの有効性に関する迅速な判断が期待されると考えられた。よって、有効性の早期判断のために事前に計画された中間解析の利用及び必ずしも事前に計画されていないが感染状況等を踏まえた公衆衛生上の必要性から有効性に基づく早期中止を行う状況を取り上げ、新興感染症として考えられるシナリオにおける性能についてシミュレーションを用いた試行的な検討を行った。

新興感染症に対するワクチン開発を想定するにあたり、ワクチン開発の蓋然性が高いと想定される感染症のタイプに関する感染症専門医へのインタビュー結果に基づき、新型インフルエンザ及び新型コロナウイルス感染症を含む呼吸器感染症に関するシナリオを検討することとし、感染性が高いが致死率が比較的低い新型コロナウイルス感染症を例としたシナリオ及び感染性は限定的であるが致死率が比較的高い中東呼吸器症候群(MERS)を例としたシナリオについ

でそれぞれ検討した。シミュレーションにより検討する試験としては、先行する治験により有効性が期待でき安全性が許容されると考えられる適切な用量が設定されている前提で、承認申請のための検証的試験として計画され終了までに比較的長期間を必要とすると考えられる、感染症発症までの時間を主要評価項目としたプラセボ対照試験を想定した。感染症の蔓延状況により、有効性に関する早期の判断と承認に関わる意思決定が必要となる場合を想定し、予定した中間解析及び予定していない早期中止について、その動作特性をシミュレーションにより確認した。また、新型コロナウイルス感染症を例としたシナリオについては、開始から短期間（Carbis Bay G7 Summitにおいて提案された 100 Days mission¹⁾を参考として 3 か月）で結論が得られる治験を実施する場合のデザインを検討した。

各シナリオについて想定するパラメータ、シミュレーション条件を含む検討の詳細については、別添 1 に示す。

C. 研究結果

1. 新興感染症のワクチンを開発する際に考慮すべき要素及び取り得る治験デザイン

新興感染症の流行に伴い、ワクチン開発のための治験の計画に際して予め想定し考慮すべき点、及び関連して取り得る治験のデザインについては以下のとおりである。

• 感染症に関する情報量

新興感染症の発生から流行の初期にかけては、感染症の感染力、発症機序、発症後の

重症化の程度、致死率等の情報が十分ではない。よって、初期に開発されるワクチンについては、検証的試験として基本的には感染症発症といった臨床的なイベントを主要評価項目として発症予防効果を検討する、プラセボ対照ランダム化比較試験により有効性及び安全性を評価する必要がある²⁾。この場合、感染症の実際の感染力の高さや発症率等の情報が少ない状況で、比較的大規模な治験を計画することになり得るため、逐次群計画（中間解析）や、治験途中での治験計画の変更を可能とするアダプティブデザインを利用し、治験実施中の有効性の評価に基づく有効中止／無効中止といった意思決定や被験者数の再設定等のデザインの変更を可能とする計画とすることは有用と考えられる。これらの治験途中の中止や計画変更については、変更内容やその基準を事前に計画しておくことが重要である。

感染症及びワクチン双方の情報量が少ない状況では、複数の小規模な治験による複数用量の検討、年齢や背景因子による複数の投与対象（コホート）の検討が、特定地域での治験により行われる可能性がある。この場合、個々の治験が特定の目的を持ち、かつ一定の共通性のあるデザインを取ることで、事前に計画された統合解析により、より多くの被験者数を必要とする発症予防効果の評価を行うことも考えられる。このような統合解析を検証的な結果として取り扱うためには、理想的には統合する各治験の計画段階で統合可能性を考慮したデザインとし、具体的な統合解析の方法等についてはいずれの治験の結果も明らかではない時点で計画し規定しておくことが重要である。また、結果の評価においては、治験毎の

結果も踏まえて統合の妥当性について説明する必要がある。

ワクチンの開発や発症予防の機序等の研究が進むのに伴い、ワクチン接種後の中和抗体価と発症予防効果の相関に関する情報が集積されることにより、有効性の代替指標として免疫原性の指標を主要評価項目とした有効性の評価が可能となることが考えられる³⁾。ただし、ワクチンが予防効果を示す免疫原性の指標の閾値や、予防効果における細胞性免疫の寄与等について一定の情報が必要となることに留意が必要である。また、安全性情報の十分性にも配慮が必要である。

予防効果に関連する免疫原性の指標や感染症に関する一定の疫学的情報の蓄積により、追加投与の効果のような特定の検討内容については、必ずしもランダム化されていない対照(外部対照)との比較により一定の評価が可能となる可能性がある。

- 感染症対策や変異株等による感染率等の変化

感染症対策の推進による感染率及び発症率の低下や、感染力の異なる変異株の流行等により、特に感染症流行初期のワクチンの治験の主要評価項目となる発症率等が計画時に想定していた値から変化する可能性がある。このような変化に対応するため、発症率等に関する情報収集に基づき盲検下で被験者数を変更することが考えられる。検証的試験については、考えられ得る計画の変更についてその要否の判断方法も含め、予め計画しておくことが有用である。

なお、特定のワクチンの有効性及び安全性に関する実施地域や時期の異なる複数の

治験の評価や統合解析の評価、治験の外部情報の利用に際しても、感染症対策や変異株等による発症率の差異や変化の影響があることには留意が必要である。

- 既承認ワクチン接種率の増加

感染症発生早期に開発が開始され、発症予防効果が検証され承認されたワクチン(初発ワクチン)の接種が進むことにより、後続ワクチンの治験の被験者となり得る対象者が減少し、実施中の治験について被験者の集積率が低下する可能性がある。また、既承認ワクチンが存在することにより、実施中のプラセボ対照試験のデザインについて配慮が必要となる可能性や⁴⁾、新たにプラセボ対照臨床試験を計画、実施することが困難となる可能性がある。一定時期以降の後続ワクチンの検証的試験は、既承認ワクチンを対照とした非劣性試験として計画、実施することが現実的となると考えられる³⁾。この時点で計画される治験としては、前述「感染症に関する情報量」の内容も踏まえ、計画段階で利用可能な有効性の指標について検討した上で主要評価項目を設定し、臨床的に適切な非劣性マージンを設定する必要がある。

- 同時に開発されるワクチンの増加

後続ワクチンの開発段階では、「既承認ワクチン接種率の増加」で述べたとおり、既承認ワクチンの接種が進むことによる潜在的な被験者の減少に加え、複数の後続ワクチンが同時に開発されることにより、さらに被験者の集積が困難となる可能性がある。

複数のワクチンが比較的短期間に同時に開発される状況では、それらの有効性評価

をより効率的に行うために、一つの共通した治験実施計画書(マスタープロトコル)の下で、共通するプラットフォームを用いて複数の被験ワクチンを評価することも可能と考えられ(プラットフォーム試験^{3),5)}、特に既承認ワクチンを共通の実対照ワクチンとして用いて評価を行う場合に有用であると考えられる。プラットフォーム試験では、複数の企業が開発するワクチンが評価され、場合によっては大規模かつ長期にわたるプラットフォームの管理が必要となるため、プラットフォームを通じた治験の管理が適切になされるよう、治験実施体制の整備が重要である。また、「感染症対策や変異株等による感染率等の変化」で述べた感染率等の時期による影響を想定すると、複数の被験ワクチンに対して共通の対照ワクチン群の結果を比較対象として用いる場合でも、基本的には各被験ワクチンについて同時期に組み入れられた実対照ワクチン群の被験者との比較を行うことが望ましいと考えられる。

- その他の要素

新興感染症が世界的に流行する状況では、開発中のワクチンに関する海外データの利用や、本邦を含む国際共同治験⁶⁾による評価も承認申請において重要な位置付けとなる。評価に際しては地域による感染状況や変異株の分布、ワクチン接種率の違い等を考慮する必要がある。その他の民族的要因⁷⁾についても事前に検討しておくことが重要である。

この他、新型コロナウイルス感染症に対するワクチンの治験では、検証的試験の主要な解析として従来一般的に用いられてい

る統計的仮説検定ではなく、ベイズ流統計手法を用いている例がある⁸⁾。当該手法の利用は、ワクチンの治験において有効性に関する柔軟な意思決定を行う上で有用な場合があると考えられる。また、アダプティブデザインを利用した試験やプラットフォーム試験の解析手法としても利用されてきている。ベイズ流統計手法の利用に際しては、用いる事前情報の適切性や解析の動作特性の評価等が重要であり、規制当局との事前の協議が推奨される。

2. 有効性早期判断のためのデザインの性能評価

新型コロナウイルス感染症シナリオについて、ワクチン群とプラセボ群の割付比を1:1とするランダム化比較試験において、プラセボ群の6ヶ月時点の発症割合を1%、VEの期待値を60%として、ハザード比に基づくVEの95%信頼区間の下限が30%を超えること⁹⁾を検証する状況を想定した。組入れ期間3ヶ月、各被験者の観察期間の最大期間を6ヶ月と設定し、試験開始から9ヶ月時点で有効性の評価が実施できる被験者数を算出したところ¹⁰⁾、19350例が必要とされた。なお、VEを50%~70%、プラセボ群の6ヶ月時点での発症割合を0.50%~1.50%に変動させたところ、必要被験者数は99959~6058例と変動したが、ほとんどの組み合わせで10000例を超えていた(別添1、表1)。早期有効中止を目的とした中間解析をイベント数が全体の50%、75%の時点で実施した場合、それぞれ試験開始から解析までの時間は5ヶ月、6ヶ月と短縮されたが、各時点での検出力は20%、

60%であった。なお、中間解析の多重性の調整方法としては Lan-DeMets 法に基づく O'Brien-Fleming 型の α 消費関数を使用している。一方、多重性調整を行わない公衆衛生上の理由での早期中止をイベント数が全体の 40%、60%、80%の時点で実施した場合、それぞれ解析までの時間は 4 ヶ月、5 ヶ月、6 ヶ月と短縮されたが、各時点での検出力は 46%、65%、76%となった（別添 1、表 2）。シミュレーション結果において早期終了 2（60%時点、81 イベント）の α エラーが 5.38%と 5%を超えていたが、シミュレーション回数を増加させて確認したところ、シミュレーション回数の不足のためと考えられた（別添 1、表 3）。

同様の条件で 3 ヶ月以内の試験終了を想定した場合（組入れ期間 1 ヶ月、各被験者の観察期間の最大期間を 2 ヶ月）には、必要なイベント数 135 イベントに対して、必要被験者数は 57388 例と算出された。試験開始から解析までの時間は中間解析から最終解析で 2.0~3.0 ヶ月、公衆衛生上の理由での早期中止では 1.8~2.6 ヶ月と短縮されたが、検出力等は前述の結果と変わらなかった（別添 1、表 4）。ロバスト分散を用いたポアソン回帰¹¹⁾によるリスク比に基づく解析も実施したが、結果は変わらなかった（別添 1、表 5）。

MERS シナリオについては、イベントを死亡、プラセボ群の 1 ヶ月時点の死亡割合を 35%とし、その他の設定は新型コロナウイルス感染症シナリオと同様として検討した。必要なイベント数は 135 イベントに対して必要被験者数は 2 群 534 例となった。解析までの時間は中間解析から最終解析で 1.0~1.5 ヶ月、公衆衛生上の理由での早期

中止では 0.9~1.3 ヶ月と短縮されたが、各時点での検出力は他の条件と比較して大きな違いはなかった（別添 1、表 6）。

D. 考察

1. 新興感染症のワクチンを開発する際に考慮すべき要素及び取り得る治験デザイン

治験計画時に考慮すべき要素としては、感染症に関する情報量、感染症対策や変異株等による感染率等の変化、既承認ワクチン接種率の増加、開発されるワクチンの増加等、時間的に変化する要因が多い。特に初期のワクチン開発で大規模かつ長期にわたり実施する治験が必要となる場合、十分な体制で当該治験の計画を迅速に開始するとともに、上述の時間的な変化の可能性について予め認識し可能な限り考慮することが重要である。

感染症発生初期のワクチン開発では、感染状況の拡大等により、治験実施中にワクチンの有効性に関する早期の意思決定のための中間解析や計画変更のためのアダプティブデザインは有用と考えられる。また、初期に開始した治験の統合解析により発症予防効果等を評価することも考えられる。特にこれらのデザインについては共通して、多重性による第一種の過誤確率の増大や、結果に基づくデザイン変更によるバイアスの回避のため、その解析方法や変更内容、基準の詳細が事前に計画され、治験実施計画書等で規定されていることが重要となる。盲検解除を伴わない変更についても、可能な限りその可能性を事前に考慮することが

重要である。既承認ワクチンや開発されるワクチン数の増加等の変化も、有効性について検証すべき内容に影響を与える可能性があるが、治験計画時に適切なデザインを選択するとともに、実施中にこれらの変化が計画に与える影響についても認識し、必要に応じた対応を予め想定しておくこと重要である。

複数の治験の統合や、マスタープロトコルを用いて複数のワクチンをプラットフォーム試験で評価する状況では、一貫した治験管理が適切になされるよう、治験実施計画書やデータ収集等の標準化を含む適切な実施のための治験実施体制の確立が重要である。また、国際共同治験の実施や海外データの利用、さらに最終的に複数地域での承認取得を視野に入れる状況では、治験計画時に考慮すべき要素の各地域における状況を踏まえて、治験の計画や結果の利用を検討できる体制が必要となる。

ワクチンの有効性に関する早期の意思決定が求められる状況においても、ワクチンの安全性評価、さらにはベネフィットリスクのバランスの評価も重要である。治験実施途中の早期の意思決定の際には、その時点での安全性データベースの内容を精査した上で、許容可能なベネフィットリスクバランスがあることを確認する必要がある。特に、有効性が免疫原性の指標で評価される場合には、想定される発症予防効果等を踏まえてそのバランスを検討する必要があることから注意が必要である。

なお、新興感染症に対するワクチン開発にあたっては規制当局への相談は重要であり、特に治験実施中の治験デザインの変更の計画や、その他想定される変更の必要性、

マスタープロトコルやベイズ流統計手法といった比較的新規の手法の利用については、詳細について議論しておくことが有用と考えられる。

2. シミュレーションによる性能評価について

新型コロナウイルス感染症シナリオでは、VE を 50%~70%、プラセボ群の 6 ヶ月時点での発症割合を 0.50%~1.50%の範囲内とすると、多くの場合被験者数が 10000 例以上必要であるとされた。初期のワクチン開発では治験計画時点での情報が少なく被験者数が保守的に見積もられる可能性もあり、治験は大規模となり最終的な結果が得られるまで長期間を要すると予想され、状況の変化への計画変更等による対応の事前の考慮や、有効性による早期中止を事前に考慮することが有用と考えられる。

事前に計画していない公衆衛生上の観点からの早期中止について、感染症蔓延状況等に応じていずれかの段階で早期中止し、その時点の結果で有効性の判断を行うことを想定して検討したが、例えばイベント数が全体の 40%の時点での解析では、検出力も 46%程度と低い状況であった。一般的に、事前に計画されていない治験デザインの変更や試験中止は避けるべきであるが、状況等に応じて検討せざるを得ない場合でも、当該時点での有効性の結果の不確実性について十分に考慮するとともに、当該時点の感染状況、得られている安全性の情報、承認後に入手可能と想定される情報等も含めて判断する必要がある。死亡率が高いとされる MERS を例としたシナリオでは、必要イ

イベント数に対する早期中止時点でのイベント数の割合が同様である場合には、検出力について新型コロナウイルス感染症シナリオと類似の結果が得られた。このシナリオでは、イベント発現割合が高いことから有効性評価に必要な治験の規模は比較的小さく、予定される治験の終了までの期間も短い。しかしながら、感染症の重篤性を考えるとより迅速な判断が必要とされる可能性もある。状況の変化による結果の不確実性について考慮すべき点を事前に想定しておくことで、必要に応じた検討が迅速に行えることが望ましい。

本研究では最近の事例を踏まえて、比較的限られた条件でのシミュレーション検討を行ったが、実際の進行感染症発生時、特に初期のワクチン開発においては、その時点で得られている感染症に関する情報の不確実性も踏まえて、実施する治験について様々な観点での想定、検討が必要となる。治験計画に関する規制当局との相談において、計画中の治験に関する各種条件によるシミュレーション結果等を踏まえて議論を行うことは有益であると考えられる。

E. 結語

本研究では、全世界に様々な側面から影響を与えた新型コロナウイルス感染症と、そのワクチン開発に関する経験や現状を踏まえ、将来的な新興感染症に対するワクチンを開発する際に想定される状況、状況に応じて考慮すべき要素、有用と考えられる治験デザインについて検討した。また、有効性に関する早期判断が必要となる状況を想定し、その際に採用し得るデザインについて、シミュレーションを用いて性能評価を

実施した。

特に初期のワクチン開発では感染症等に関する情報量も少なく、また、感染症をとりまく環境は時間的に大きく変化すると考えられる。ワクチン開発のための治験の計画に際して予め想定し考慮すべき点、及び関連して取り得る治験のデザインについては、その実施を可能とする実施体制の構築について、また、感染症の状況によっては有効性に関する早期の判断を行ってワクチンを供給すべきとされる可能性も踏まえ、治験途中での意思決定に伴う不確実性等については、必要時に議論できるよう予め認識しておくことが重要である。平時にこのような想定に関する議論や検討を行うことにより、実際の感染症発生時のワクチン開発者及び審査担当者の双方による必要な対応が可能となると考える。

謝辞: 本分担研究には、独立行政法人医薬品医療機器総合機構新薬審査第四部の浅野淳一主任専門員、新薬審査第五部の加藤凌輔審査専門員が研究協力者として参加した。

F. 研究発表

特になし

G. 知的所有権の取得状況

特になし

<参考文献>

- 1) G7 Summit 2021, 100 Days mission to respond to future pandemic threats. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/992762/100_D

[ays_Mission_to_respond_to_future_pandemic_threats__3_.pdf](#) (Accessed 23, May 2022).

- 2) 独立行政法人医薬品医療機器総合機構 ワクチン等審査部, 新型コロナウイルス (SARS-CoV-2) ワクチンの評価に関する考え方. 令和2年9月2日.
- 3) 独立行政法人医薬品医療機器総合機構 ワクチン等審査部, 新型コロナウイルス (SARS-CoV-2) ワクチンの評価に関する考え方 (補遺3) 免疫原性に基づく新型コロナウイルスワクチンの評価の考え方. 令和3年10月22日.
- 4) 独立行政法人医薬品医療機器総合機構 ワクチン等審査部. 新型コロナウイルス (SARS-CoV-2) ワクチンの評価に関する考え方 (補遺2) プラセボ対照試験の被験者等に対する倫理的配慮について. 令和3年6月11日.
- 5) International Council for Harmonization of Technical Requirement for Pharmaceuticals for Human Use, E8(R1) General Considerations for Clinical Studies. 6 October, 2021.
- 6) 厚生労働省, 国際共同治験の計画及びデザインに関する一般原則に関するガイドラインについて. 薬生薬審発 0612 第1号, 平成30年6月12日.
- 7) 厚生労働省, 外国臨床データを受け入れる際に考慮すべき民族的要因について. 医薬審第672号平成10年8月11日
- 8) 独立行政法人医薬品医療機器総合機構, コミナティ筋注特例承認に係る報告書. 令和3年2月8日.
- 9) World Health Organization, CONSIDERATIONS FOR EVALUATION OF COVID19 VACCINES – Points to consider for manufacturers of COVID 19 vaccines. 25 November, 2020.
- 10) Schoenfeld D, The Asymptotic Properties of Nonparametric Tests for Comparing Survival Distributions. *Biometrika*, 68, 316-319, 1981.
- 11) Zou G, A Modified Poisson Regression Approach to Prospective Studies with Binary Data. *Am J Epidemiol*, 159:702–706, 2004.

【別添 1】実施したシミュレーションの詳細

新型コロナウイルス感染症シナリオについて、治験デザイン、主要評価項目、有効性の判断基準等は以下のとおりとした。

- 既に確立した用量のワクチンとプラセボを比較する、プラセボ対照ランダム化比較試験を想定する。
- 主要評価項目は既承認の新型コロナウイルス感染症ワクチンの第 3 相試験を参考に、感染症発症確定までの時間とする。
- ハザード比 $HR = \lambda_E / \lambda_P$ を Cox 比例ハザードモデルで推定し、ハザード比に基づくワクチン効果 Vaccine efficacy (VE) ($= (1 - HR) \times 100$) の 95%CI の下限値が 30% を超えるか否かを確認する。
 - t_E : 本剤群の被験者のイベントまでの時間
 - t_P : プラセボ群の被験者のイベントまでの時間
 - λ_E : 本剤群のハザード $t_E \sim Exponential(\lambda_E)$
 - λ_P : プラセボ群のハザード $t_P \sim Exponential(\lambda_P)$
 - 帰無仮説 $H_0 : VE \leq 30\%$
 - ハザード比の下での帰無仮説 $H_0 : HR \geq 0.7$
- 早期有効中止を目的した中間解析の実施を計画する。多重性の調整方法としては Lan-DeMets 法に基づく O'Brien-Fleming 型の α 消費関数を使用する。この計画に際し、必要イベント数の再設定は計画せず、試験開始前に被験者数を保守的に見積もることを想定する。
- 計画された中間解析とは別に、公衆衛生上の理由による事前規定していない早期の試験終了が実施されることを想定する。この場合、結果によらず試験は必ず終了するものとする。例えば、早期有効中止を目的とした中間解析を 50% のイベントが起きた時点で実施したものの、早期中止せずに継続するが、60% のイベントが起きた時点で公衆衛生上の理由により事前規定していない早期の試験終了が起きるといった状況を想定する。

シミュレーション条件については、以下のとおりとした。

- ワクチン群とプラセボ群の割付比は 1:1 とする
- プラセボ群の 6 ヶ月時点の発症割合を 1% と仮定する。プラセボ群の発症割合から λ_P を算出すると、 $\exp(-\lambda_P \cdot 0.5) = 0.99 \Rightarrow \lambda_P = 0.02$ となる。
- VE の期待値を 60% と仮定し、 $\lambda_P = 0.02$ としたことから $\lambda_E = 0.008$ となる。
- 組み入れ期間は 3 ヶ月、各被験者の観察期間の最大期間を 6 ヶ月と設定して、試験開始から 9 カ月時点で有効性の評価が実施できるように被験者数を設定する（観察期間に登録期間は考慮しない）。
- 6 ヶ月時点の脱落割合は 2% と仮定し、指数分布を仮定する ($\lambda_{censor} = 0.04$)。

- 試験開始から組入れまでの期間には一様分布を仮定する (*Uniform(0,3)*)。
- 有意水準は両側 5%とする。この場合、仮説検定の検出力として 90%を確保するために必要なイベント数は 135 イベント、必要被験者数は 2 群 19350 例となった (Schoenfeld の方法、表 1)
- 早期有効中止を目的とした中間解析のタイミングは、イベント数が全体の 50% (68 イベント時点)、75% (102 イベント時点) の時点とした。なお、この場合、両側有意水準は 50%の時点では 0.0042、75%の時点では 0.0194、最終解析では 0.043 となる。
- 公衆衛生上の理由による事前規定していない早期終了のタイミングは、イベント数が全体の 40% (54 イベント時点)、60% (81 イベント時点)、80% (108 イベント時点) の時点を検討することとした。なお、早期終了時の有意水準は残りの全ての α を消費することとした。両側有意水準は 40%の時点では 0.05、60%の時点では 0.048、80%の時点では 0.043 となる。
- 最終解析は 135 イベント時点で実施するものとする。なお、135 イベントを得るために、シミュレーションにおける各被験者の最大の観察期間は 1 年と設定した。
- シミュレーション回数は 10,000 回とした。

なお、Carbis Bay G7 Summit において提案された 100 Days mission を参考として 3 か月以内に試験を終了する場合として、以下の条件でもシミュレーションを実施した。その他の設定については、前述のシミュレーションと同様とした。

- 組み入れ期間は 1 ヶ月、各被験者の観察期間の最大期間を 2 ヶ月と設定して、試験開始から 3 ヶ月時点で有効性の評価を行うこととする (観察期間に登録期間は考慮しない)。
- 有意水準は両側 5%とする。この場合、仮説検定の検出力として 90%を確保するために必要なイベント数 135 イベントに対して、必要被験者数は 2 群 57388 例となる (Schoenfeld の方法)。
- ロバスト分散を用いたポアソン回帰 (Zou G, 2004) によるリスク比に基づく解析も実施した。この場合、リスク比の信頼区間の上限が 0.7 を下回れば有効と判定する。

プラセボ群の 6 ヶ月時点のイベント発現割合及び VE の期待値の組み合わせに対して、有意水準両側 5%、検出力 90%を確保するために必要な被験者数を算出した結果を表 1 に示す。

シミュレーションにより得られた、中間解析及び最終解析、並びに各早期終了における α エラー、検出力、有意になった場合のハザード比 (真値は 0.4) の推定値、及び試験開始から解析までの時間を表 2 に示す。なお、検出力の計算において、各解析時点よりも前に実施された中間解析は考慮されていない。また、シミュレーション結果において早期終了 60% 時点 (81 イベント) の α エラーが 5.38%と当該時点の有意水準である 4.8%を超えていたことから、その原因を検討するため、シミュレーション回数を 2 万回、10 万回と増加させ、

α エラーとの関係を確認した。結果を表 3 に示す。結果から、数値の乖離はシミュレーション回数の不足によるものと考えられた。

3 ヶ月以内の試験終了を想定したシミュレーションの結果を表 4 に示した。また、ロバスト分散を用いたポアソン回帰によるリスク比に基づく解析を行う場合の検討を行い、Cox 比例ハザードモデルの場合との比較を行った結果を表 5 に示した。解析方法による差は見られなかった。

感染性は限定的であるが致死率が比較的高い MERS を想定したシナリオについて、試験デザインは新型コロナウイルス感染症を想定したシナリオと同様としたが、シミュレーション条件は以下のとおりとした。また、以下に示した以外の条件は新型コロナウイルス感染症を想定したシナリオと同様とした

- 主要評価項目は死亡までの時間とする。
- ワクチン群とプラセボ群の割付比は 1:1 とする。
- プラセボ群の 1 カ月時点の死亡割合を 35%と仮定する。プラセボ群の死亡割合から λ_p を算出すると、 $\exp(-\lambda_p \cdot 1) = 0.65 \Rightarrow \lambda_p = 0.43$ となる。
- VE の期待値を 60%と仮定し、 $\lambda_p = 0.43$ としたことから $\lambda_E = 0.17$ となる。
- 組み入れ期間は 1 カ月、各被験者の観察期間の最大期間を 1 カ月と設定して、試験開始から 2 カ月時点で有効性の評価が実施できるように被験者数を設定する（観察期間に登録期間は考慮しない）。
- 1 カ月時点の脱落割合は 4%と仮定し、指数分布を仮定する ($\lambda_{censor} = 0.04$)。
- 試験開始から組み入れまでの期間には一様分布を仮定する ($Uniform(0,1)$)。
- 有意水準は両側 5%とする。この場合、仮説検定の検出力として 90%を確保するために必要なイベント数は 135 イベント、必要被験者数は 2 群 534 例となった (Schoenfeld の方法)。

シミュレーションにより得られた、各解析及び中止時点における α エラー、検出力、有意になった場合のハザード比 (真値は 0.4) の推定値、及び試験開始から解析までの時間を表 6 に示す。

表 1. 試験デザインの仮定と目標被験者数

プラセボ群の 6 カ月時点の発症割合	VEの期待値		
	50%	60%	70%
0.50%	99959	38717	18189
0.75%	66626	25806	12124
1.00%	49959	19350	9091
1.25%	39958	15477	7271
1.50%	33292	12895	6058

表 2. シミュレーション結果（新型コロナウイルス感染症シナリオ）

解析目的	時点 イベント数	α エラー (各解析時点で 用いた有意水準)	検出力	有意になった 場合のハザード比 (真値 0.4)	試験開始から 解析までの時間
中間解析 1	50%時点 68	0.034% (0.420%)	20%	0.26	5 カ月
中間解析 2	75%時点 102	1.94% (1.94%)	60%	0.35	6 カ月
最終解析	100%時点 135	4.30% (4.30%)	84%	0.38	8 カ月
早期終了 1	40%時点 54	4.80% (5.00%)	46%	0.30	4 カ月
早期終了 2	60%時点 81	5.38% (4.80%)	65%	0.35	5 カ月
早期終了 3	80%時点 108	4.30% (4.30%)	76%	0.37	6 カ月

表3. シミュレーション回数と α エラーの関係（下段は各解析時点で用いた有意水準）

解析目的	時点 イベント数	1万回	2万回	10万回
中間解析1	50%時点 68	0.03% (0.42%)	0.25% (0.42%)	0.24% (0.42%)
中間解析2	75%時点 102	1.94% (1.94%)	1.79% (1.94%)	1.62% (1.94%)
最終解析	100%時点 135	4.30% (4.30%)	4.19% (4.30%)	3.84% (4.30%)
早期終了1	40%時点 54	4.80% (5.00%)	4.69% (5.00%)	4.35% (5.00%)
早期終了2	60%時点 81	5.38% (4.80%)	4.91% (4.80%)	4.56% (4.80%)
早期終了3	80%時点 108	4.30% (4.30%)	3.97% (4.30%)	3.74% (4.30%)

表4. シミュレーション結果（新型コロナウイルス感染症シナリオ、3か月以内に試験を終了）

解析目的	時点 イベント数	α エラー (各解析時点で 用いた有意水準)	検出力	有意になった 場合のハザード比 (真値は0.4)	試験開始から 解析までの時間
中間解析1	50%時点 68	0.22% (0.420%)	19%	0.26	2.0 カ月
中間解析2	75%時点 102	1.46% (1.94%)	60%	0.35	2.5 カ月
最終解析	100%時点 135	3.94% (4.30%)	85%	0.38	3.0 カ月
早期終了1	40%時点 54	4.18% (5.00%)	46%	0.31	1.8 カ月
早期終了2	60%時点 81	4.84% (4.80%)	66%	0.35	2.2 カ月
早期終了3	80%時点 108	4.24% (4.30%)	76%	0.37	2.6 カ月

表 5. Cox 比例ハザードモデルとポアソン回帰モデルの検出力の比較（新型コロナウイルス感染症シナリオ）

解析目的	時点 イベント数	検出力	
		Cox 比例ハザードモデル	ロバスト分散を用いた ポアソン回帰モデル
中間解析 1	50%時点 68	18%	18%
中間解析 2	75%時点 102	60%	60%
最終解析	100%時点 135	84%	84%
早期終了 1	40%時点 54	46%	46%
早期終了 2	60%時点 81	65%	65%
早期終了 3	80%時点 108	75%	75%

表 6. シミュレーション結果（MERS シナリオ）

解析目的	時点 イベント数	α エラー (各解析時点で 用いた有意水準)	検出力	有意になった場合 のハザード比 (真値は 0.4)	試験開始から 解析までの時間
中間解析 1	50%時点 68	0.2% (0.420%)	19%	0.27	1.0 カ月
中間解析 2	75%時点 102	1.58% (1.94%)	62%	0.35	1.2 カ月
最終解析	100%時点 135	3.76% (4.30%)	86%	0.38	1.5 カ月
早期終了 1	40%時点 54	3.92% (5.00%)	47%	0.31	0.9 カ月
早期終了 2	60%時点 81	4.26% (4.80%)	65%	0.35	1.1 カ月
早期終了 3	80%時点 108	4.08% (4.30%)	77%	0.37	1.3 カ月