The NEW ENGLAND JOURNAL of MEDICINE

Perspective

APRIL 15, 2021

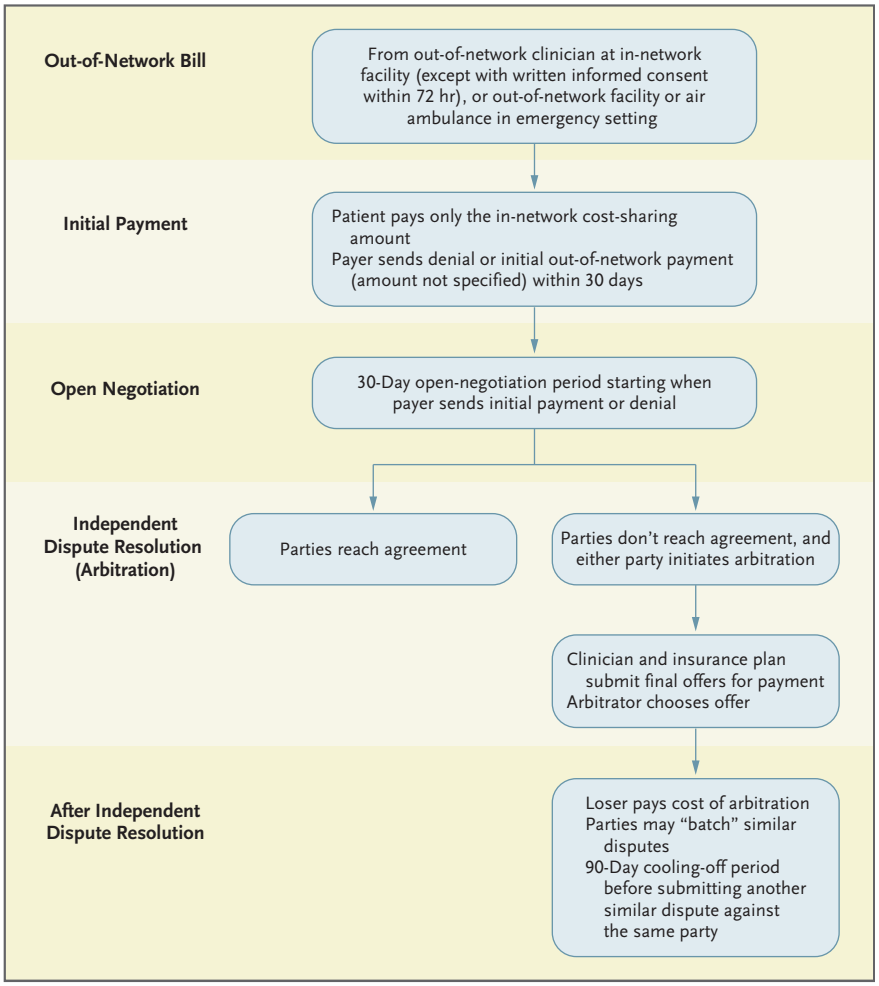
No More Surprises — New Legislation on Out-of-Network Billing

Karan R. Chhabra, M.D., Erin Fuse Brown, J.D., M.P.H., and Andrew M. Ryan, Ph.D.

The passage of the No Surprises Act — which banned "surprise billing" in many scenarios — on December 27, 2020, was an unexpected step forward during an otherwise dysfunctional year of U.S. policymaking. As many as one in five patients visiting an emergency department or undergoing elective surgery receives an out-of-network bill from a clinician whom they had no ability to choose, and more than 70% of ambulance rides are out of network.[1] Since insurance plans aren't required to pay out-of-network providers their full charges, clinicians may bill the patient for the difference between the insurance payment and their charges. These surprise bills can lead to thousands of dollars in unanticipated costs and have been nearly impossible for patients to avoid. Folded into the 2020 year-end spending and Covid-19 relief package, the new legislation will benefit patients and is likely to have little effect on most physicians who don't engage in surprise billing. Its effects on health insurance premiums, networks, and overall health care costs remain unclear, but they could be favorable.

Surveys show that unexpected medical bills are Americans' top financial fear. Nearly a dozen surprise-billing proposals were introduced in the 116th Congress. Despite bipartisan support, these proposals sparked intense disagreement within the health care industry about how out-of-network clinicians should be reimbursed. Insurers, employers, and consumer groups favored setting a benchmark price for services based on in-network rates, whereas hospitals and clinicians favored an arbitration process that would determine reimbursement on a case-by-case basis. In 2019, a year-end compromise was thwarted by a campaign funded by private-equity firms that control large physician-staffing companies notorious for using surprise billing as a business tactic (e.g., Envision and TeamHealth). A year later, Congress acquiesced to many of the demands of physician groups and passed the No Surprises Act.

Effective January 1, 2022, patients receiving out-of-network emergency services, air-ambulance transportation, or out-of-network nonemergency services at in-network facilities may be billed only the amount they would owe for an in-network provider. The law applies to all health plans, including employer-based, small-group, and individual-coverage plans. Out-of-network providers and insurers will have 30 days to agree on payment and then may invoke a binding arbitration process, in which each party submits a final offer and an arbitrator chooses between the two (see diagram). The arbitrator is instructed to consider the median in-network rate for the service, previous con-



Arbitration Process under the No Surprises Act.

tracted rates between the parties, and specific information about the patient's disease and the clinician's experience but not provider charges or Medicare rates.

The law also advances billing and payment transparency. Three days before scheduled procedures, clinicians and insurers must inform patients of their expected out-of-pocket costs and clinicians' network status. Only after receiving this information and information on in-network alternatives and consenting to out-of-network bills can patients be balance-billed. This notice-and-consent exception doesn't apply to emergency services, urgent or unanticipated care, situations in which there are no in-network alternatives, or "ancillary" services, such as anesthesiology, radiology, pathology, or neonatology. In other words, patients cannot be balance billed in these cases or for these services, even if they provide consent.

Recent evidence may help predict the new law's effects. A similar arbitration process has been in place for several years in New York and New Jersey. One key difference is that arbitrators in these states are instructed to consider the 80th percentile of provider charges for a given service, which is typically many times higher than the median in-network rate. For example, the median in-network rate for a comprehensive emergency department evaluation in New York is $320, whereas the 80th percentile of charges is $1,211.[2] Clinicians won the majority of decisions in 2018 in both New York and New Jersey, with awards gravitating toward the 80th percentile of charges.[2,3] Because providers can receive generous arbitration awards by staying out of network, they have the upper hand in negotiating in-network rates with payers, who may prefer to pay high in-network rates over going to arbitration. This dynamic may inflate prices in the long run.

California's surprise-billing ban, by contrast, established a benchmark for out-of-network reimbursements, set at the higher of the payer's local average in-network rate or 125% of the Medicare rate, coupled with an optional arbitration process that has been used infrequently. After this law was enacted in 2017, the share of out-of-network claims in affected specialties decreased from 21.5% to 17.8%.[4] Benchmarking reduces the incentive for physicians to be out of network, since reimbursement for out-of-network services is pegged to average in-network rates. It may also reduce long-term spending, because it doesn't allow physicians to seek higher reimbursements using an arbitration approach anchored at a higher rate. This approach may reduce the negotiating leverage of physicians in hospital-based specialties linked to surprise billing. The decrease in out-of-network services suggests that the policy hasn't substantially disrupted California's provider networks, though questions remain about its effect on physician reimbursement.[4]

The No Surprises Act blends these approaches and may prevent unfair practices on both sides. Unlike in New York and New Jersey, arbitrators will be prohibited from considering charges and will instead refer to median in-network rates for services. This approach may help avoid the inflationary effects seen in these states. On the other hand, unlike California's policy, the legislation doesn't set a benchmark price and requires arbitrators to consider case-specific nuances, such as the clinician's expertise and both parties' history of good-faith negotiation — which may prevent insurers from unfairly dropping clinicians from their networks. The law will probably reduce reimbursements for providers who use surprise billing as a business tactic, such as large physician-staffing firms in emergency medicine and anesthesia. The Congressional Budget Office estimates that the law will reduce payments for some clinicians, reduce insurance premiums by up to 1%, and save the federal government nearly $17 billion over 10 years.

The law's transparency provisions — particularly the requirements to provide advance price and network-participation information — may have a larger effect on day-to-day practice than its balance-billing provisions. Providing an advance explanation of benefits for scheduled procedures requires providers to anticipate all clinicians involved in the procedure and submit their identifiers and billing codes to insurance plans, and requires insurers to cross-reference this information against provider directories and records of patients' deductibles and out-of-pocket maximums. Although standard in fields such as dentistry (and certainly worth pursuing), this process would represent a seismic change for clinicians and insurers — particularly for underresourced practices and hospitals.

An audio interview with Dr. Ryan is available at NEJM.org

Before passage of the No Surprises Act, most states had laws protecting patients from surprise bills, although they have historically applied to only fully insured health plans, which cover a minority of commercially insured people.[5] The new law defers to states' various approaches for determining out-of-network rates, including binding arbitration, nonbinding arbitration, benchmarks, or other methods. The benefits of state deference are that states can test approaches and can pass more protective standards if federal ones prove inadequate. The downsides are that state deference permits state laws that may err too far in favor of clinicians or insurers, could allow states to undermine federal protections, and leaves providers, arbitrators, regulators, and patients with a confusing patchwork of standards.

Although the new legislation is fairly comprehensive, more work on surprise billing remains. The law's omission of ground-ambulance surprise bills is an important weakness, and the ground-ambulance advisory committee it created may not be up to the delicate task of designing a policy that could upset local governments — some of which rely on balance billing to sustain their ambulance corps. Researchers and policymakers will need to evaluate the law's effects on network participation (since it may induce low-paid providers to go out of network or insurers to drop high-priced providers), in-network prices, physician supply, and overall health care spending.

Despite some flaws, the No Surprises Act is a major victory for the public. Like any compromise, it is imperfect and will require close scrutiny as it unfolds. Yet in a time of tremendous economic uncertainty, it represents an important step toward reducing financial harm to patients and restoring trust in the health care system.

Disclosure forms provided by the authors are available at NEJM.org.

From the Department of Surgery, Brigham and Women's Hospital, Boston (K.R.C.), the College of Law, Georgia State University, Atlanta (E.F.B.), and the Center for Evaluating Health Reform and the School of Public Health, University of Michigan, Ann Arbor (A.M.R.).

This article was published on March 17, 2021, at NEJM.org.

1. Chhabra KR, McGuire K, Sheetz KH, Scott JW, Nuliyalu U, Ryan AM. Most patients undergoing ground and air ambulance transportation receive sizable out-of-network bills. Health Aff (Millwood) 2020; 39:777-82.
2. Adler L. Experience with New York's arbitration process for surprise out-of-network bills. Washington, DC: Brookings Institution. October 24, 2019 (https://www.brookings.edu/blog/usc-brookings-schaeffer-on-health-policy/2019/10/24/experience-with-new-yorks-arbitration-process-for-surprise-out-of-network-bills/).
3. Chartock BL, Adler L, Ly B, Duffy E, Trish E. Arbitration over out-of-network medical bills: evidence from New Jersey payment disputes. Health Aff (Millwood) 2021; 40:130-7.
4. Adler L, Duffy E, Ly B, Trish E. California saw reduction in out-of-network care from affected specialties after 2017 surprise billing law. Washington, DC: the Brookings Institution. September 26, 2019 (https://www.brookings.edu/blog/usc-brookings-schaeffer-on-health-policy/2019/09/26/california-saw-reduction-in-out-of-network-care-from-affected-specialties-after-2017-surprise-billing-law/).
5. The Commonwealth Fund. State balance-billing protections. November 30, 2020 (https://www.commonwealthfund.org/publications/maps-and-interactives/2020/jul/state-balance-billing-protections).

DOI: 10.1056/NEJMp2035905

Remote Patient Monitoring — Overdue or Overused?

Keizra Mecklai, B.S., Nicholas Smith, B.S., Ariel D. Stern, Ph.D., and Daniel B. Kramer, M.D., M.P.H.

The Covid-19 pandemic has challenged clinicians to find innovative ways to provide essential services while minimizing risks for themselves and their patients. These approaches increasingly leverage remote patient monitoring (RPM), using technology to support treatment for chronic conditions. As the use of RPM services grows, clinicians, payers, and patients face important questions regarding the volume, value, and appropriate use of this care model.

RPM has long been integrated into focused areas of disease management, such as care of patients with pacemakers or implantable cardioverter–defibrillators. RPM for these patients can reduce costs and supplement or replace in-office care, while offering convenience and heightened surveillance for clinical events. In recent years, RPM technology has expanded into new areas, including chronic and acute care management for multiple common conditions. Devices used in patients' homes now capture physiological parameters such as weight, blood pressure, oxygen saturation, and blood glucose levels and transmit these data to clinicians for review. For example, wrist-worn pulse oximeters transmitting oxygen-saturation data may be used to monitor lung function in patients with chronic obstructive pulmonary disease, and continuous glucose monitors may wirelessly transmit to physicians information about blood-sugar control in diabetic patients at different times of day and between office visits.

In 2019, the Centers for Medicare and Medicaid Services (CMS) issued a final rule on changes to the Medicare Part B Physician Fee Schedule establishing three new billing codes for Chronic Care RPM.[1] These codes allowed reimbursement for initial setup of RPM devices and associated patient education; collection and interpretation of physiological data; and RPM treatment management services. A 2020 update expanded coverage for RPM services and created an add-on code for reimbursement for an additional 20 minutes of RPM services per patient per month, raising the maximum to 40 minutes per month.[2]

In response to Covid-19 and associated legislation,[3] CMS expanded RPM coverage further, specifying that it is not limited to patients with chronic conditions but also includes those with acute conditions such as Covid-19. The interim rule also established that for the duration of the national emergency, consent for RPM services can be obtained just once a year for both new and established patients. Providers are also permitted to waive copayments for services rendered outside an "in-person face-to-face" encounter, including telehealth and RPM. This confluence of technological advancement and assurance of reimbursement in a fee-for-service environment — particularly as health care providers lose revenue because of the pandemic — may lead to dramatic increases in RPM utilization and expenditures.

RPM has the potential to enhance management of acute and chronic conditions and to help personalize treatment plans with the use of high-frequency health data. It is possible, although not yet demonstrated at scale, that evidence-based RPM can improve clinical outcomes for individual patients while, at the health systems level, reducing downstream health care costs, such as those associated with preventable hospital admissions. There are, however, several reasons to worry about a short-term explosion in RPM expenditures.

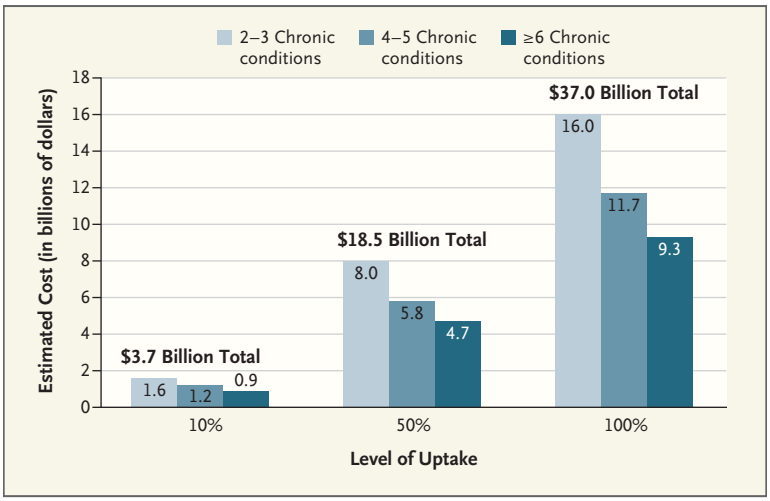
First, makers of RPM tools can currently pursue marketing approval (if needed) and subsequent reimbursement coverage under standards that do not require demonstration of clinical effectiveness in overall disease management. A pulse oximetry system for patients with chronic lung disease, for example, may have to meet certain engineering and manufacturing standards but does not need to be shown to improve patient outcomes to be legally marketed. For these devices in general, the Food and Drug Administration (FDA) places the burden on health care providers to "develop appropriate processes and procedures to assess and manage risks associated with the integration of [radiofrequency] wireless technology into medical systems."[4] In the FDA's risk-based classification of devices, most RPM devices will not be considered high-risk, so meeting the statutory standard of reasonable assurance of safety and effectiveness generally will not require clinical trials, nor will the software for many commercial wear-

ables, which is expected to be regulated through the FDA's Digital Health Software Precertification Program in the future.

Second, to date CMS has offered few stipulations on what specifications or standards must be met for coverage of an RPM device. Even well-studied devices in common diseases, such as hypertension, heart failure, and atrial fibrillation, have shown highly variable benefits of different products and care pathways.[5] Randomized, controlled trials have revealed variable effects on outcome measures such as hospital readmission, cardiovascular mortality, or all-cause mortality. Eventually, high-quality, prospective studies either designed as clinical trials or leveraging real-world data will be necessary to support the clinical case for RPM systems.

Third, even without high-quality clinical data, expansion of fee-for-service reimbursement for RPM services provides incentives for rapid uptake. With more and more devices available, providers may enroll large numbers of patients in RPM programs with little regard for who will see a meaningful benefit. Alternative payment models such as bundled payments may shift these incentives, but fee-for-service reimbursement remains a dominant feature of U.S. health care. The costs of RPM expansion may also be borne partially by patients. RPM could increase out-of-pocket expenditures, depending on coinsurance and access to devices, since one established RPM Current Procedural Terminology (CPT) code allows providers to bill for up to 30 minutes per patient per month without having to communicate with the patient or caregiver.

With few data to guide forecasting, whether RPM services and associated expenditures will grow rapidly remains to be seen. However, we estimated the potential impact of RPM services on Medicare expenditures with a simple model integrating the number of beneficiaries, chronic conditions per beneficiary, utilization of RPM, and reimbursement per RPM service (see the Supplementary Appendix, available at NEJM.org). A conservative estimate would assume that RPM enrollment would be limited to patients with multiple chronic conditions, yet dissemination in that subpopulation alone could translate into annual expenditures exceeding $18 billion, even with just 50% uptake (see graph).



Estimated Cost to Medicare of Remote Patient Monitoring (RPM) Services.

Beneficiaries are categorized according to the number of documented chronic conditions. The bars indicate the estimated contribution of each category to overall expenditures at a given level of uptake of RPM services.

This estimate is based on the assumption that 68% of Medicare fee-for-service beneficiaries — about 25.4 million patients as of September 2020, according to CMS — have two or more chronic conditions. The maximum annual cost per patient enrolled in an RPM program is $1,460, according to the 2020 CMS Fee Schedule. This cost comprises monthly fees for device supply and data transmission ($62.44, CPT code 99454) and for collection and interpretation of physiological data ($59.19, CPT code 99091). It is unrealistic to believe that 100% of eligible patients will enroll, but even with a 10% enrollment rate, the annual cost to Medicare could reach $3.7 billion — just under 1% of total 2018 Medicare Part A and B expenditures (see Supplementary Appendix). Additional costs might be accrued as Medicare Advantage and other private payers adopt similar coverage and reimbursement.

Research is urgently needed to elucidate which patients benefit most from RPM services and which devices and specifications provide the highest clinical value. This information will enable professional societies to publish evidence-based guidelines on who should enroll in RPM programs and which devices and support systems should be deployed to maximize the clinical impact of RPM and the collection of health data. Such studies would also provide evidence to enable CMS

to set standards for RPM devices to qualify for coverage. Furthermore, private-sector efforts to create transparency regarding the usability, validation, and data-security profiles of biosensors will support clinicians and researchers in technology-adoption decisions.

The CMS rule changes and the pandemic have resulted in rapid and sweeping expansion of reimbursement for telehealth and RPM technologies and services without evidence-based coverage decisions. Given social-distancing recommendations and the desire to enhance patient safety, RPM provides promising solutions for accessible, data-driven care while reducing exposure risks. As RPM tools evolve, we may have opportunities to learn from other countries; for example, Germany's 2019 Digital Healthcare Act, which provides for insurance coverage of certain digital health applications, includes provisions for evidence generation as a requirement for ongoing reimbursement. Rigorous, ongoing evaluation of RPM devices and platforms will be essential for elucidating their value and driving coverage decisions and adoption programs for the most effective solutions.

Disclosure forms provided by the authors are available at NEJM.org.

From Harvard Medical School (K.M., N.S., D.B.K.), Harvard Business School (K.M., N.S., A.D.S.), the Harvard–MIT Center for Regulatory Science (A.D.S.), and the Richard A. and Susan F. Smith Center for Outcomes Research in Cardiology, Beth Israel Deaconess Medical Center (D.B.K.) — all in Boston; and the Health Innovation Hub, German Federal Ministry of Health, Berlin (A.D.S.).

Ms. Mecklai and Mr. Smith contributed equally to this article.

This article was published on April 10, 2021, at NEJM.org.

1. Medicare program; revisions to payment policies under the Physician Fee Schedule and other revisions to part B for CY 2019; Medicare Shared Savings Program requirements; Quality Payment Program; Medicaid Promoting Interoperability Program; Quality Payment Program — extreme and uncontrollable circumstance policy for the 2019 MIPS payment year; provisions from the Medicare Shared Savings Program — accountable care organizations — pathways to success; and expanding the use of telehealth services for the treatment of opioid use disorder under the Substance Use-Disorder Prevention that Promotes Opioid Recovery and Treatment (SUPPORT) for Patients and Communities Act. Federal Register. November 23, 2018 (https://www.federalregister.gov/documents/2018/11/23/2018-24170/medicare-program-revisions-to-payment-policies-under-the-physician-fee-schedule-and-other-revisions#h-81).
2. H.R. 6074, Coronavirus Preparedness and Response Supplemental Appropriations Act, 116th Congress (2019-2020) (https://www.congress.gov/bill/116th-congress/house-bill/6074).
3. Centers for Medicare & Medicaid Services. Medicare and Medicaid Programs: policy and regulatory revisions in response to the COVID-19 public health emergency. 2020 (https://www.cms.gov/files/document/covid-final-ifc.pdf).
4. Food and Drug Administration. Wireless medical devices (https://www.fda.gov/medical-devices/digital-health-center-excellence/wireless-medical-devices).
5. American Heart Association. Using remote patient monitoring technologies for better cardiovascular disease outcomes: guidance, 2019 (https://www.heart.org/-/media/files/about-us/policy-research/policy-positions/clinical-care/remote-patient-monitoring-guidance-2019.pdf).

DOI: 10.1056/NEJMp2033275

HISTORY OF MEDICINE

Race, Policing, and History — Remembering the Freedom House Ambulance Service

Matthew L. Edwards, M.D.

Americans protesting violent policing of Black communities are calling for law-enforcement budgets to be reallocated to community health services. Although such proposals are sometimes dismissed as naive or unrealistic, history provides an example of a transfer of power and resources from police to health services that benefited Black communities enormously. Pittsburgh's Freedom House Enterprises (FHE) Ambulance Service not only supplanted the police in a role in which law-enforcement officers were not effective, but also reimagined the role of Black citizens in improving the community's health and helped establish national standards for emergency medical care.

Freedom House was a community-based sociomedical program that aspired to "encourage Black enterprise" during the 1960s and 1970s by training Black community members to provide emergency medical services (EMS) (see photo).[1] At the time, police officers and morticians without medical training supplied most prehospital "care," generally providing transportation without medical treatment. Even Pennsylvania Governor David Lawrence's 1966

to set standards for RPM devices to qualify for coverage. Furthermore, private-sector efforts to create transparency regarding the usability, validation, and data-security profiles of biosensors will support clinicians and researchers in technology-adoption decisions.

The CMS rule changes and the pandemic have resulted in rapid and sweeping expansion of reimbursement for telehealth and RPM technologies and services without evidence-based coverage decisions. Given social-distancing recommendations and the desire to enhance patient safety, RPM provides promising solutions for accessible, data-driven care while reducing exposure risks. As RPM tools evolve, we may have opportunities to learn from other countries; for example, Germany's 2019 Digital Healthcare Act, which provides for insurance coverage of certain digital health applications, includes provisions for evidence generation as a requirement for ongoing reimbursement. Rigorous, ongoing evaluation of RPM devices and platforms will be essential for elucidating their value and driving coverage decisions and adoption programs for the most effective solutions.

Disclosure forms provided by the authors are available at NEJM.org.

From Harvard Medical School (K.M., N.S., D.B.K.), Harvard Business School (K.M., N.S., A.D.S.), the Harvard–MIT Center for Regulatory Science (A.D.S.), and the Richard A. and Susan F. Smith Center for Outcomes Research in Cardiology, Beth Israel Deaconess Medical Center (D.B.K.) — all in Boston; and the Health Innovation Hub, German Federal Ministry of Health, Berlin (A.D.S.).

Ms. Mecklai and Mr. Smith contributed equally to this article.

This article was published on April 10, 2021, at NEJM.org.

1. Medicare program; revisions to payment policies under the Physician Fee Schedule and other revisions to part B for CY 2019; Medicare Shared Savings Program requirements; Quality Payment Program; Medicaid Promoting Interoperability Program; Quality Payment Program — extreme and uncontrollable circumstance policy for the 2019 MIPS payment year; provisions from the Medicare Shared Savings Program — accountable care organizations — pathways to success; and expanding the use of telehealth services for the treatment of opioid use disorder under the Substance Use-Disorder Prevention that Promotes Opioid Recovery and Treatment (SUPPORT) for Patients and Communities Act. Federal Register. November 23, 2018 (https://www.federalregister.gov/documents/2018/11/23/2018-24170/medicare-program-revisions-to-payment-policies-under-the-physician-fee-schedule-and-other-revisions#h-81).
2. H.R. 6074, Coronavirus Preparedness and Response Supplemental Appropriations Act, 116th Congress (2019-2020) (https://www.congress.gov/bill/116th-congress/house-bill/6074).
3. Centers for Medicare & Medicaid Services. Medicare and Medicaid Programs: policy and regulatory revisions in response to the COVID-19 public health emergency. 2020 (https://www.cms.gov/files/document/covid-final-ifc.pdf).
4. Food and Drug Administration. Wireless medical devices (https://www.fda.gov/medical-devices/digital-health-center-excellence/wireless-medical-devices).
5. American Heart Association. Using remote patient monitoring technologies for better cardiovascular disease outcomes: guidance, 2019 (https://www.heart.org/-/media/files/about-us/policy-research/policy-positions/clinical-care/remote-patient-monitoring-guidance-2019.pdf).

DOI: 10.1056/NEJMp2033275

HISTORY OF MEDICINE

Race, Policing, and History — Remembering the Freedom House Ambulance Service

Matthew L. Edwards, M.D.

Americans protesting violent policing of Black communities are calling for law-enforcement budgets to be reallocated to community health services. Although such proposals are sometimes dismissed as naive or unrealistic, history provides an example of a transfer of power and resources from police to health services that benefited Black communities enormously. Pittsburgh's Freedom House Enterprises (FHE) Ambulance Service not only supplanted the police in a role in which law-enforcement officers were not effective, but also reimagined the role of Black citizens in improving the community's health and helped establish national standards for emergency medical care.

Freedom House was a community-based sociomedical program that aspired to "encourage Black enterprise" during the 1960s and 1970s by training Black community members to provide emergency medical services (EMS) (see photo).[1] At the time, police officers and morticians without medical training supplied most prehospital "care," generally providing transportation without medical treatment. Even Pennsylvania Governor David Lawrence's 1966

death, which was partially attributable to inadequate EMS care, failed to galvanize improvements in emergency care.[2] Moreover, EMS quality was often worse in Black communities. In this bleak environment, Freedom House enabled a group of disadvantaged Black laypeople to establish a model for paramedic training that ultimately set the U.S. standard.

The problems affecting Pittsburgh residents in the 1960s were similar to those we face today. The National Advisory Commission on Civil Disorders argued that the urban riots of the 1960s were a response to structural racism and socioeconomic inequity.[3] Cities throughout the United States were rife with racist systems that precluded equal access to education, housing, employment, political opportunities, and social services.[2] Black citizens were subject to unfair treatment by the carceral system and inadequate access to medical care.[3] President Lyndon Johnson's War on Poverty program sought to remedy these inequities by expanding civil rights and promoting public welfare, education, urban development, and public health programs.[3] The Opportunities Industrialization Center and the War on Poverty initiative's new Office of Economic Opportunity (OEO) increased employment opportunities by implementing job-training programs such as the Freedom House Ambulance Service.[3]

After World War II, municipal laws authorized police departments to provide emergency medical services.[2] Many Black Americans relied on the police for EMS because they could not afford private hospital transport and because White operators of such services often avoided Black communities. Though the Emergency Medical Treatment and Labor Act of 1986 would eventually guarantee the right to emergency response and treatment regardless of one's background or ability to pay, such provisions did not exist in the 1960s.[3] Moreover, the minimum training standards for emergency responders fell short of evolving treatment standards.[3]

Freedom House Paramedics with Ambulance.

Photo credit: University of Pittsburgh

Then, as now, Black citizens faced discrimination and abuse by police and disproportionate rates of arrest and incarceration. Activists' efforts to secure redress were rebuffed, as police leadership cited difficulties obtaining police-misconduct convictions.[2] Consequently, many Black people felt a sense of indignity and fear when forced to rely on police officers for transportation to the hospital.[2,3]

So a biracial group of Pittsburgh leaders approached physician Peter Safar for guidance on equipping ambulance vehicles to transport Black patients to and from the hospital. In Baltimore during the 1950s, Safar and James Elam had not only proved the superiority of mouth-to-mouth ventilation in resuscitation, but demonstrated that laypeople could learn principles of artificial respiration. Safar and Elam went on to develop public-awareness videos and tools such as the "Resusci Anne" doll to teach and simulate artificial respiration. At the University of Pittsburgh Medical Center, Safar had achieved renown for developing cardiopulmonary resuscitation principles, creating the first multidisciplinary critical care unit in the United States, and designing mobile intensive care units.

Safar agreed to provide consultation on emergency vehicles in exchange for the chance to train Black community members to provide prehospital transport. The OEO helped recruit a Black workforce to undergo training, in a local approach to addressing racial and health inequities.

Fire and police departments vigorously opposed the ambulance service, which they saw as a threat to their autonomy.[3] But even in this hostile environment, which persisted from the start of the program in 1967 until its demise in 1975, Freedom House proved largely successful. One study comparing its services with police services found that Freedom House paramedics provided improper treatment in only 11% of cases, as compared with 62% by the police.[2]

In 1973, Safar recruited physician Nancy Caroline to lead the Freedom House training program.[3,4] Over her 2 years as medical director of what she called an "audacious, improbable experiment,"[2] she transformed its leadership and was a steady presence in trainees' lives. The program used classroom, hospital-based, and field training to teach basic anatomy, physiology, disease recognition and diagnosis, and common emergency conditions. Caroline connected with the paramedics personally while delivering rigorous training by regularly participating in ambulance rides and providing clinical oversight. She understood the importance of a sound foundation in critical care medicine, but as a Jewish female physician, she also recognized the importance of offering a sense of dignity and belonging to marginalized Black Americans.

Freedom House became the pilot course for EMS training for the U.S. Department of Transportation and the Federal Interagency Committee on Emergency Medical Services.[3,4] Freedom House paramedics and the surrounding communities were proud of their accomplishments. Previously deemed "unemployable," many trainees pursued advanced degrees, municipal leadership and state-level administrative roles, and advanced training in education, medicine, and allied health fields.[1] Although some were able to find employment in municipal services, state-level leadership, and public health administration,[1,5] others found themselves cast out and once again unemployed when the service was disbanded.[1]

White Freedom House employees had a different experience. As Caroline wrote, "for eight years, [Black Freedom House trainees] had stuck with the organization while they watched white trainees leave FHE to assume high administrative positions with City and County EMS agencies. . . . They [White trainees] had all done their apprenticeship with FHE, and now they were in control and Freedom House was odd man out."[3]

Although Safar viewed work toward racial equity, economic opportunity, and health care access as complementary to his vision of national EMS standards, political opposition to his vision intensified in Pittsburgh.[3] FHE initially received local, state, and federal funding, but Pittsburgh's administration proposed to reduce funding for social welfare programs over time and cited Freedom House's cost as a deterrent to further support. The city government's subsequent creation of a more expensive and predominantly White citywide "superambulance" service, however, suggested that money was not the issue. Freedom House began preparing to close, and its board and city officials voted to dissolve it on September 22, 1975.[3]

The superambulance service employed predominantly White workers, jettisoning the social goals of Freedom House by excluding the Black men and women who had pioneered EMS standards. Seeing prehospital care as the "weakest link" in a continuum of critical care medicine,[3,5] Safar focused primarily on nationalizing EMS standards — improving systems and protocols for emergency medical care through national organizations such as the National Research Council. In prioritizing standardization over community development and protection of oppressed groups, Safar's choice set the tone for subsequent national structuring of police and emergency response systems, which relegated these functions to mostly White male professionals from outside the communities they serve.

The history of Freedom House is inspiring as an example of how physicians and a multiracial group of citizens can recognize a need for resources in Black communities and coordinate a progressive community-development program that raises standards of excellence for all. But it's also a cautionary tale, demonstrating how well-intentioned leaders with their own agendas can undermine the goals of Black community empowerment while co-opting the products of Black innovation. History reveals various models of community partnerships, from the top-down political sponsorship of Safar to the bottom-up coalition work of Caroline. The Freedom House story suggests that though community health initiatives offer an alternative to systems like policing, the same racial and power dynamics can

affect both. Whether recent calls to restructure community investment grow out of protests against police brutality or the need to rethink community health services, they raise critical questions about why we invest so heavily in racist systems that see communities of color as in need of help, but not equipped to also offer it, rather than supporting programs that protect and involve these citizens as agents.

Disclosure forms provided by the author are available at NEJM.org.

From the Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA.

This article was published on April 10, 2021, at NEJM.org.

1. Caroline NL. Emergency! 'Freedom House' saved lives — yours and theirs, but now it is mostly shunted aside and forgotten. Pittsburgh Magazine. Vol. 8. 1977:42-6, 84-6.
2. Trotter JW, Day JN. Race and renaissance: African Americans in Pittsburgh since World War II. Pittsburgh: University of Pittsburgh Press, 2013.
3. Edwards ML. Pittsburgh's Freedom House Ambulance Service: the origins of emergency medical services and the politics of race and health. J Hist Med Allied Sci 2019;74:440-66.
4. Safar PJ. On the history of emergency medical services. Bull Anesth Hist 2001;19:1-11.
5. Freedom House ambulance service. 99% Invisible. Episode no. 405. 2020 (podcast) (https://99percentinvisible.org/episode/freedom-house-ambulance-service/).

DOI: 10.1056/NEJMp2035467

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 APRIL 15, 2021 VOL. 384 NO. 15

Hypothermic Machine Perfusion in Liver Transplantation — A Randomized Trial

R. van Rijn, I.J. Schurink, Y. de Vries, A.P. van den Berg, M. Cortes Cerisuelo, S. Darwish Murad, J.I. Erdmann, N. Gilbo, R.J. de Haas, N. Heaton, B. van Hoek, V.A.L. Huurman, I. Jochmans, O.B. van Leeuwen, V.E. de Meijer, D. Monbaliu, W.G. Polak, J.J.G. Slangen, R.I. Troisi, A. Vanlander, J. de Jonge, and R.J. Porte, for the DHOPE-DCD Trial Investigators*

ABSTRACT

BACKGROUND

Transplantation of livers obtained from donors after circulatory death is associated with an increased risk of nonanastomotic biliary strictures. Hypothermic oxygenated machine perfusion of livers may reduce the incidence of biliary complications, but data from prospective, controlled studies are limited.

METHODS

In this multicenter, controlled trial, we randomly assigned patients who were undergoing transplantation of a liver obtained from a donor after circulatory death to receive that liver either after hypothermic oxygenated machine perfusion (machine-perfusion group) or after conventional static cold storage alone (control group). The primary end point was the incidence of nonanastomotic biliary strictures within 6 months after transplantation. Secondary end points included other graft-related and general complications.

RESULTS

A total of 160 patients were enrolled, of whom 78 received a machine-perfused liver and 78 received a liver after static cold storage only (4 patients did not receive a liver in this trial). Nonanastomotic biliary strictures occurred in 6% of the patients in the machine-perfusion group and in 18% of those in the control group (risk ratio, 0.36; 95% confidence interval [CI], 0.14 to 0.94; P=0.03). Postreperfusion syndrome occurred in 12% of the recipients of a machine-perfused liver and in 27% of those in the control group (risk ratio, 0.43; 95% CI, 0.20 to 0.91). Early allograft dysfunction occurred in 26% of the machine-perfused livers, as compared with 40% of control livers (risk ratio, 0.61; 95% CI, 0.39 to 0.96). The cumulative number of treatments for nonanastomotic biliary strictures was lower by a factor of almost 4 after machine perfusion, as compared with control. The incidence of adverse events was similar in the two groups.

CONCLUSIONS

Hypothermic oxygenated machine perfusion led to a lower risk of nonanastomotic biliary strictures following the transplantation of livers obtained from donors after circulatory death than conventional static cold storage. (Funded by Fonds NutsOhra; DHOPE-DCD ClinicalTrials.gov number, NCT02584283.)

The authors' full names, academic degrees, and affiliations are listed in the Appendix. Address reprint requests to Dr. Porte at the Department of Surgery, Section of Hepatobiliary Surgery and Liver Transplantation, University Medical Center Groningen, P.O. Box 30.001, 9700RB Groningen, the Netherlands, or at r.j.porte@umcg.nl.

*A complete list of collaborators and their roles in the DHOPE-DCD trial is provided in the Supplementary Appendix, available at NEJM.org.

This article was published on February 24, 2021, at NEJM.org.

N Engl J Med 2021;384:1391-401.
DOI: 10.1056/NEJMoa2031532

A Quick Take is available at NEJM.org

Nonanastomotic biliary strictures are a major complication after liver transplantation, resulting in cholestasis and cholangitis and, frequently, in the use of biliary interventions or even retransplantation (i.e., transplantation of a new liver graft and removal of the first graft).[1,2] The incidence of nonanastomotic biliary strictures is approximately 3 times as high after the transplantation of livers obtained from donors after circulatory death as after the transplantation of livers obtained from brain-dead donors.[2,3] Nevertheless, liver grafts from donors after circulatory death are increasingly used for transplantation owing to persistent donor-organ shortage.[4,5]

Ischemia–reperfusion injury is a key mechanism in the pathogenesis of bile-duct injury and the subsequent development of biliary strictures after transplantation.[1,3] Although conventional static cold preservation provides some protection against ischemia–reperfusion injury, more-advanced preservation methods are needed to improve outcomes after transplantation of livers obtained from donors after circulatory death and to increase the frequency of their use.[4]

Oxygenated ex situ machine perfusion is a dynamic preservation method that has been developed to reduce the incidence and severity of ischemia–reperfusion injury and to improve outcomes after organ transplantation.[6-9] Preclinical studies have shown that a short period (1 to 2 hours) of hypothermic oxygenated machine perfusion restores mitochondrial function and reduces the production of radical oxygen species and damage-associated molecular patterns after transplantation.[10-12] This relatively simple technique can be performed after static cold storage.[6,8] The first clinical experiences suggested that this preservation method was safe, reduced the incidence of hepatobiliary preservation injury, and was associated with improved early graft function, as compared with static cold preservation alone.[13-16] Although these findings were promising and have increased the interest in machine-based preservation techniques, they were based on small single-center cohorts without a randomized control group. We conducted a multicenter, randomized, controlled trial to compare hypothermic oxygenated machine perfusion with static cold preservation in the transplantation of livers from donors after circulatory death, with the incidence of nonanastomotic biliary strictures as the primary end point.

METHODS

TRIAL DESIGN AND OVERSIGHT

The DHOPE-DCD (Dual Hypothermic Oxygenated Perfusion of DCD Liver Grafts in Preventing Nonanastomotic Biliary Strictures after Transplantation) trial was investigator-initiated and was designed as a multicenter, prospective, two-group, randomized, controlled, clinical trial. The trial was conducted in six liver-transplantation centers in Europe. Centralized balanced-block randomization (in blocks of six) was computer-generated, with stratification according to trial center and primary sclerosing cholangitis as an indication for transplantation (yes or no). Patients were randomly assigned in a 1:1 ratio to receive a liver preserved either with hypothermic oxygenated machine perfusion after static cold preservation during transportation (machine-perfusion group) or with static cold preservation alone (control group). Randomization took place immediately after a donor liver had been deemed to be suitable and had been accepted by the transplantation surgeon for a recipient. The trial did not interfere with the regular process of organ allocation or acceptance.

The trial protocol, which is available with the full text of this article at NEJM.org, has been published previously.[17] The protocol was approved by research ethics committees at each trial site and medical-device regulatory bodies in each country. Patients and the organ-procurement teams were unaware of the trial-group assignments. The authors designed and implemented the trial and collected and analyzed the data. The first and last authors wrote the first draft of the manuscript, and all the authors contributed to the subsequent versions. All the authors vouch for the accuracy and completeness of the data and for the fidelity of the trial to the protocol. Access to the data was not restricted by confidentiality agreements.

Fonds NutsOhra supported this trial. Bridge to Life provided the machine-perfusion fluid (Belzer MPS UW machine-perfusion solution) free of charge. Each participating center covered the costs for the purchase of a machine-perfusion device, and training of perfusionists was pro-

vided by the manufacturer (Organ Assist) as part of their regular after-sales responsibilities. The funding organization and the companies had no role in the trial design; the collection, management, analysis, or interpretation of the data; or the writing of the manuscript and the decision to submit it for publication.

TRIAL PATIENTS

Patients 18 years of age or older who were undergoing liver-only transplantation with a graft from a donor after circulatory death (in controlled circumstances) were eligible for inclusion in the trial. Patients were excluded if the body weight of the donor was less than 40 kg or if the donor was positive for the human immunodeficiency virus or hepatitis B or C virus. Patients were also excluded if they were undergoing transplantation for fulminant liver failure or for primary nonfunction after a previous transplantation, were incapable of providing informed consent, were positive for the human immunodeficiency virus, or had a contraindication to undergoing magnetic resonance cholangiography. All the patients provided written informed consent.

Donor livers were obtained, preserved, and transported to the transplantation centers according to standard practice, with the use of conventional static cold preservation. The transplantation surgery and postoperative care were performed according to standard local practice.

HYPOTHERMIC OXYGENATED MACHINE PERFUSION

The Liver Assist device (Organ Assist) was used for ex situ machine perfusion of the liver. The device enables pressure-controlled, dual perfusion through the portal vein and the hepatic artery with the use of two centrifugal pumps, providing continuous portal flow and a pulsatile arterial flow at 60 beats per minute. The perfusion device was primed with 4 liters of cold Belzer machine-perfusion solution (Bridge to Life), supplemented with 3 mmol of glutathione per liter of solution (Biomedica). The perfusion pressure was 25 mm Hg for the hepatic artery and 5 mm Hg for the portal vein. The temperature of the perfusion fluid was 10°C. Oxygenation was provided by 500 ml per minute of 100% oxygen flow to each oxygenator.[15] The minimum protocol-stipulated duration of machine perfusion was 2 hours, a duration that is considered to be sufficient to restore mitochondria and intrahepatic ATP and to protect organs against ischemia–reperfusion injury.[10,11,15] Additional details are provided in the Supplementary Appendix, available at NEJM.org.

END-POINT MEASURES

The primary end point was the incidence of symptomatic nonanastomotic biliary strictures at 6 months after transplantation. The occurrence of nonanastomotic biliary strictures was assessed primarily by the medical teams of the participating centers on the basis of the presence of the following prespecified criteria: any irregularity or narrowing of the lumen of the intrahepatic or extrahepatic donor bile ducts, excluding the biliary anastomosis, diagnosed with the use of cholangiography (preferably, magnetic resonance cholangiography), in combination with clinical symptoms (e.g., jaundice or cholangitis) or an elevation of cholestatic laboratory variables, in the presence of a patent hepatic artery. All clinical data, including data from cholangiographies, were submitted to the central data center for review. To avoid reporting bias, magnetic resonance cholangiography was performed after 6 months, in accordance with the study protocol, to detect radiologic evidence of cholangiopathy (nonanastomotic strictures) in patients who had not already received a diagnosis in the preceding time period. Nonanastomotic biliary strictures are typically detected 3 to 4 months after transplantation, and an observation period of 6 months was therefore considered to be appropriate for the detection of clinically meaningful events.[1,18] All the cholangiographies, both in patients who were symptomatic and in those who were asymptomatic, were reviewed by two independent radiologists who were unaware of the preservation method and clinical symptoms. In the case of discordant readings, a third radiologist was consulted for cases that could not be settled by consensus.

Secondary end points included intraoperative postreperfusion syndrome,[19,20] defined as a decrease of more than 30% in the mean systemic arterial blood pressure within 10 minutes after reperfusion, with or without a doubling of the norepinephrine dose; primary nonfunction, defined as liver failure, without an identifiable

cause, that necessitated retransplantation or led to death within 7 days after transplantation; early allograft dysfunction, assessed according to the Olthoff criteria[21]; and durations of stay in the intensive care unit and hospital. Other secondary end points included thrombosis of the hepatic artery or portal vein, anastomotic biliary strictures or leakage, and use of renal-replacement therapy within 6 months after transplantation. Serum markers of hepatobiliary injury and function were recorded daily during the first week and at 1 month, 3 months, and 6 months after transplantation. Patient and graft survival were recorded up to 1 year after the transplantation.

STATISTICAL ANALYSIS

The trial was powered to detect a clinically relevant difference in the incidence of symptomatic nonanastomotic biliary strictures between the two trial groups. On the basis of previous reports about the transplantation of livers obtained from donors after circulatory death, we presumed an incidence of 29% among livers that had been preserved by static cold storage, and we expected that the incidence with machine perfusion would be 11% (proportional reduction, 60%).[22-25] On the basis of a power of 80% and a 5% significance level (two-sided test) in two independent groups, we calculated that 77 livers would be needed in each trial group. We aimed to include 1 additional patient per trial group, resulting in 78 patients per group.

All end-point analyses were prespecified in the protocol and statistical analysis plan, which was finalized before the database was locked. The primary end point was analyzed with the use of a chi-square test, as well as in a log-binomial regression model with calculation of risk ratios. Prespecified covariates in this model were based on relevant literature and included stratification factors (trial site and primary sclerosing cholangitis) and donor-specific risk factors (donor risk index and donor warm-ischemia time, defined as the time period between circulatory arrest and in situ cold flush-out in the donor).[26,27] For consistency with the original protocol, we also analyzed the results using logistic-regression modeling and report them in the Supplementary Appendix. Time-to-event outcomes were analyzed with the use of Kaplan–Meier curves with a log-rank test and Cox proportional-hazards regression model with the calculation of hazard ratios. Secondary binary end points were assessed by means of a chi-square test or log-binomial regression to adjust for stratification factors. Continuous (log-transformed) outcomes were compared with the use of an independent Student's t-test. Missing data were assumed to be missing at random, and multiple imputation was performed when more than 10% of all the patients had missing data for a specific variable. There was no adjustment for multiplicity in analyses of secondary end points, and these analyses should be considered to be exploratory. Tests were two-sided, and results are reported with 95% confidence intervals. A P value of less than 0.05 was considered to indicate statistical significance. Analyses were performed with the use of SPSS software, version 23.0 (SPSS).

RESULTS

PATIENTS

From January 2016 through July 2019, we assessed a total of 245 patients for eligibility, of whom 160 underwent randomization. After randomization, four transplantations were canceled before any trial procedure was started. In one case, the liver was intended to undergo machine perfusion, and in three cases, the liver had been assigned to the control group. The reasons for cancelation were massive steatosis in two livers and a nonreconstructable damaged artery in another liver. These three livers had initially been deemed transplantable and had been accepted; they were secondarily rejected on the basis of this new information. In one patient, pseudomyxoma peritonei was detected after laparotomy; the transplantation was canceled and the liver was allocated to another patient outside the trial. This resulted in the inclusion of 78 patients in the machine-perfusion group and 78 patients in the control group (Fig. S1 in the Supplementary Appendix).

The baseline characteristics of the donors and recipients were well matched in the two trial groups (Table 1). Inherent to the intervention, the static cold-ischemia time was slightly shorter in the machine-perfusion group than in the control group (6 hours 11 minutes vs. 6 hours 49 minutes) and the total preservation time was longer (8 hours 44 minutes vs. 6 hours 49 minutes).

Table 1. Characteristics of the Donors and Recipients at Baseline.*

Characteristic	Machine Perfusion (N=78)	Control (N=78)
Donor characteristics		
Age — yr		
Median	52	49
Interquartile range	43–57	37–59
Male sex — no. (%)	52 (67)	51 (65)
Donor risk index†		
Median	2.12	2.12
Interquartile range	1.84–2.38	1.86–2.42
Body-mass index‡		
Median	25	25
Interquartile range	23–27	21–28
Preservation characteristics		
Time from withdrawal of life support to aortic flush-out — min		
Median	29	27
Interquartile range	22–33	21–35
Time from circulatory arrest in the donor to aortic flush-out — min		
Median	11	11
Interquartile range	8–13	8–15
Static cold-ischemia time§		
Median	6 hr 11 min	6 hr 49 min
Interquartile range	5 hr 16 min–6 hr 55 min	5 hr 56 min–7 hr 57 min
Machine-perfusion time		
Median	2 hr 12 min	NA
Interquartile range	2 hr 00 min–2 hr 33 min	NA
Total preservation time¶		
Median	8 hr 44 min	6 hr 49 min
Interquartile range	7 hr 46 min–9 hr 16 min	5 hr 56 min–7 hr 57 min
Recipient characteristics		
Age — yr		
Median	60	60
Interquartile range	52–65	52–65
Male sex — no. (%)	55 (71)	52 (67)
Laboratory MELD score‖		
Median	14	16
Interquartile range	10–19	10–22
Renal-replacement therapy — no. (%)	3 (4)	2 (3)

* Data on additional characteristics, including causes of death of the donor and indications for transplantation, are provided in Table S1. NA denotes not applicable.

† The donor risk index is a scoring system that was developed to quantitatively predict the risk of post-transplantation graft failure in liver transplantation, on the basis of donor risk factors.[27]

‡ The body-mass index is the weight in kilograms divided by the square of the height in meters.

§ The static cold-ischemia time was defined as time between aortic cold flush-out in the donor to reperfusion in the recipient, minus the machine perfusion time. P<0.001 for the comparison of the two groups.

¶ The total preservation time was defined as the time between aortic cold flush-out in the donor to reperfusion in the recipient. P<0.001 for the comparison of the two groups.

‖ The laboratory Model of End-Stage Liver Disease (MELD) score ranges from 6 to 40, with higher scores indicating more advanced disease. The laboratory MELD score is based on original laboratory variables; MELD exception points, which are used to assign increased priority on the waiting list to patients whose severity of illness or risk of complications is not captured by the laboratory MELD score, are not included in the scores shown here.

Table 2. Primary and Secondary End Points.*

Outcome	Machine Perfusion (N=78)	Control (N=78)	Treatment Effect (95% CI)	P Value
Primary end point†				
Nonanastomotic biliary strictures — no. (%)	5 (6)	14 (18)		0.03
Unadjusted risk ratio			0.36 (0.14 to 0.94)	0.03
Adjusted risk ratio			0.35 (0.14 to 0.92)	0.03
Secondary end points				
Postreperfusion syndrome				
>30% decrease in systemic mean arterial pressure — no./total no. (%)	9/72 (12)	19/70 (27)	0.43 (0.20 to 0.91)‡	
>30% decrease in systemic mean arterial pressure or ≥100% increase in norepinephrine dose — no./total no. (%)	20/72 (28)	33/72 (46)	0.59 (0.38 to 0.92)‡	
Serum potassium after reperfusion — mmol/liter§	4.1±0.7	4.4±1.1	−0.4 (−0.1 to −0.6)	
Graft-related complication — no. (%)				
Early allograft dysfunction¶	20 (26)	31 (40)	0.61 (0.39 to 0.96)	
Primary nonfunction	0	1 (1)	NA	
Hepatic-artery thrombosis	2 (3)	2 (3)	0.94 (0.12 to 7.19)‡	
Portal-vein thrombosis	0	2 (3)	NA	
Biliary anastomotic stricture	23 (29)	22 (28)	1.07 (0.52 to 2.20)‡	
Biliary anastomotic leakage	6 (8)	8 (10)	0.69 (0.22 to 2.13)‡	
Renal failure leading to dialysis — no. (%)	7 (9)	7 (9)	0.79 (0.27 to 2.34)‡	
Median duration of stay (interquartile range) — days				
In the intensive care unit	2 (2 to 5)	2 (1 to 4)	NA	
In the hospital	15 (12 to 20)	15 (12 to 26)	NA	
Retransplantation within 6 mo — no. (%)	3 (4)	6 (8)	0.49 (0.12 to 1.94)‖	
Primary nonfunction — no.	0	1		
Hepatic-artery thrombosis — no.	2	1		
Severe liver laceration — no.**	0	2		
Nonanastomotic biliary strictures — no.	0	2		
Secondary liver dysfunction in the context of multiorgan failure of unknown origin — no.	1	0		
Death of patient within 6 mo — no. (%)	6 (8)	4 (5)	1.46 (0.41 to 5.21)‖	
Multiorgan failure — no.	2	0		
Sepsis — no.	0	3		
Respiratory failure — no.	2	1		
Anoxic brain injury — no.	1	0		
Hemophagocytic syndrome — no.	1	0		

* Plus–minus values are means ±SD. Because of an absence of events in one group or an obvious lack of difference, some treatment differences were not assessed (NA). The widths of the confidence intervals have not been adjusted for multiplicity, and so the inferences drawn from them may not be reproducible.

† The P value for the first assessment of the primary end point is from a chi-square test. The other two P values are based on the unadjusted and adjusted log-binomial regression analysis. For the adjusted analysis, the risk ratio and 95% confidence interval were adjusted for prespecified covariates, including stratification factors (transplantation center and primary sclerosing cholangitis) and established donor risk factors (donor warm-ischemia time and donor risk index).

‡ The treatment effect is expressed as risk ratio and 95% confidence interval, with adjustment for stratification factors.

§ Data were available for 54 patients in the machine-perfusion group and for 60 in the control group. The results of statistical testing are after multiple imputations. The treatment effect is expressed as the mean difference and 95% confidence interval.

¶ Early allograft dysfunction was defined as any one of the following clinical indicators: a bilirubin level of at least 171 μmol per liter (10 mg per deciliter) on postoperative day 7; an international normalized ratio of at least 1.6 on postoperative day 7; or alanine aminotransferase and aspartate aminotransferase levels of more than 2000 U per liter within the first 7 postoperative days. Data were available for all patients. The treatment effect is expressed as a risk ratio and 95% confidence interval, with adjustment for stratification factors.

‖ The treatment effect is expressed as a hazard ratio and 95% confidence interval, with adjustment for stratification factors.

** Liver laceration occurred during the donor hepatectomy and caused severe bleeding and subcapsular hematoma after reperfusion in the recipient, necessitating gauze packing and listing for retransplantation.

PRIMARY END POINT

Symptomatic nonanastomotic biliary strictures occurred in 5 of 78 patients (6%) in the machine-perfusion group and in 14 of 78 (18%) in the control group (risk ratio, 0.36; 95% confidence interval [CI], 0.14 to 0.94; P=0.03). When the analysis was adjusted for stratification factors and prespecified donor risk factors in the log-binomial regression model, the result remained essentially the same (Table 2). These findings were confirmed in the time-to-event analyses that used the Kaplan–Meier method and Cox regression analysis (hazard ratio, 0.32; 95% CI, 0.11 to 0.89; P=0.03; P=0.03 also by the log-rank test) (Fig. 1).

All 19 patients who had symptomatic nonanastomotic strictures received the diagnosis before the trial magnetic resonance cholangiography was performed at 6 months after transplantation, and all had clinical symptoms or cholestatic laboratory tests (or both) that supported this diagnosis (Table 3). Blinded review of the cholangiograms in symptomatic patients confirmed radiologic evidence of nonanastomotic strictures.

A sensitivity analysis that involved all the patients who completed 6 months of follow-up, including the trial magnetic resonance cholangiography, did not change the conclusion (Table S2). Given the small between-group difference in the static cold-ischemia time, we conducted a post hoc sensitivity analysis with this variable as a covariate in the log-binomial regression model; the conclusion did not change.

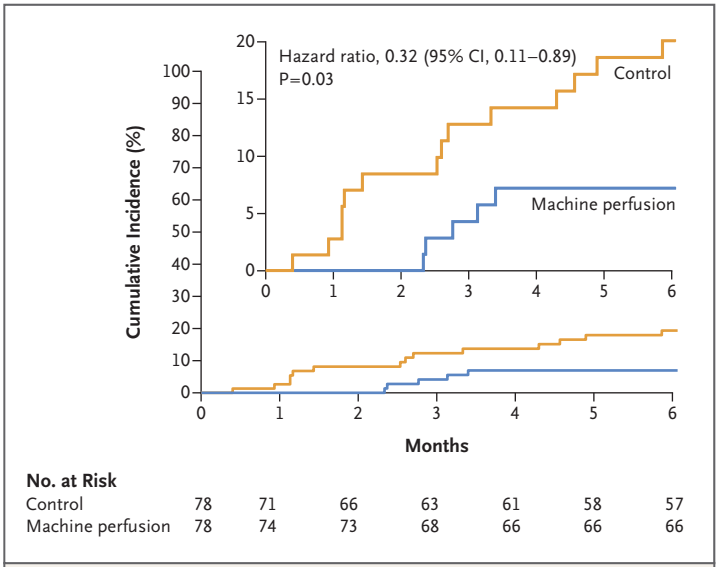


Figure 1. Cumulative Incidence of Symptomatic Nonanastomotic Biliary Strictures.

Shown are the time-to-event Kaplan–Meier curves for symptomatic nonanastomotic biliary strictures within 6 months after liver transplantation (primary end point). The hazard ratio was adjusted for stratification factors (transplantation center and primary sclerosing cholangitis) and for prespecified, established donor risk factors (donor warm-ischemia time and donor risk index); the P value is from a Cox regression analysis. P=0.03 also by the log-rank test. The inset shows the same data on an enlarged y axis.

SECONDARY END-POINT MEASURES

Intraoperatively, the postreperfusion syndrome, which was defined as a decrease of more than 30% in the mean arterial blood pressure, occurred less frequently in recipients of a machine-perfused liver than in those in the control group (12% vs. 27%; adjusted risk ratio, 0.43; 95% CI, 0.20 to 0.91). This difference remained when we included increased inotropic support in the definition (Table 2). In line with this, the mean (±SD) serum potassium levels immediately after transplantation were lower in the machine-perfusion group than in the control group (4.1±0.7 mmol per liter vs. 4.4±1.1 mmol per liter; mean difference, −0.4 mmol per liter; 95% CI, −0.1 to −0.6).

Early allograft dysfunction occurred in 20 machine-perfused livers (26% of the patients), as compared with 31 control livers (40% of the patients) (adjusted risk ratio, 0.61; 95% CI, 0.39 to 0.96). There were no cases of primary nonfunction in the machine-perfusion group, but one case was observed in the control group.

The cumulative number of treatments for nonanastomotic biliary strictures and related complications within 6 months after transplantation was lower by a factor of almost 4 in the machine-perfusion group than in the control group (Table 3). Two patients, both in the control group, underwent retransplantation because of severe nonanastomotic strictures. There were no between-group differences in the incidence of anastomotic biliary leakage or strictures (Table 2). Results of the blinded review of all cholangiograms are presented in Tables S3 through S7.

Laboratory analyses of serum liver-function tests are presented in Figure S2. In accordance with the higher percentage of patients with symptomatic nonanastomotic biliary strictures in the control group than in the machine-perfusion group, serum cholestasis markers in the control group were higher than those in the

Table 3. Clinical Symptoms, Laboratory Abnormalities, and Biliary Interventions in Patients with Symptomatic Nonanastomotic Biliary Strictures (NAS).*

Variable	Time from Transplantation to First Signs of NAS	Cholestatic Laboratory Tests at Time of Detection of NAS†			NAS-Related Biliary Treatment or Intervention			
		γ-Glutamyl-transferase	Alkaline Phosphatase	Bilirubin‡	Antibiotics for Cholangitis	Endoscopic or Percutaneous Stenting	Readmission	Reoperation
	no. of days	*U/liter*		*μmol/liter*	*no. of events*			
Machine-perfusion group								
Patient no.								
1	94	149	316	30	1	—	1	—
2	70	219	222	157	1	1	1	—
3	83	91	217	30	—	2	2	—
4	102	48	139	14	1§	—	1	—
5	71	743	547	9	1	2	1	—
Total no. of events					4	5	6	—
Control group								
Patient no.								
6	34	991	754	40	1	—	—	—
7	28	338	962	320	1	—	1	Retransplantation for NAS
8	12	63	157	8	—	—	—	Resection of bile duct¶
9	81	832	1150	31	1	2	1	—
10	176	618	650	30	—	1	1	—
11	34	1091	250	48	1	2	3	—
12	129	695	653	30	1	1	1	—
13	147	226	733	22	1	1	1	—
14	137	1430	1065	77	1	5	1	—
15	43	366	323	19	4	1	4	—
16	76	189	144	9	—	2	2	—
17	35	370	272	19	1	2	—	—
18	100	560	317	200	1	4	1	—
19	78	2801	4017	376	1	1	1	Retransplantation for NAS
Total no. of events					14	22	17	3

* All the patients had radiologically confirmed nonanastomotic strictures of the donor bile ducts that were detected before the trial magnetic resonance cholangiography at 6 months after transplantation and in the presence of a patent hepatic artery. Cholangiographic details are provided in the Supplementary Appendix. Interventions for other types of biliary complications, such as anastomotic strictures, are excluded from this overview.

† Laboratory results were based on the nearest available prespecified sample point at 1 month, 3 months, or 6 months after transplantation.

‡ To convert the values for serum bilirubin from micromoles per liter to milligrams per deciliter, divide by 17.1.

§ The patient had clinical evidence of bacterial cholangitis. Cholangiography revealed multiple nonanastomotic strictures of the intrahepatic donor bile ducts. Despite treatment with antibiotic agents and ursodeoxycholic acid, the values on the cholestatic laboratory tests worsened; at 6 months after transplantation, the alkaline phosphatase level was 378 IU per liter and the γ-glutamyltransferase level was 263 U per liter.

¶ The patient underwent resection of a necrotic extrahepatic bile duct. There were multiple strictures of the intrahepatic bile ducts on cholangiography.

machine-perfusion group at 3 months (alkaline phosphatase and bilirubin) and 6 months (alkaline phosphatase). There were no relevant differences between the two groups in the use of renal-replacement therapy, in the durations of stay in the intensive care unit or hospital, or in graft and patient survival at 1 year (Table 2 and Fig. S3).

SAFETY AND ADVERSE EVENTS

The distribution of patients for whom adverse events were reported was similar in the two groups (Table 4). There was no relevant clinical difference between the two groups in the severity of adverse events (Table S8).

Table 4. All Reported Adverse Events within 6 Months after Transplantation.*

Event	Machine Perfusion (N=78)	Control (N=78)
Total no. of events	644	694
Infection — no. (%)	131 (20)	162 (23)
Rejection of transplanted liver — no. (%)	9 (1)	16 (2)
Renal event — no. (%)	47 (7)	36 (5)
Hepatic event — no. (%)	91 (14)	111 (16)
Cardiovascular event — no. (%)	52 (8)	52 (7)
Respiratory event — no. (%)	36 (6)	29 (4)
Neurologic event — no. (%)	62 (10)	55 (8)
Gastrointestinal event — no. (%)	43 (7)	51 (7)
Hematologic event — no. (%)	39 (6)	41 (6)
Dermatologic event — no. (%)	19 (3)	12 (2)
Endocrine event — no. (%)	20 (3)	25 (4)
Cancer — no. (%)	3 (<1)	1 (<1)
Miscellaneous adverse event — no. (%)	91 (14)	103 (15)
Device error — no. (%)†	1 (<1)	0

* The data shown are the numbers of reported adverse events; the percentages are based on the total number of reported adverse events (rather than on the total number of patients). Patients could have had more than one event, and no statistical test was applied to these data. Percentages may not total 100 because of rounding.

† Leakage of the disposable tubing set was reported in one case before machine perfusion was started. After replacement of this disposable set, high flows were noted owing to a malfunctioning pressure sensor caused by a user error. This was immediately corrected without injury to the liver.

DISCUSSION

In this trial involving patients receiving a liver graft from a donor after circulatory death, those who had been randomly assigned to receive the liver graft after hypothermic oxygenated machine perfusion had a risk of symptomatic nonanastomotic biliary strictures within 6 months after transplantation that was approximately two thirds lower than those who had been randomly assigned to receive the liver graft after conventional static cold preservation alone. The lower incidence of this type of cholangiopathy was both statistically and clinically significant.

Nonanastomotic biliary strictures are a result of incomplete recovery from biliary ischemia–reperfusion injury, resulting in fibrotic narrowing of the bile-duct lumen and obstruction of bile flow.[1,3] Although some patients can be treated with endoscopic or percutaneous interventions, strictures are often resistant to dilatations and stenting, and retransplantation may remain the only definitive therapy.[18,22,24] In the present trial, the cumulative number of interventions for nonanastomotic biliary strictures and antibiotic therapy for related cholangitis was lower by a factor of almost 4 among machine-perfused livers than among control livers. Two patients in the control group underwent retransplantation within 6 months because of severe cholangiopathy.

The protective effect of machine perfusion was also shown by the lower risk of postreperfusion syndrome and early allograft dysfunction. Graft reperfusion is often accompanied by hepatic release of potassium and circulatory instability.[28] In a clinical pilot study,[29] a reduction in serum potassium levels was observed after the transplantation of hypothermic machine-perfused livers, and this benefit was confirmed in the current prospective trial.

Important advantages of hypothermic machine perfusion over other dynamic preservation methods, such as normothermic machine perfusion, are its relative simplicity and intrinsic safety. Technical malfunction leading to insufficient hepatic perfusion would not immediately be detrimental because the organ is maintained at low temperature. This situation differs from normothermic machine perfusion, in which device or operator errors result in warm ischemia and may lead to organ loss.[30-32] Another advantage of hypothermic machine perfusion is that it is effective after static cold storage. Although a transportable hypothermic perfusion device is currently under clinical investigation (ClinicalTrials.gov number, NCT03484455), it is still undetermined whether this provides additional benefit.

Despite the restoration of ATP, hepatic metabolism remains suppressed and livers do not produce bile during this type of machine perfusion. Although the release of mitochondrial flavin mononucleotide into the perfusate has been correlated with hepatic function after transplantation, it remains unknown whether this also predicts the risk of cholangiopathy.[33] In contrast to normothermic machine perfusion, hypothermic machine perfusion is, therefore, currently not considered to be a tool for viability testing before transplantation; rather, it is a method to reduce the incidence of ischemia–reperfusion injury. This makes it suited for donor livers with an increased risk of development of ischemia-related complications, such as livers obtained from donors after circulatory death. To this end, hypothermic and normothermic machine perfusion serve different goals and are not competing techniques. The two techniques can be applied sequentially with complementary benefits.[34-36] Whether hypothermic machine perfusion is also beneficial in the transplantation of livers obtained from brain-dead donors is the subject of ongoing clinical trials (NCT01317342 and NCT03124641).

In the present trial, machine perfusion did not have an effect on patient or graft survival. Given the high percentage of patients who survive after liver transplantation and the relatively low risk of graft loss, much larger trials would be needed to detect an effect on these outcome measures.

Reimbursement of this new technology by health care funders will involve a health-economic evaluation. Costs for transplantation of a liver from a donor after circulatory death are 25 to 30% higher than those for transplantation of livers from brain-dead donors, mainly because of the higher incidence of biliary complications.[37,38] The prevention of post-transplantation cholangiopathy may not only increase the acceptance for transplantation of liver grafts obtained from donors after circulatory death but may also make the use of machine perfusion cost-effective.

In this randomized trial involving patients who underwent transplantation of a liver obtained from a donor after circulatory death, we found that hypothermic oxygenated machine perfusion led to a lower incidence of symptomatic nonanastomotic biliary strictures than conventional static cold preservation.

Supported by a grant for applied medical research (1404-012) from Fonds NutsOhra.

Dr. Jochmans reports receiving travel support from Astellas Pharma; and Dr. Porte, receiving donated supplies from Bridge to Life. No other potential conflict of interest relevant to this article was reported.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

A data sharing statement provided by the authors is available with the full text of this article at NEJM.org.

We thank the patients who participated in this trial and the doctors, transplantation and research coordinators, operating room staff, and organ perfusionists working at the six collaborating transplantation centers for their contributions to this trial.

APPENDIX

The authors' full names and academic degrees are as follows: Rianne van Rijn, M.D., Ph.D., Ivo J. Schurink, B.Sc., Yvonne de Vries, M.D., Ph.D., Aad P. van den Berg, M.D., Ph.D., Miriam Cortes Cerisuelo, M.D., Ph.D., Sarwa Darwish Murad, M.D., Ph.D., Joris I. Erdmann, M.D, Ph.D., Nicholas Gilbo, M.D., Ph.D., Robbert J. de Haas, M.D., Ph.D., Nigel Heaton, M.D., Ph.D., Bart van Hoek, M.D., Ph.D., Volkert A.L. Huurman, M.D., Ph.D., Ina Jochmans, M.D., Ph.D., Otto B. van Leeuwen, Ph.D., Vincent E. de Meijer, M.D., Ph.D., Diethard Monbaliu, M.D., Ph.D., Wojciech G. Polak, M.D., Ph.D., Jules J.G. Slangen, M.D., Roberto I. Troisi, M.D., Ph.D., Aude Vanlander, M.D., Ph.D., Jeroen de Jonge, M.D., Ph.D., and Prof. Robert J. Porte, M.D., Ph.D.

The authors' affiliations are as follows: the Department of Surgery, Section of Hepatobiliary Surgery and Liver Transplantation (R.R., Y.V., O.B.L., V.E.M., R.J.P.), the Departments of Gastroenterology and Hepatology (A.P.B.) and Radiology (R.J.H., J.J.G.S.), University of Groningen, University Medical Center Groningen, Groningen, the Departments of Surgery (I.J.S., W.G.P., J.J.) and Gastroenterology and Hepatology (S.D.M.), Erasmus University Medical Center, Rotterdam, and the Departments of Surgery (J.I.E., V.A.L.H.) and Gastroenterology and Hepatology (B.H.), Leiden University Medical Center, Leiden — all in the Netherlands; the Institute of Liver Studies, Kings College Hospital NHS Foundation Trust, London (M.C.C., N.H.); the Transplantation Research Group, the Department of Microbiology, Immunology, and Transplantation, Katholieke Universiteit Leuven, and the Department of Abdominal Transplantation Surgery and Coordination, University Hospitals Leuven, Leuven (N.G., I.J., D.M.), and the Department of Transplant Surgery, Ghent University Hospital, Ghent (R.I.T., A.V.) — both in Belgium.

REFERENCES

1. de Vries Y, von Meijenfeldt FA, Porte RJ. Post-transplant cholangiopathy: classification, pathogenesis, and preventive strategies. Biochim Biophys Acta Mol Basis Dis 2018;1864:1507-15.
2. O'Neill S, Roebuck A, Khoo E, Wigmore SJ, Harrison EM. A meta-analysis and meta-regression of outcomes including biliary complications in donation after cardiac death liver transplantation. Transpl Int 2014;27:1159-74.
3. Foley DP, Fernandez LA, Leverson G, et al. Biliary complications after liver transplantation from donation after cardiac death donors: an analysis of risk factors and long-term outcomes from a single center. Ann Surg 2011;253:817-25.
4. Tullius SG, Rabb H. Improving the

supply and quality of deceased-donor organs for transplantation. N Engl J Med 2018;378:1920-9.
5. Summers DM, Counter C, Johnson RJ, Murphy PG, Neuberger JM, Bradley JA. Is the increase in DCD organ donors in the United Kingdom contributing to a decline in DBD donors? Transplantation 2010;90:1506-10.
6. Dutkowski P, Guarrera JV, de Jonge J, Martins PN, Porte RJ, Clavien PA. Evolving trends in machine perfusion for liver transplantation. Gastroenterology 2019;156:1542-7.
7. Xu J, Buchwald JE, Martins PN. Review of current machine perfusion therapeutics for organ preservation. Transplantation 2020;104:1792-03.
8. Lüer B, Koetting M, Efferz P, Minor T. Role of oxygen during hypothermic machine perfusion preservation of the liver. Transpl Int 2010;23:944-50.
9. Monbaliu D, Brassil J. Machine perfusion of the liver: past, present and future. Curr Opin Organ Transplant 2010;15:160-6.
10. Dutkowski P, Furrer K, Tian Y, Graf R, Clavien PA. Novel short-term hypothermic oxygenated perfusion (HOPE) system prevents injury in rat liver graft from non-heart beating donor. Ann Surg 2006;244:968-76.
11. de Rougemont O, Breitenstein S, Leskosek B, et al. One hour hypothermic oxygenated perfusion (HOPE) protects nonviable liver allografts donated after cardiac death. Ann Surg 2009;250:674-83.
12. Schlegel A, Kron P, Graf R, Clavien PA, Dutkowski P. Hypothermic oxygenated perfusion (HOPE) downregulates the immune response in a rat model of liver transplantation. Ann Surg 2014;260:931-7.
13. Guarrera JV, Henry SD, Samstein B, et al. Hypothermic machine preservation in human liver transplantation: the first clinical series. Am J Transplant 2010;10:372-81.
14. Dutkowski P, Schlegel A, de Oliveira M, Müllhaupt B, Neff F, Clavien PA. HOPE for human liver grafts obtained from donors after cardiac death. J Hepatol 2014;60:765-72.
15. van Rijn R, Karimian N, Matton APM, et al. Dual hypothermic oxygenated machine perfusion in liver transplants donated after circulatory death. Br J Surg 2017;104:907-17.
16. van Rijn R, van Leeuwen OB, Matton APM, et al. Hypothermic oxygenated machine perfusion reduces bile duct reperfusion injury after transplantation of donation after circulatory death livers. Liver Transpl 2018;24:655-64.
17. van Rijn R, van den Berg AP, Erdmann JI, et al. Study protocol for a multicenter randomized controlled trial to compare the efficacy of end-ischemic dual hypothermic oxygenated machine perfusion with static cold storage in preventing non-anastomotic biliary strictures after transplantation of liver grafts donated after circulatory death: DHOPE-DCD trial. BMC Gastroenterol 2019;19:40-6.
18. Buis CI, Verdonk RC, Van der Jagt EJ, et al. Nonanastomotic biliary strictures after liver transplantation. 1. Radiological features and risk factors for early vs. late presentation. Liver Transpl 2007;13:708-18.
19. Aggarwal S, Kang Y, Freeman JA, Fortunato FL, Pinsky MR. Postreperfusion syndrome: cardiovascular collapse following hepatic reperfusion during liver transplantation. Transplant Proc 1987;19:Suppl 3:54-5.
20. Hilmi I, Horton CN, Planinsic RM, et al. The impact of postreperfusion syndrome on short-term patient and liver allograft outcome in patients undergoing orthotopic liver transplantation. Liver Transpl 2008;14:504-8.
21. Olthoff KM, Kulik L, Samstein B, et al. Validation of a current definition of early allograft dysfunction in liver transplant recipients and analysis of risk factors. Liver Transpl 2010;16:943-9.
22. Abt P, Crawford M, Desai N, Markmann J, Olthoff K, Shaked A. Liver transplantation from controlled non-heart-beating donors: an increased incidence of biliary complications. Transplantation 2003;75:1659-63.
23. Dubbeld J, Hoekstra H, Farid W, et al. Similar liver transplantation survival with selected cardiac death donors and brain death donors. Br J Surg 2010;97:744-53.
24. Croome KP, McAlister V, Adams P, Marotta P, Wall W, Hernandez-Alejandro R. Endoscopic management of biliary complications following liver transplantation after donation from cardiac death donors. Can J Gastroenterol 2012;26:607-10.
25. Meurisse N, Vanden Bussche S, Jochmans I, et al. Outcomes of liver transplantations using donations after circulatory death: a single-center experience. Transplant Proc 2012;44:2868-73.
26. Taner CB, Bulatao IG, Perry DK, et al. Asystole to cross-clamp period predicts development of biliary complications in liver transplantation using donation after cardiac death donors. Transpl Int 2012;25:838-46.
27. Feng S, Goodrich NP, Bragg-Gresham JL, et al. Characteristics associated with liver graft failure: the concept of a donor risk index. Am J Transplant 2006;6:783-90.
28. Patrono D, Romagnoli R. Postreperfusion syndrome, hyperkalemia and machine perfusion in liver transplantation. Transl Gastroenterol Hepatol 2019;4:68.
29. Burlage LC, Hessels L, van Rijn R, et al. Opposite acute potassium and sodium shifts during transplantation of hypothermic machine perfused donor livers. Am J Transplant 2019;19:1061-71.
30. Selzner M, Goldaracena N, Echeverri J, et al. Normothermic ex vivo liver perfusion using steen solution as perfusate for human liver transplantation: first North American results. Liver Transpl 2016;22:1501-8.
31. Bral M, Gala-Lopez B, Bigam D, et al. Preliminary single-center Canadian experience of human normothermic ex vivo liver perfusion: results of a clinical trial. Am J Transplant 2017;17:1071-80.
32. Nasralla D, Coussios CC, Mergental H, et al. A randomized trial of normothermic preservation in liver transplantation. Nature 2018;557:50-6.
33. Muller X, Schlegel A, Kron P, et al. Novel real-time prediction of liver graft function during hypothermic oxygenated machine perfusion before liver transplantation. Ann Surg 2019;270:783-90.
34. de Vries Y, Matton APM, Nijsten MWN, et al. Pretransplant sequential hypo- and normothermic machine perfusion of suboptimal livers donated after circulatory death using a hemoglobin-based oxygen carrier perfusion solution. Am J Transplant 2019;19:1202-11.
35. Boteon YL, Laing RW, Schlegel A, et al. The impact on the bioenergetic status and oxidative-mediated tissue injury of a combined protocol of hypothermic and normothermic machine perfusion using an acellular haemoglobin-based oxygen carrier: the cold-to-warm machine perfusion of the liver. PLoS One 2019;14(10):e0224066.
36. van Leeuwen OB, de Vries Y, Fujiyoshi M, et al. Transplantation of high-risk donor livers after ex situ resuscitation and assessment using combined hypo- and normothermic machine perfusion: a prospective clinical trial. Ann Surg 2019;270:906-14.
37. Jay CL, Lyuksemburg V, Kang R, et al. The increased costs of donation after cardiac death liver transplantation: caveat emptor. Ann Surg 2010;251:743-48.
38. van der Hilst CS, Ijtsma AJ, Bottema JT, et al. The price of donation after cardiac death in liver transplantation: a prospective cost-effectiveness study. Transpl Int 2013;26:411-18.

ORIGINAL ARTICLE

Trial of Psilocybin versus Escitalopram for Depression

Robin Carhart-Harris, Ph.D., Bruna Giribaldi, B.Sc., Rosalind Watts, D.Clin.Psy., Michelle Baker-Jones, B.A., Ashleigh Murphy-Beiner, M.Sc., Roberta Murphy, M.D., Jonny Martell, M.D., Allan Blemings, M.Sc., David Erritzoe, M.D., and David J. Nutt, M.D.

ABSTRACT

BACKGROUND

Psilocybin may have antidepressant properties, but direct comparisons between psilocybin and established treatments for depression are lacking.

METHODS

In a phase 2, double-blind, randomized, controlled trial involving patients with long-standing, moderate-to-severe major depressive disorder, we compared psilocybin with escitalopram, a selective serotonin-reuptake inhibitor, over a 6-week period. Patients were assigned in a 1:1 ratio to receive two separate doses of 25 mg of psilocybin 3 weeks apart plus 6 weeks of daily placebo (psilocybin group) or two separate doses of 1 mg of psilocybin 3 weeks apart plus 6 weeks of daily oral escitalopram (escitalopram group); all the patients received psychological support. The primary outcome was the change from baseline in the score on the 16-item Quick Inventory of Depressive Symptomatology–Self-Report (QIDS-SR-16; scores range from 0 to 27, with higher scores indicating greater depression) at week 6. There were 16 secondary outcomes, including QIDS-SR-16 response (defined as a reduction in score of >50%) and QIDS-SR-16 remission (defined as a score of ≤5) at week 6.

RESULTS

A total of 59 patients were enrolled; 30 were assigned to the psilocybin group and 29 to the escitalopram group. The mean scores on the QIDS-SR-16 at baseline were 14.5 in the psilocybin group and 16.4 in the escitalopram group. The mean (±SE) changes in the scores from baseline to week 6 were −8.0±1.0 points in the psilocybin group and −6.0±1.0 in the escitalopram group, for a between-group difference of 2.0 points (95% confidence interval [CI], −5.0 to 0.9) (P=0.17). A QIDS-SR-16 response occurred in 70% of the patients in the psilocybin group and in 48% of those in the escitalopram group, for a between-group difference of 22 percentage points (95% CI, −3 to 48); QIDS-SR-16 remission occurred in 57% and 28%, respectively, for a between-group difference of 28 percentage points (95% CI, 2 to 54). Other secondary outcomes generally favored psilocybin over escitalopram, but the analyses were not corrected for multiple comparisons. The incidence of adverse events was similar in the trial groups.

CONCLUSIONS

On the basis of the change in depression scores on the QIDS-SR-16 at week 6, this trial did not show a significant difference in antidepressant effects between psilocybin and escitalopram in a selected group of patients. Secondary outcomes generally favored psilocybin over escitalopram, but the analyses of these outcomes lacked correction for multiple comparisons. Larger and longer trials are required to compare psilocybin with established antidepressants. (Funded by the Alexander Mosley Charitable Trust and Imperial College London's Centre for Psychedelic Research; ClinicalTrials.gov number, NCT03429075.)

From the Centre for Psychedelic Research, Department of Brain Sciences, Faculty of Medicine, Imperial College London, London. Address reprint requests to Dr. Carhart-Harris at Imperial College London, Burlington Danes Bldg., 160 Du Cane Rd., London W12 0NN, United Kingdom, or at r.carhart-harris@imperial.ac.uk.

Dr. Carhart-Harris, Ms. Giribaldi, and Dr. Watts contributed equally to this article.

N Engl J Med 2021;384:1402-11.
DOI: 10.1056/NEJMoa2032994

Major depressive disorder affects approximately 10% of the general population in the United Kingdom, impairs patients' lives, and is costly to society.[1] Selective serotonin-reuptake inhibitors are first-line treatments for major depressive disorder; however, these drugs take several weeks to work and, in some patients, do not induce a response.[2] Escitalopram, a selective serotonin-reuptake inhibitor, is representative of the currently used antidepressants in terms of safety and efficacy.[2,3]

The psychedelic compound psilocybin is the phosphorylated ester of its metabolite, psilocin (4-OH-N,N-dimethyltryptamine). Psilocybin and psilocin occur naturally in the psychoactive psilocybe genus of mushrooms. As with other traditional psychedelic substances,[4,5] the main effects of psilocin occur through serotonin 5-hydroxytryptamine type 2A (5-HT_{2A}) receptor agonism, which is part of a pathway implicated in depression.[4-6] Psilocybin showed promise as an adjunct to psychotherapy for mood disorders and addiction in the mid-20th century.[7,8]

One open-label trial[9] and four randomized, controlled clinical trials[10-13] of psilocybin for depression and anxiety have been conducted.[5,9-13] Reductions in depressive symptoms after the administration of one or two doses of psilocybin were observed in trials across several patient populations,[9-14] including a small open-label trial involving patients with treatment-resistant depression,[9,14] the results of which informed the current trial. We performed a phase 2, double-blind, randomized, controlled trial involving patients with long-standing, moderate-to-severe major depressive disorder to compare psilocybin with escitalopram over a 6-week period.

Methods

Trial Oversight

A Schedule 1 drug license from the U.K. Home Office was obtained by the investigators, and the trial was approved by the Brent Research Ethics Committee, the U.K. Medicines and Healthcare Products Regulatory Agency, the Health Research Authority, the Imperial College London Joint Research Compliance and General Data Protection Regulation Offices, and the risk assessment and trial management review board at the trial site (the National Institute for Health Research [NIHR] Imperial Clinical Research Facility [CRF]). Psilocybin was provided by COMPASS Pathways, and escitalopram and placebo were provided by the Pharmacy Manufacturing Unit at Guy's and St. Thomas's Hospital.

This was an investigator-initiated, university-sponsored trial. All medicinal products under investigation were stored and dispensed by Invicro. Trial visits occurred at the NIHR CRF from January 2019 through March 2020. The first author designed the trial and wrote the first draft of the manuscript with assistance from the second author. The second through seventh authors performed the trial and collected the data, and the eighth author analyzed the data. Clinical oversight of the trial was provided by the third, penultimate, and last authors, and the overall trial was overseen by the last author. The authors vouch for the accuracy and completeness of the data and for the fidelity of the trial to the protocol (available with the full text of this article at NEJM.org). There was no industry involvement in the collection or analysis of the data, and no agreements were in place between the authors and any commercial entity.

Patients

Men and women between the ages of 18 and 80 years were recruited formally through trial networks, informally through social media, and through other sources, which directed patients to a recruitment website. The main exclusion criteria were an immediate family or personal history of psychosis, medically significant health conditions that make a person unsuitable to participate in the trial (as assessed by a physician), a history of serious suicide attempts, a positive pregnancy test, contraindications to taking selective serotonin-reuptake inhibitors or undergoing magnetic resonance imaging (MRI), previous use of escitalopram (although previous use of psilocybin was allowed), or suspected or known presence of a preexisting psychiatric condition (e.g., borderline personality disorder) that could jeopardize rapport between the patient and their two mental health caregivers within the trial. Additional details about the trial exclusion criteria are provided in the protocol.

Information about the trial, including inclusion and exclusion criteria, was made available online at the Centre for Psychedelic Research website (www.imperial.ac.uk/psychedelic-research-centre), the ClinicalTrials.gov website, the MQ mental health research recruitment platform (www.mqmentalhealth.org/home/), and the ISRCTN

Registry website. Volunteers initiated contact by emailing the recruitment coordinator after hearing about the trial. Most of the recruited patients referred themselves. Candidates were sent a patient information sheet and invited to a telephone screening. Assessments with the 17-item Hamilton Depression Scale (HAM-D-17) were performed by means of a video call; a score of at least 17 (indicating moderate-to-severe major depressive disorder) on a scale that ranges from 0 to 52, with higher scores indicating greater depression, was required for trial enrollment. Confirmation of a diagnosis of depression and medical history were obtained from the patient's general physician. Eligible patients then underwent face-to-face physical and mental health assessments with a trial psychiatrist, which was followed by their first psychological support session (see the protocol). The patients discontinued any use of a psychiatric medication before starting the trial, with full discontinuation occurring at least 2 weeks before starting a trial medication; any use of psychotherapy was stopped at least 3 weeks before starting a trial medication.

After the telephone screening, each patient was assigned to two supervising mental health professionals. The role of these mental health professionals was to build a therapeutic alliance with the patient before, during, and after each day of dosing. (Additional details are provided in Section S2.8 of the Supplementary Appendix, available at NEJM.org.) One of the pair was a clinical psychologist, psychotherapist, or psychiatrist, and the other could be an equivalent grade clinician or trainee. The mental health professionals were present for all trial visits. Baseline assessments were completed 7 to 10 days before trial visit 1.

TRIAL DESIGN

Randomization (performed with the use of a random-number generator) was implemented by staff members who were not part of the research team. (Details regarding the randomization process are provided in Section S2.6.) All the patients provided written informed consent and, after screening, were required to attend six visits over a 6-week trial period. Procedures for the ingestion of psychotherapeutic agents and size- and color-matched placebo capsules were consistent between the trial groups.

At visit 1 (baseline), all the patients underwent functional MRI, completed a battery of cognitive and affective processing tasks (data not yet analyzed), and attended a preparatory therapeutic session. At visit 2, which occurred 1 day after visit 1, the patients in the psilocybin group received 25 mg of psilocybin, and those in the escitalopram group received 1 mg of psilocybin, which was presumed to have negligible activity (dosing-day 1). To standardize expectations, all the patients were informed that they would receive psilocybin, but the dose was not disclosed to them. The medications and placebos were prepackaged with nondisclosing labels, and all the investigators and medication administering staff were unaware of the trial-group assignments. The dosing days for each patient were supervised by the two mental health professionals who had been assigned to the patient. Supervision consisted of caring for the physical and psychological well-being of the patient and responding to signs of patient discomfort during and immediately after the administration of a trial medication.[15] (Additional details regarding psychological support are provided in Section S2.8.) A trial psychiatrist assessed eligibility for discharge when the functional status of a patient had returned to the baseline level.

Before the patients left the CRF after visit 2, they received a screw-top bottle of capsules and were instructed to take one capsule each morning until their next scheduled day of psilocybin dosing. The capsules contained either microcrystalline cellulose (placebo), which were given to the patients who had received the 25-mg dose of psilocybin on dosing-day 1, or 10 mg of escitalopram, which were given to the patients who had received the 1-mg dose of psilocybin on dosing-day 1. Visit 3 occurred 1 day after dosing-day 1 and included a psychological debriefing. An additional debriefing by telephone or video call occurred 1 week later.

At visit 4, which occurred 3 weeks after dosing-day 1, the patients received their second dose of psilocybin or placebo (dosing-day 2), and at visit 5 (the next day), a psychological integration session involving open, attentive listening was held. After dosing-day 2, the patients were asked to take two capsules each morning (either placebo in the psilocybin group or an increased dose of 20 mg of escitalopram in the escitalopram group) for the next 3 weeks.

Three weeks after visit 5, the patients returned

for their final trial visit (visit 6) for the assessment of the primary outcome. The structure of this visit was similar to that of visit 1 and involved the performance of functional MRI (6 weeks after the first), cognitive and affective processing tasks, final clinician-rated assessments, and psychological debriefing. After these assessments, the patient and the trial staff were informed of the trial-group assignment, and a trial psychiatrist discussed future treatment options. In the escitalopram group, discontinuation of the trial drug was managed by the patients and their general physicians. After week 6, the patients were followed for 6 months by the investigators, but these data have not yet been fully collected. The initial trial design included a placebo group that was to receive 1 mg of psilocybin and placebo, but this group was not included in the final protocol because it was determined that a trial involving three groups would be too complex and expensive to conduct and power adequately, given the resources that were available at the time. The data obtained from an imaging group in the trial, in which functional MRI was used to predict responses to the trial drugs, have not been analyzed.

OUTCOMES

The primary clinical outcome was the change from baseline in the score on the 16-item Quick Inventory of Depressive Symptomatology–Self-Report (QIDS-SR-16; scores range from 0 to 27, with higher scores indicating greater depression) at 6 weeks. Secondary outcomes included response at 6 weeks according to the QIDS-SR-16 (defined as a decrease in score of ≥50% from baseline); remission at 6 weeks according to the QIDS-SR-16 (defined as a score of 0 to 5); change in the score on the 14-item QIDS-SR (QIDS-SR-14) from the day before to the day after dosing-day 1; and the changes from baseline to week 6 in the scores on the Beck Depression Inventory 1A (BDI-1A), the 17-item Hamilton Depression Rating Scale (HAM-D-17), and the Montgomery and Asberg Depression Rating Scale (MADRS). Other secondary outcomes were the changes from baseline to 6 weeks in the scores on the Flourishing Scale (FS), the Spielberger's Trait Anxiety Inventory (STAI), the Brief Experiential Avoidance Questionnaire (BEAQ),[16] the Work and Social Adjustment Scale (WSAS), the Snaith Hamilton Anhedonia Pleasure Scale (SHAPS), the Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS), and the Suicidal Ideation Attributes Scale (SIDAS), as well as the scores at 6 weeks on the Psychotropic-Related Sexual Dysfunction Questionnaire (PRSexDQ), the Laukes Emotional Intensity Scale (LEIS),[17] and the Emotional Breakthrough Inventory,[18] which assessed acute subjective experiences after each dosing day (Fig. S4 and Table S5). An investigator-constructed patient-rated scale (the Post-Treatment Changes Scale [PTCS]) was used as a safety outcome measure for assessing post-treatment side effects and other phenomena that previous work has associated with psychedelic compounds or selective serotonin-reuptake inhibitors (Section S2.11 and Table S2). Additional details of these outcomes are provided in the protocol.

ADVERSE EVENTS

Adverse events were recorded at every visit and telephone call from dosing-day 1 through week 6. Adverse events were assessed by asking "how have you been since your last visit?" or on the basis of events that were observed at the trial site. Additional details of the criteria used for the reporting of adverse events are provided in the protocol. All adverse events that occurred or worsened between dosing-day 1 and week 6 were recorded and coded with the use of the *Medical Dictionary for Regulatory Activities*, version 23.0.

STATISTICAL ANALYSIS

The clinical component of the trial was powered on the basis of data from previous trials[10,14] and on an assumption of equal variance for both trial drugs with respect to the primary outcome and the ability to detect a difference between the groups at a two-sided level of $P<0.05$ with 80% power. This would require 20 patients per trial group, and we proposed recruiting a minimum of 30 patients per group (60 in total for the trial). Additional details are provided in Sections 4.2.1 and 10 of the protocol. All the patients who had undergone randomization were included in an intention-to-treat analysis.

The change from baseline in the score on the QIDS-SR-16 at week 6 (the primary outcome) was compared between the trial groups with the use of repeated-measures analysis of covariance (ANCOVA), with adjustment for baseline scores. Logistic regression, with adjustment for baseline scores, was used to analyze the secondary out-

comes of response and remission according to the QIDS-SR-16, as well as the additional outcomes of response and remission according to the BDI-1A, the HAM-D-17, and the MADRS. The changes from baseline to week 6 in the scores on the HAM-D-17, the QIDS-SR-14, the MADRS, the WEMWBS, the FS, the BEAQ, the WSAS, the SHAPS, the STAI, and the LEIS were analyzed with the use of ANCOVA or repeated-measures ANCOVA, with adjustment for baseline (if possible). The changes from baseline to week 6 in the scores on the BDI-1A and the SIDAS were analyzed with the use of the permutation test stratified according to baseline scores. The score at 6 weeks on the PRSexDQ was analyzed with the use of a Wilcoxon test. The score at 6 weeks on the PTCS was analyzed with the use of the Jonckheere–Terpstra trend test.

The results are presented as means, adjusted for baseline values. There was no imputation for missing data except for the WSAS, for which missing data were imputed with the overall mean calculated from nonmissing data. Because of the absence of a prespecified plan for adjustment of confidence intervals for multiple comparisons of secondary outcomes, P values are not reported and no clinical conclusions can be drawn from these data.

RESULTS

PATIENTS

Approximately 1000 patients underwent screening by telephone (103 of whom also attended a formal screening visit). A total of 891 patients did not meet inclusion criteria (19 of whom had a coexisting psychiatric condition), and 50 declined to participate (Section S2.7). Thus, 59 patients were enrolled and underwent randomization; 30 were assigned to the psilocybin group and 29 to the escitalopram group. Of the 59 patients enrolled, 23 (39%) had completely discontinued a psychiatric medication before entering the trial, and 4 (7%) had discontinued psychotherapy. In the escitalopram group, 5 of 29 patients did not complete the protocol requirements: 4 stopped taking their escitalopram capsules because of adverse events, and 1 missed dosing-day 2 and subsequent visits owing to restrictions related to coronavirus disease 2019 (Covid-19). One patient in the escitalopram group guessed that the capsules contained escitalopram and reduced the dose by half (from 20 mg to 10 mg) because of perceived adverse events; a reduction in the escitalopram dose to 10 mg was permitted in the protocol because it reflects clinical practice. In the psilocybin group, 3 of 30 patients did not complete all dosing procedures: 2 missed dosing-day 2 and subsequent visits because of Covid-19–related restrictions, and 1 stopped taking daily placebo capsules after guessing their content.

The mean age of the patients enrolled in the trial was 41 years; 20 (34%) were women and most were White. Depression had been present for a mean of 22 years among the patients in the psilocybin group and for a mean of 15 years among those in the escitalopram group; QIDS-SR-16 scores at baseline were 14.5 and 16.4, respectively. There was more alcohol use among the patients in the escitalopram group than in the psilocybin group; other characteristics were similar in the groups (Table 1).

EFFICACY OUTCOMES

The mean (±SE) change from baseline in the score on the QIDS-SR-16 at week 6 (the primary outcome) was −8.0±1.0 in the psilocybin group and −6.0±1.0 in the escitalopram group (difference, −2.0; 95% confidence interval [CI], −5.0 to 0.9; P=0.17), indicating no significant difference between the trial groups (Fig. 1 and Table 2). A per-protocol analysis produced similar results (Table S1).

The results of the secondary-outcome analyses are provided in Figure 1, Table 2, and Figures S3 and S4. A QIDS-SR-16 response at 6 weeks occurred in 21 patients (70%) in the psilocybin group and in 14 patients (48%) in the escitalopram group (difference, 22 percentage points; 95% CI, −3 to 48, indicating no significant difference) (Table 2). QIDS-SR-16 remission at week 6 occurred in 17 patients (57%) in the psilocybin group and in 8 patients (28%) in the escitalopram group (difference, 28.1 percentage points; 95% CI, 2.3 to 53.8) (Table 2). Other secondary measures of depression (changes from baseline to week 6 in the scores on the BDI-1A, HAM-D-17, and MADRS) and the between-group differences in the scores on other scales mostly favored psilocybin over escitalopram, although the confidence intervals for the between-group differences were

Table 1. Demographic and Clinical Characteristics of the Patients at Baseline.*

Characteristic	Psilocybin (N=30)	Escitalopram (N=29)
Demographic		
Age (range) — yr	43.3±11.7 (21–64)	39.1±9.7 (22–60)
Female sex — no. (%)	11 (37)	9 (31)
White race — no. (%)†	28 (93)	24 (83)
Employment status — no. (%)		
Employed	21 (70)	21 (72)
Student	2 (7)	3 (10)
Unemployed	7 (23)	5 (17)
University level education — no. (%)	22 (73)	23 (79)
No previous psilocybin use — no. (%)	22 (73)	21 (72)
Weekly alcohol use (range) — g‡	36.8±43.1 (0–160)	67.7±66.6 (0–240)
Discontinued psychiatric medication for trial — no. (%)	11 (37)	12 (41)
Clinical		
Duration of illness (range) — yr	22.1±10.7 (3–44)	15.1±11.0 (2–46)
No. of psychiatric medications previously used (range)	2.2±1.6 (0–6)	1.8±1.5 (0–5)
Previous use of psychotherapy — no. (%)	28 (93)	26 (90)
QIDS-SR-16 score at pretreatment baseline (range)§	14.5±3.9 (7–23)	16.4±4.1 (6–22)
HAM-D-17 score at pretreatment baseline (range)¶	19.2±2.3 (16–23)	18.4±3.4 (11–26)
BDI-1A score at pretreatment baseline (range)‖	29.1±6.8 (16–41)	28.7±7.0 (10–44)

* Plus–minus values are means ±SD. Pretreatment baseline was 7 to 10 days before dosing-day 1.
† Race was reported by the patients.
‡ To convert grams to U.K. units, divide by 8.
§ The scores on the 16-item Quick Inventory of Depressive Symptomatology–Self-Report (QIDS-SR-16) range from 0 to 27, with higher scores indicating greater depression.
¶ The scores on the 17-item Hamilton Depression Rating Scale (HAM-D-17) range from 0 to 50, with higher scores indicating greater depression. At screening, which was typically a few weeks before pretreatment baseline, all the patients had a score of at least 17 on the HAM-D-17. The depression scores reported in this table are from pretreatment baseline and not screening.
‖ The scores on the Beck Depression Inventory 1A (BDI-1A) range from 0 to 63, with higher scores indicating greater depression.

not adjusted for multiple comparisons (Table 2). Ratings on the Emotional Breakthrough Inventory are provided in Figure S7. With respect to the primary and secondary outcomes, the absolute values that were not adjusted for baseline values (Table S12) were in the same general direction as those in the adjusted analyses. Multiple imputation was performed for two patients with missing baseline values on the WSAS, and the results were similar to those in the main (baseline-adjusted) analysis (Section S13). A post hoc analysis for the imbalanced use of alcohol between the trial groups showed results in the same direction as those in the main analysis (Section S12).

SAFETY

No serious adverse events were observed in either trial group. The percentage of patients reporting adverse events was similar in the two groups: 26 (87%) in the psilocybin group and 24 (83%) in the escitalopram group (Table 3, and Fig. S6). The percentage of patients who had increased anxiety and dry mouth was higher in the escitalopram group than in the psilocybin group. Adverse events in the psilocybin group typically occurred within 24 hours after the dosing day; the most common adverse event was headache. A complete list of the adverse events that occurred in the trial groups is provided in Table S5.

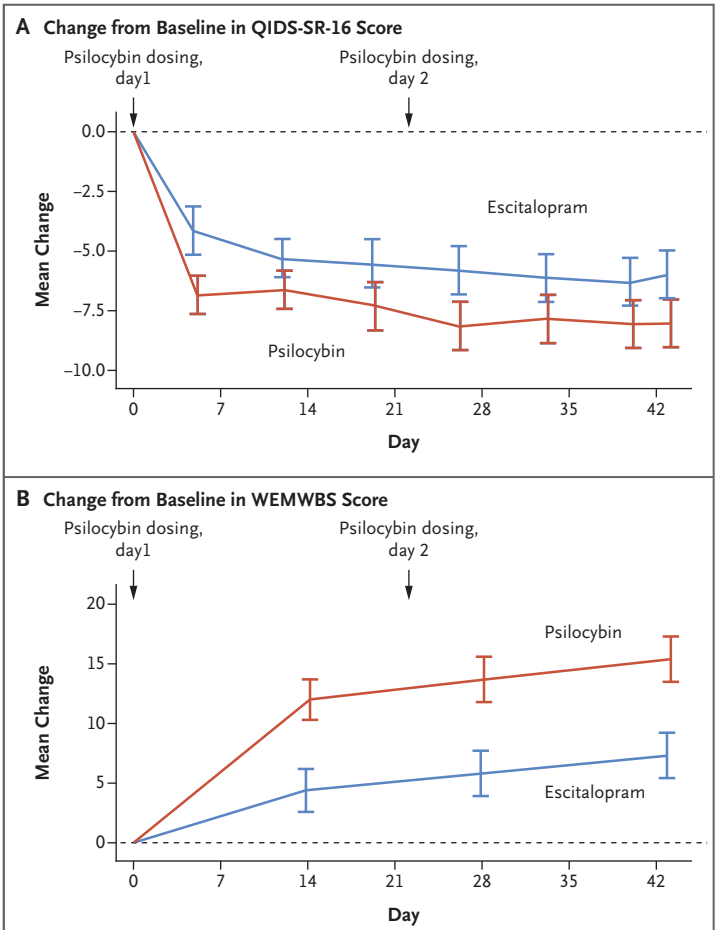


Figure 1. Change in Depression Severity and in Well-Being over 6 Weeks.

Panel A shows the mean change from baseline in the score on the 16-item Quick Inventory of Depressive Symptomatology–Self-Report (QIDS-SR-16; on which scores range from 0 to 27, with higher scores indicating greater depression). Panel B shows the mean change in the score on the Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS; on which scores range from 14 to 70, with higher scores indicating greater mental well-being). These were the only outcomes for which there were data every week (QIDS-SR-16) or every 2 weeks (WEMWBS) and for which there were prespecified hypotheses (Section S2.1 in the Supplementary Appendix). P values are not shown because there was no correction for multiple comparisons in the analyses of the WEMWBS (a secondary outcome) or of the outcomes at any intermediate time points. I bars indicate standard errors.

When cued to report on specific emotional and side-effect–related phenomena through the PTCS (a description of this scale is provided in Section S2.11), patients in the psilocybin group reported greater perceived improvements in the ability to cry and feel compassion, intense emotion, and pleasure and reported feeling less drowsy than those in the escitalopram group (Table S2). No cases of visual perceptual changes, psychotic symptoms, or dependency-related behaviors were observed or reported in either trial group at 6 weeks.

DISCUSSION

In this 6-week randomized trial comparing psilocybin with escitalopram in patients with longstanding, mild-to-severe depression, the change in depression scores on the QIDS-SR-16 at week 6 (the primary outcome) did not differ significantly between the trial groups. Secondary outcomes generally favored psilocybin over escitalopram; however, the confidence intervals for the between-group differences were not adjusted for multiple comparisons, and no conclusions can be drawn from these data. In both trial groups, the scores on the depression scales at week 6 were numerically lower than the baseline scores, but the absence of a placebo group in the trial limits conclusions about the effect of either agent alone. The incidence of adverse events was similar in the trial groups, and no serious adverse events occurred. The percentages of patients who had anxiety, dry mouth, sexual dysfunction, or reduced emotional responsiveness were higher in the escitalopram group than in the psilocybin group.[19] Four patients in the escitalopram group stopped taking their daily capsules entirely, and 1 patient halved the dose because of perceived adverse events. No patient in the psilocybin group requested to cancel the second psilocybin dose. Three patients were unable to attend sessions to receive the second psilocybin dose owing to the Covid-19 lockdown (2 patients in the psilocybin group and 1 in the escitalopram group). The most common adverse event in the psilocybin group was transient headache reported within 24 hours after the day of psilocybin dosing. The incidence of headache was similar to those reported in previous studies of psilocybin.[9,10,13,20]

Acute subjective effects of psilocybin relating to the psychedelic experience were not included as adverse events in our trial, because previous studies have suggested that they may have a mediating influence on positive outcomes.[21] The altered quality of conscious experience typically induced by a 25-mg dose of psilocybin adds complexity to this treatment model, because it requires that psychological support be provided to patients during and after treatment sessions.[15] This requirement informed this trial's screening

Table 2. Primary and Secondary Outcomes.*

Outcome	Psilocybin (N=30)	Escitalopram (N=29)	Difference (95% CI)†
Primary			
Change in QIDS-SR-16 score at 6 wk — points	−8.0±1.0	−6.0±1.0	−2.0 (−5.0 to 0.9)‡
Secondary			
Depression-related outcomes			
Change in QIDS-SR-14 score from the day before to the day after dosing-day 1 — points	−5.7±0.9	−2.8±0.9	−3.0 (−5.5 to −0.4)
QIDS-SR-16 response at 6 wk — no. (%)§	21 (70)	14 (48)	22 (−3 to 48)
QIDS-SR-16 remission at 6 wk — no. (%)¶	17 (57)	8 (28)	28 (2 to 54)
Change in HAM-D-17 score at 6 wk — points	−10.5±1.0	−5.1±1.0	−5.3 (−8.2 to −2.4)
Change in MADRS score at 6 wk — points	−14.4±1.7	−7.2±1.7	−7.2 (−12.1 to −2.4)
Change in BDI-1A score at 6 wk — points	−18.4 (−22.6 to −13.8)	−10.8 (−14.3 to −7.3)	−7.6 (−13.3 to −1.8)
Change in WEMWBS score at 6 wk — points	15.4±1.9	7.3±1.9	8.1 (2.6 to 13.5)
Change in FS score at 6 wk — points	14.4±1.7	9.0±1.7	5.4 (0.5 to 10.3)
Change in STAI score at 6 wk — points	−17.6±2.2	−8.5±2.2	−9.0 (−15.2 to −2.8)
Change in BEAQ score at 6 wk — points	−10.5±2.2	−1.0±2.3	−9.5 (−15.9 to −3.1)
Change in WSAS score at 6 wk — points	−9.7±1.7	−3.8±1.7	−5.8 (−10.7 to −1.0)
Change in SHAPS score at 6 wk — points	−4.7±0.6	−2.5±0.6	−2.2 (−3.8 to −0.6)
Change in SIDAS score at 6 wk — points	−2.0 (−4.3 to 0.0)	−0.8 (−3.4 to 2.0)	−1.3 (−6.5 to −0.3)
PRSexDQ score at 6 wk	0 (0 to 0)	3 (0 to 7)	−2 (−4 to 0)
LEIS score at 6 wk	4.1±0.9	−2.2±1.0	6.3 (3.6 to 9.0)

* Changes in scores represent the mean change from baseline and are reported as mean ±SE, except for the changes in the BDI-1A and Suicidal Ideation Attributes Scale (SIDAS) scores, which are reported as mean (95% confidence interval). The PRSexDQ score at 6 weeks is reported as mean ±SE, and the LEIS score at 6 weeks is reported as mean (95% confidence interval). Scores range from 0 to 60 on the Montgomery and Asberg Depression Rating Scale (MADRS), from 20 to 80 on the Spielberger's Trait Anxiety Inventory (STAI), from 15 to 90 on the Brief Experiential Avoidance Questionnaire (BEAQ), from 0 to 40 on the Work and Social Adjustment Scale (WSAS), from 0 to 14 on the Snaith Hamilton Anhedonia Pleasure Scale (SHAPS), and from 0 to 50 on the SIDAS; greater reductions from baseline on all of these scales indicate greater reductions in symptom severity or impairment. Scores on the Psychotropic-Related Sexual Dysfunction Questionnaire (PRSexDQ) range from 0 to 15, with higher scores indicating greater dysfunction. Scores ranges from 14 to 70 on the Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS) and from 8 to 56 on the Flourishing Scale (FS) range; greater increases from baseline on these scales indicate greater improvements. Scores on the Laukes Emotional Intensity Scale (LEIS) range from −34 to +34, with positive scores indicating an increased intensity of emotional responsiveness and negative scores a reduced intensity of emotional responsiveness. The analysis of each efficacy outcome was generated from statistical models, as described in the statistical analysis plan, available in the protocol. All values shown were adjusted for the baseline value. Unadjusted values are provided in Table S12 in the Supplementary Appendix.

† The confidence intervals for the secondary outcomes have not been corrected for multiple comparisons, and no clinical conclusions can be drawn from these data.

‡ P=0.17.

§ A QIDS-SR-16 response was defined as a reduction in score of more than 50% from baseline. The difference between the groups is expressed as percentage points.

¶ QIDS-SR-16 remission was defined as a score of 5 or lower at week 6. The difference between the groups is expressed as percentage points.

criteria that excluded patients with preexisting psychiatric conditions believed to be incompatible with the limited psychological support that could be made available within the trial. This exclusion criterion may have biased the trial sample toward patients who could receive psilocybin without unacceptable side effects. However, psychological support was provided for both groups in this trial, and it is possible that the adjunctive psychological support improved outcomes among those in the escitalopram group.

A limitation of the trial is the brief duration of escitalopram treatment, because this drug has a delayed therapeutic action on depression.[22] Had the course of escitalopram been extended, it is possible that better efficacy would have been observed among the patients in the escitalopram group. Patients who received the 25-mg dose of

Table 3. Adverse Events Reported during the 6-Week Trial Period and on Dosing-Day 1.*

Event	6-Wk Trial Period		Dosing-Day 1	
	Psilocybin (N=30)	Escitalopram (N=29)	Psilocybin (N=30)	Escitalopram (N=29)
	number of patients (percent)			
Any adverse event	26 (87)	24 (83)	15 (50)	8 (28)
Serious adverse event	0	0	0	0
Related adverse event†	22 (73)	23 (79)	15 (50)	6 (21)
Adverse event reported in ≥3 patients during the full trial period				
Headache	20 (67)	15 (52)	13 (43)	5 (17)
Nausea	8 (27)	9 (31)	4 (13)	0
Fatigue	2 (7)	7 (24)	0	0
Anxiety	0	4 (14)	0	0
Dry mouth	0	4 (14)	0	0
Migraine	3 (10)	1 (3)	0	0
Palpitations	1 (3)	3 (10)	0	0
Sleep disorder	1 (3)	3 (10)	0	0
Diarrhea	1 (3)	2 (7)	0	0
Feeling abnormal	0	3 (10)	0	0
Feeling jittery	2 (7)	1 (3)	0	0
Vomiting	2 (7)	1 (3)	0	0

* These were the most prevalent adverse events that were reported during the trial.
† Whether an adverse event was related to the therapeutic intervention was determined by the study clinician through dialogue with each patient. Events deemed "probably" or "definitely" related were counted.

psilocybin rated the intensity of acute subjective effects higher than patients who received the 1-mg dose (Fig. S7). We did not assess the effectiveness of blinding within each treatment group. It was assumed that the active comparator design would mitigate expectancy bias, but we cannot be confident that guessing of the trial-group assignment or biased expectations in favor of psilocybin did not influence the results. Although efforts were made to recruit patients by external referrals, most of the recruited volunteers referred themselves, and many expressed a preference for psilocybin over escitalopram. This created a selected trial population and limits generalization of the results.

The patients in the trial were not from diverse ethnic or socioeconomic backgrounds. Strategies to improve recruitment of more diverse study populations are needed in studies of psilocybin for depression. Also, average symptom severity scores at baseline were in the range for moderate depression, thus limiting extrapolations to patients with severe depressive symptoms or treatment-resistant depression.

Psychedelic agents have been shown to enhance suggestibility,[23] and their psychological effects are assumed to be context-dependent.[24,25] In other words, the content and subjective quality of the psychedelic experience is influenced by a person's memories, perceptions, and degree to which the environment is supportive at the time of administration of the agent. In a study in which various psychedelic compounds were administered to rats, the compounds were shown to increase dendritic arbor complexity, promote dendritic spine growth, and stimulate synapse formation in the rat cortex, mediated by serotonin 5-HT_{2A} receptor agonism,[26] all of which are forms of neuronal plasticity that may be related to the principle that responses to psychedelics are especially dependent on contextual conditions.[24,25]

This trial comparing psilocybin with escitalopram in a selected group of patients showed that

the change in scores for depression at 6 weeks did not differ significantly between the trial groups. Secondary outcomes mostly favored psilocybin over escitalopram, but the confidence intervals for the between-group differences were not adjusted for multiple comparisons. Larger and longer trials are needed to compare psilocybin with established treatments for depression.

The views expressed are those of the authors and not necessarily those of the National Health Service, the National Institute for Health Research (NIHR), or the Department of Health and Social Care.

Supported by a private donation from the Alexander Mosley Charitable Trust and by the founding partners of Imperial College London's Centre for Psychedelic Research. Infrastructure support was provided by the NIHR Imperial Biomedical Research Centre and the NIHR Imperial Clinical Research Facility.

Dr. Carhart-Harris reports receiving consulting fees from COMPASS Pathways, Entheon Biomedical, Mydecine, Synthesis Institute, Tryp Therapeutics, and Usona Institute; Dr. Giribaldi, receiving consulting fees from SmallPharma; Dr. Watts, receiving advisory board fees from Usona Institute and being employed by Synthesis Institute; Dr. Baker-Jones, receiving fees for facilitating meetings from Synthesis Institute; Dr. Erritzoe, receiving consulting fees from Field Trip and Mydecine; and Dr. Nutt, receiving consulting fees from Awakn, H. Lundbeck, and Psyched Wellness, advisory board fees from COMPASS Pathways, and lecture fees from Takeda Medical Research Foundation and owning stock in Alcarelle. No other potential conflict of interest relevant to this article was reported.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

A data sharing statement provided by the authors is available with the full text of this article at NEJM.org.

We thank Ms. Renee Harvey, Dr. Graham Campbell, Dr. Benjamin Waterhouse, Dr. Frederico Magalhães, Mr. James Close, Dr. Leor Roseman, Ms. Hilary Platt, Mr. Gregory Donaldson, and Dr. Chris Timmermann for their voluntary assistant guide roles; Dr. Meg Spriggs and Ms. Laura Kärtner for their support with data collection; Dr. Louise Paterson and Dr. Robin Tyacke for their assistance with randomization, blinding, and drug accountability; Ms. Ghazel Mukhtar for her assistance with administrative tasks; and Dr. Tim Read for his clinical supervision and guidance.

REFERENCES

1. Malhi GS, Mann JJ. Depression. Lancet 2018;392:2299-312.
2. Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. Lancet 2018;391:1357-66.
3. Cipriani A, Santilli C, Furukawa TA, et al. Escitalopram versus other antidepressive agents for depression. Cochrane Database Syst Rev 2009;2:CD006532.
4. Nichols DE. Psychedelics. Pharmacol Rev 2016;68:264-355.
5. Nutt D, Erritzoe D, Carhart-Harris R. Psychedelic psychiatry's brave new world. Cell 2020;181:24-8.
6. Madsen MK, Fisher PM, Burmester D, et al. Psychedelic effects of psilocybin correlate with serotonin 2A receptor occupancy and plasma psilocin levels. Neuropsychopharmacology 2019;44:1328-34.
7. Rucker JJH, Iliff J, Nutt DJ. Psychiatry & the psychedelic drugs: past, present & future. Neuropharmacology 2018;142:200-18.
8. Carhart-Harris RL, Goodwin GM. The therapeutic potential of psychedelic drugs: past, present, and future. Neuropsychopharmacology 2017;42:2105-13.
9. Carhart-Harris RL, Bolstridge M, Rucker J, et al. Psilocybin with psychological support for treatment-resistant depression: an open-label feasibility study. Lancet Psychiatry 2016;3:619-27.
10. Griffiths RR, Johnson MW, Carducci MA, et al. Psilocybin produces substantial and sustained decreases in depression and anxiety in patients with life-threatening cancer: a randomized double-blind trial. J Psychopharmacol 2016;30:1181-97.
11. Ross S, Bossis A, Guss J, et al. Rapid and sustained symptom reduction following psilocybin treatment for anxiety and depression in patients with life-threatening cancer: a randomized controlled trial. J Psychopharmacol 2016;30:1165-80.
12. Grob CS, Danforth AL, Chopra GS, et al. Pilot study of psilocybin treatment for anxiety in patients with advanced-stage cancer. Arch Gen Psychiatry 2011;68:71-8.
13. Davis AK, Barrett FS, May DG, et al. Effects of psilocybin-assisted therapy on major depressive disorder: a randomized clinical trial. JAMA Psychiatry 2020 November 4 (Epub ahead of print).
14. Carhart-Harris RL, Bolstridge M, Day CMJ, et al. Psilocybin with psychological support for treatment-resistant depression: six-month follow-up. Psychopharmacology (Berl) 2018;235:399-408.
15. Johnson M, Richards W, Griffiths R. Human hallucinogen research: guidelines for safety. J Psychopharmacol 2008;22:603-20.
16. Chawla N, Ostafin B. Experiential avoidance as a functional dimensional approach to psychopathology: an empirical review. J Clin Psychol 2007;63:871-90.
17. Opbroek A, Delgado PL, Laukes C, et al. Emotional blunting associated with SSRI-induced sexual dysfunction: do SSRIs inhibit emotional responses? Int J Neuropsychopharmacol 2002;5:147-51.
18. Roseman L, Haijen E, Idialu-Ikato K, Kaelen M, Watts R, Carhart-Harris R. Emotional breakthrough and psychedelics: validation of the Emotional Breakthrough Inventory. J Psychopharmacol 2019;33:1076-87.
19. Ng CG, Wong SK, Loh HS, Yee A. Anhedonia among patients with major depressive disorder: a comparison between patients on escitalopram and healthy controls. Clin Ter 2014;165(6):e384-e390.
20. Johnson MW, Sewell RA, Griffiths RR. Psilocybin dose-dependently causes delayed, transient headaches in healthy volunteers. Drug Alcohol Depend 2012;123:132-40.
21. Roseman L, Nutt DJ, Carhart-Harris RL. Quality of acute psychedelic experience predicts therapeutic efficacy of psilocybin for treatment-resistant depression. Front Pharmacol 2018;8:974.
22. Trivedi MH, Rush AJ, Wisniewski SR, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. Am J Psychiatry 2006;163:28-40.
23. Carhart-Harris RL, Kaelen M, Whalley MG, Bolstridge M, Feilding A, Nutt DJ. LSD enhances suggestibility in healthy volunteers. Psychopharmacology (Berl) 2015;232:785-94.
24. Carhart-Harris RL, Roseman L, Haijen E, et al. Psychedelics and the essential importance of context. J Psychopharmacol 2018;32:725-31.
25. Brouwer A, Carhart-Harris RL. Pivotal mental states. J Psychopharmacol 2020 November 11 (Epub ahead of print).
26. Ly C, Greb AC, Cameron LP, et al. Psychedelics promote structural and functional neural plasticity. Cell Rep 2018;23:3170-82.

ORIGINAL ARTICLE

BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Mass Vaccination Setting

Noa Dagan, M.D., Noam Barda, M.D., Eldad Kepten, Ph.D., Oren Miron, M.A., Shay Perchik, M.A., Mark A. Katz, M.D., Miguel A. Hernán, M.D., Marc Lipsitch, D.Phil., Ben Reis, Ph.D., and Ran D. Balicer, M.D.

ABSTRACT

BACKGROUND

As mass vaccination campaigns against coronavirus disease 2019 (Covid-19) commence worldwide, vaccine effectiveness needs to be assessed for a range of outcomes across diverse populations in a noncontrolled setting. In this study, data from Israel's largest health care organization were used to evaluate the effectiveness of the BNT162b2 mRNA vaccine.

METHODS

All persons who were newly vaccinated during the period from December 20, 2020, to February 1, 2021, were matched to unvaccinated controls in a 1:1 ratio according to demographic and clinical characteristics. Study outcomes included documented infection with the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), symptomatic Covid-19, Covid-19–related hospitalization, severe illness, and death. We estimated vaccine effectiveness for each outcome as one minus the risk ratio, using the Kaplan–Meier estimator.

RESULTS

Each study group included 596,618 persons. Estimated vaccine effectiveness for the study outcomes at days 14 through 20 after the first dose and at 7 or more days after the second dose was as follows: for documented infection, 46% (95% confidence interval [CI], 40 to 51) and 92% (95% CI, 88 to 95); for symptomatic Covid-19, 57% (95% CI, 50 to 63) and 94% (95% CI, 87 to 98); for hospitalization, 74% (95% CI, 56 to 86) and 87% (95% CI, 55 to 100); and for severe disease, 62% (95% CI, 39 to 80) and 92% (95% CI, 75 to 100), respectively. Estimated effectiveness in preventing death from Covid-19 was 72% (95% CI, 19 to 100) for days 14 through 20 after the first dose. Estimated effectiveness in specific subpopulations assessed for documented infection and symptomatic Covid-19 was consistent across age groups, with potentially slightly lower effectiveness in persons with multiple coexisting conditions.

CONCLUSIONS

This study in a nationwide mass vaccination setting suggests that the BNT162b2 mRNA vaccine is effective for a wide range of Covid-19–related outcomes, a finding consistent with that of the randomized trial.

From the Clalit Research Institute, Innovation Division, Clalit Health Services, Tel Aviv (N.D., N.B., E.K., O.M., S.P., M.A.K., R.D.B.), and the School of Public Health, Faculty of Health Sciences, Ben Gurion University of the Negev, Be'er Sheva (O.M., M.A.K., R.D.B.) — both in Israel; University of Michigan School of Public Health, Ann Arbor (M.A.K.); and the Department of Biomedical Informatics (N.D., N.B.), Harvard Medical School (B.R.), the Departments of Epidemiology and Biostatistics (M.A.H.), and the Center for Communicable Disease Dynamics, Departments of Epidemiology and of Immunology and Infectious Diseases (M.L.), Harvard T.H. Chan School of Public Health, Harvard–MIT Division of Health Sciences and Technology (M.A.H.), and the Predictive Medicine Group, Computational Health Informatics Program, Boston Children's Hospital (B.R.) — all in Boston. Address reprint requests to Dr. Balicer at the Clalit Research Institute, Innovation Division, Clalit Health Services, Tel Aviv, Israel, or at Rbalicer@clalit.org.il.

Drs. Dagan and Barda contributed equally to this article.

This article was published on February 24, 2021, at NEJM.org.

N Engl J Med 2021;384:1412-23.
DOI: 10.1056/NEJMoa2101765

Mass vaccination campaigns using newly approved vaccines against the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)[1,2] are beginning in many parts of the world. Randomized clinical trials of mRNA-based vaccines reported efficacies for preventing coronavirus 2019 (Covid-19) in the range of 94%[2] to 95%.[1]

Although randomized clinical trials are considered the "gold standard" for evaluating intervention effects, they have notable limitations of sample size and subgroup analysis, restrictive inclusion criteria, and a highly controlled setting that may not be replicated in a mass vaccine rollout. For example, the phase 3 trial of the BNT162b2 mRNA vaccine against Covid-19 included 21,720 persons who were randomly assigned to the vaccinated group, which permitted estimates of vaccine efficacy in only a small number of subpopulations.[1] Moreover, patients with chronic diseases were included only if the conditions were deemed stable by the investigators.[3] It is also important to see whether in a scaled-up vaccination program such factors as suboptimal adherence to vaccination schedules and vaccine-handling logistics influence vaccine effectiveness. Postauthorization analyses can thus meet the urgent need to evaluate the effectiveness of Covid-19 vaccines across diverse populations with a wide range of coexisting conditions, in the midst of imperfect adherence to vaccination protocols and the challenges of cold-chain maintenance and vaccine-deployment logistics.

We leveraged the integrated data repositories of Israel's largest health care organization to evaluate Covid-19 vaccine effectiveness for five outcomes: documented SARS-CoV-2 infection, symptomatic Covid-19, hospitalization, severe illness, and death. Using this observational data set, we evaluated the effectiveness over time and in subpopulations defined by age, sex, and coexisting conditions.

Methods

Study Population

We analyzed data from Clalit Health Services (CHS), the largest of four integrated health care organizations in Israel, which insures 4.7 million patients (53% of the population). A description of the CHS data repositories used for this study is provided in the Supplementary Appendix (Supplementary Methods 1), available with the full text of this article at NEJM.org. Information on authors' access to these repositories as well as authors' contributions to the study is provided in Supplementary Methods 2. This study was approved by the CHS institutional review board. The study was exempt from the requirement for informed consent.

Study Design

We designed this observational study to emulate a target trial of the causal effect of the BNT162b2 vaccine on Covid-19 outcomes.[4] Eligibility criteria included an age of 16 years or older, not having a previously documented positive SARS-CoV-2 polymerase-chain-reaction (PCR) test, and being a member of the health care organization during the previous 12 months.

Population groups in which internal variability in the probability of exposure or the outcomes is high and controlling for the high variability is not feasible (e.g., high variability in infection risk among patient-facing health care workers in dedicated Covid-19 wards as compared with administrative staff) were excluded. Such population groups are persons not having a documented geostatistical living area, those who have had interactions with the health care system during the preceding 3 days that may indicate the start of symptomatic disease and may preclude vaccination, nursing home residents, persons medically confined to the home, or health care workers.

Each day during the period from December 20, 2020, to February 1, 2021, all newly vaccinated persons were matched in a 1:1 ratio to unvaccinated controls. For each person, follow-up ended at the earliest of the following events: occurrence of an outcome event, death unrelated to Covid-19, vaccination (for unvaccinated controls), vaccination of the matched control (for vaccinated persons), or the end of the study period. Newly vaccinated persons were eligible for inclusion in the study, even if they had previously been selected as a control.

We matched vaccine recipients and controls on variables associated with the probability of both vaccination and infection or severity of Covid-19: age, sex, sector (general Jewish, Arab, or ultra-Orthodox Jewish), neighborhood of residence (since disease activity and vaccination

uptake vary greatly across defined geostatistical areas), history of influenza vaccination during the preceding 5 years (0, 1 or 2, 3 or 4, or ≥5 vaccinations), pregnancy (a potential risk factor for severe Covid-19[5] and associated with the rate of vaccination owing to evolving vaccination guidelines for pregnant women), and the total number of coexisting conditions that had been identified by the Centers for Disease Control and Prevention (CDC) as risk factors for severe Covid-19 as of December 20, 2020.[6,7] (See Supplementary Methods 3 for additional information about the matching process. The protocol and statistical analysis plan are available at NEJM.org.)

The five outcomes of interest were documented SARS-CoV-2 infection confirmed by positive PCR test, documented symptomatic Covid-19, hospital admission for Covid-19, severe Covid-19 (according to National Institutes of Health criteria)[8] and death from Covid-19. Each of these outcomes includes the outcomes that follow it. In a supplementary analysis, we also evaluated an additional outcome, SARS-CoV-2 infection without documented symptoms, as an imperfect proxy for asymptomatic infection (since mild symptoms may not be documented).

Table S1 provides details on definitions of variables. Persons with missing data for smoking status or body-mass index (BMI) were dropped from the analysis.

STATISTICAL ANALYSIS

Covariate balance after matching was evaluated with the use of a plot of the mean differences between variable values (standardized for continuous variables) for the vaccinated and unvaccinated groups, with a difference of 0.1 or less considered to be acceptable.[9] Survival curves for the vaccinated and unvaccinated groups were estimated with the Kaplan–Meier estimator.[10] We considered three periods: days 14 through 20 after the first dose of vaccine, days 21 through 27 after the first dose (administration of the second dose was scheduled to occur on day 21 after the first dose), and day 7 after the second dose until the end of the follow-up. For each period, we used the Kaplan–Meier estimator with daily outcome and censoring events to compute the probability ("risk") of the outcome during the period, using matched pairs in which both persons were still at risk at the beginning of the period. We then calculated risk ratios for vaccination as compared with no vaccination and estimated the vaccine effectiveness as one minus the risk ratio. We estimated the vaccine effectiveness only in analyses in which there were more than 10 instances of an outcome across the two groups.

The period immediately after the first dose, when immunity is gradually building,[1] was excluded in the main analyses because the risk ratio is expected to be close to 1 during this period. In secondary analyses, we considered the periods from day 0 through day 20 and day 0 through day 27, to avoid a potential selection bias in the main analyses that were restricted to persons whose data remained uncensored at the beginning of each period (see Supplementary Methods 4).[11-13] We also conducted a sensitivity analysis in the 6 days after the second dose of vaccine among those who received a second dose. A further sensitivity analysis estimated the hazard ratio each day for the documented SARS-CoV-2 infection outcome.

We performed an additional sensitivity analysis to assess the potential for selection bias due to informative censoring. In this analysis, data on controls who were subsequently vaccinated were censored only after 7 days (i.e., after the period with little or no vaccine effect) plus the median time from documented Covid-19 diagnosis to the outcome being studied.

We calculated 95% confidence intervals using the percentile bootstrap method with 500 repetitions. Analyses were performed with the use of R software, version 4.0.2.

RESULTS

STUDY POPULATION

Of 1,503,216 CHS members who were vaccinated, 1,163,534 were eligible for the study and 596,618 were matched to unvaccinated controls (Fig. 1). Matched persons were younger than the eligible population overall and had a lower prevalence of chronic conditions because there was a smaller pool of potential unvaccinated matches for older vaccine recipients, owing to high vaccination rates in the older population (Table S2 and Fig. S1). The baseline characteristics of the matched persons are shown in Table 1. All variables were well balanced between the study groups (Fig. S2). About 0.6% of persons with missing data on smoking status or body-mass index were dropped from the analysis (Fig. 1). Data for 44% of the unvaccinated controls and

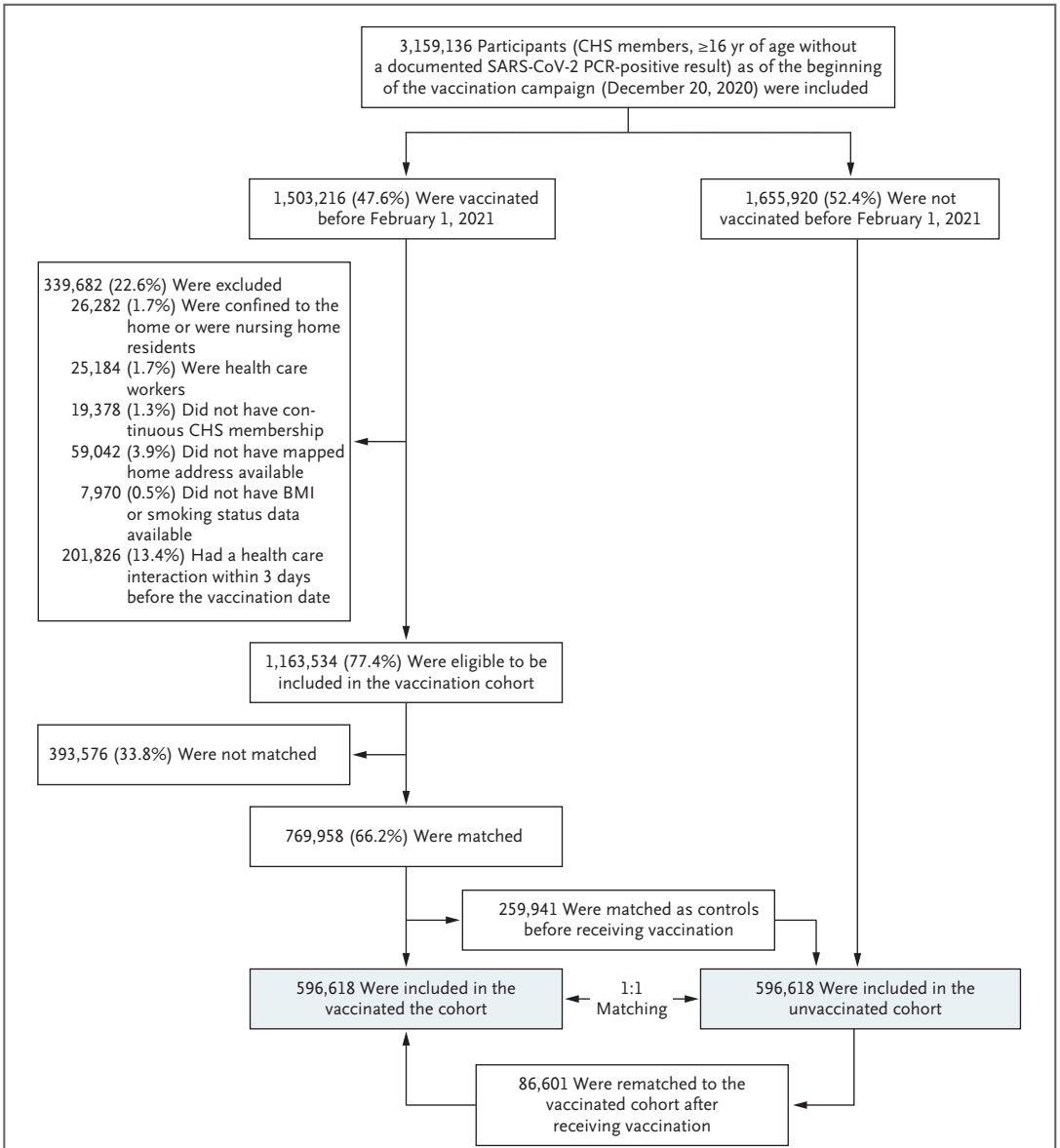


Figure 1. Study Population and Cohort Enrollment Process, December 20, 2020, to February 1, 2021.

The 1,503,216 persons vaccinated before February 1, 2021, were also required to be without a documented SARS-CoV-2 PCR-positive result before the vaccination date. Absolute numbers and percentage changes are shown for each inclusion and exclusion criterion. The exclusion process was gradual and occurred in phases; persons could have had more than one reason for exclusion. The same exclusion criteria were applied to the unvaccinated persons for each index date in which they were considered for matching. The chart focuses on the vaccinated population. CHS denotes Clalit Health Services.

their matched pairs were censored when the controls received the vaccine.

Vaccine Effectiveness

During a mean follow-up of 15 days (interquartile range, 5 to 25), 10,561 infections were documented (0.6 infections per 1000 person-days), of which 5996 (57%) were symptomatic Covid-19 illness, 369 required hospitalization, 229 were severe cases of Covid-19, and 41 resulted in death. Hospitalizations, severe disease, and death occurred at increasing time spans from diagnosis (median times, 1, 5, and 11 days, respectively; see Fig. S3). Of persons who had 21 or more days

Table 1. Demographic and Clinical Characteristics of Vaccinated Persons and Unvaccinated Controls at Baseline.*

Characteristics	Unvaccinated Controls (N=596,618)	Vaccinated Persons (N=596,618)
Median age (IQR) — yr	45 (35–62)	45 (35–62)
Age group — no. (%)		
16 to 39 yr	213,090 (35.7)	213,090 (35.7)
40 to 49 yr	130,752 (21.9)	130,752 (21.9)
50 to 59 yr	85,609 (14.3)	85,609 (14.3)
60 to 69 yr	88,153 (14.8)	88,153 (14.8)
70 to 79 yr	56,946 (9.5)	56,946 (9.5)
≥80 yr	22,068 (3.7)	22,068 (3.7)
Sex — no. (%)		
Female	298,059 (50.0)	298,059 (50.0)
Male	298,559 (50.0)	298,559 (50.0)
Population sector — no. (%)		
General Jewish	463,234 (77.6)	463,234 (77.6)
Arab	120,896 (20.3)	120,896 (20.3)
Ultra-Orthodox Jewish	12,488 (2.1)	12,488 (2.1)
No. of risk factors according to CDC criteria — no. (%)		
0	338,384 (56.7)	338,384 (56.7)
1	140,779 (23.6)	140,779 (23.6)
2	55,766 (9.3)	55,766 (9.3)
3	29,273 (4.9)	29,273 (4.9)
≥4	32,416 (5.4)	32,416 (5.4)
No. of influenza vaccinations during preceding 5 yr — no. (%)		
0	351,141 (58.9)	351,141 (58.9)
1 or 2	116,200 (19.5)	116,200 (19.5)
3 or 4	50,441 (8.5)	50,441 (8.5)
≥5	78,836 (13.2)	78,836 (13.2)
CDC "certain" risk criteria — no. of persons (%)		
Cancer	11,946 (2.0)	11,595 (1.9)
Chronic kidney disease	40,568 (6.8)	40,587 (6.8)
Chronic obstructive pulmonary disease	12,667 (2.1)	11,131 (1.9)
Heart disease	39,165 (6.6)	38,913 (6.5)
Solid-organ transplantation	495 (0.1)	435 (0.1)
Obesity: BMI, 30 to 40	100,584 (16.9)	105,476 (17.7)
Severe obesity: BMI, ≥40	9,856 (1.7)	8,920 (1.5)
Pregnancy	1,508 (0.3)	1,508 (0.3)
Sickle cell disease	98 (<0.1)	109 (<0.1)
Smoking	118,733 (19.9)	97,881 (16.4)
Type 2 diabetes mellitus	66,198 (11.1)	65,343 (11.0)
CDC "possible" risk criteria — no. of persons (%)		
Asthma	32,114 (5.4)	29,814 (5.0)

Table 1. (Continued.)

Characteristics	Unvaccinated Controls (N=596,618)	Vaccinated Persons (N=596,618)
Cerebrovascular disease	18,392 (3.1)	17,792 (3.0)
Other respiratory disease	2,198 (0.4)	2,014 (0.3)
Hypertension	101,017 (16.9)	103,028 (17.3)
Immunosuppression	15,823 (2.7)	16,180 (2.7)
Neurologic disease	25,897 (4.3)	24,111 (4.0)
Liver disease	11,109 (1.9)	9,699 (1.6)
Overweight: BMI, 25 to 30	203,296 (34.1)	212,778 (35.7)
Thalassemia	3,764 (0.6)	3,967 (0.7)
Type 1 diabetes mellitus	2,309 (0.4)	2,406 (0.4)

* The 86,601 persons who were first recruited as unvaccinated controls and then, after vaccination, were re-recruited as vaccinated persons appear in both groups. BMI denotes body-mass index (the weight in kilograms divided by the square of the height in meters), CDC Centers for Disease Control and Prevention, and IQR interquartile range.

of follow-up, 96% received a second dose of vaccine (95% of whom received it before day 24).

Figure 2 shows the cumulative incidence curves for the included outcomes, and Table 2 shows the estimated vaccine effectiveness for the main outcomes and time periods. During the period from 14 to 20 days after the first dose, the estimated vaccine effectiveness for documented infection was 46% (95% confidence interval [CI], 40 to 51); symptomatic Covid-19 illness, 57% (95% CI, 50 to 63); hospitalization, 74% (95% CI, 56 to 86); severe illness, 62% (95% CI, 39 to 80); and death, 72% (95% CI, 19 to 100). During the period from 21 to 27 days after the first dose, the estimated effectiveness for these outcomes was 60% (95% CI, 53 to 66), 66% (95% CI, 57 to 73), 78% (95% CI, 61 to 91), 80% (95% CI, 59 to 94), and 84% (95% CI, 44 to 100), respectively. In the follow-up period starting 7 days after the second dose, the vaccine effectiveness for documented infections, symptomatic illness, hospitalization, and severe disease was 92% (95% CI, 88 to 95), 94% (95% CI, 87 to 98), 87% (95% CI, 55 to 100), and 92% (95% CI, 75 to 100), respectively. The daily value for one minus the hazard ratio for the documented infection outcome is included in Figure S4; it is consistent with a gradual daily increase in vaccine effectiveness.

Table 3 shows the estimated vaccine effectiveness for documented SARS-CoV-2 infection and Covid-19 outcomes in subpopulations defined by age, sex, and coexisting conditions. The estimates are consistent with similar effectiveness across age groups and slightly lower effectiveness among patients with multiple coexisting conditions.

The estimated vaccine effectiveness for the asymptomatic infection proxy was 29% (95% CI, 17 to 39) during the period from 14 to 20 days after the first dose, 52% (95% CI, 41 to 60) 21 to 27 days after the first dose, and 90% (95% CI, 83 to 94) 7 or more days after the second dose (Table S3 and Fig. S5).

Figure S6 shows a magnification of the cumulative incidence curve for the symptomatic illness outcome, showing the divergence of the curves starting around day 12. This is shown in comparison with the same curve from an analysis minimally matched (on age and sex only) that shows an earlier and wider separation of the curves.

Table S4 shows the sensitivity analyses of vaccine effectiveness across additional follow-up periods. Cumulative effectiveness estimates starting from day 0 were lower across all outcomes. Effectiveness estimates conditional on receipt of the second dose of vaccine were higher than unconditional estimates for days 21 through 27 after the first dose.

Table S5 and Figure S7 show the results of the sensitivity analysis in which data for persons who were enrolled as controls and were then vaccinated were censored at a delay (a number of days after the vaccination date, depending on the outcome). The estimates are similar

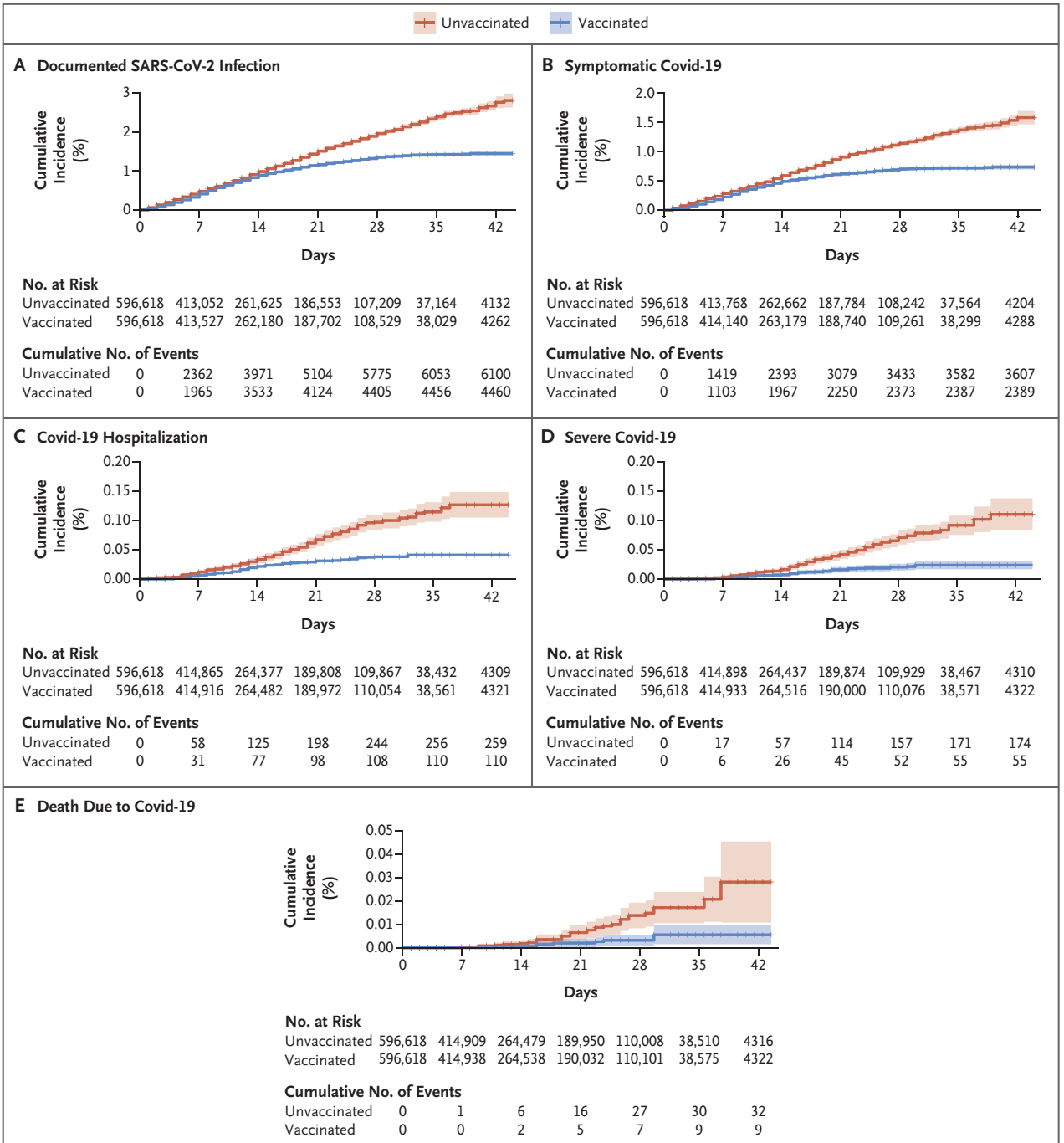


Figure 2. Cumulative Incidence of the Five Outcomes.

Cumulative incidence curves (1 minus the Kaplan–Meier risk) for the various outcomes are shown, starting from the day of administration of the first dose of vaccine. Shaded areas represent 95% confidence intervals. The number at risk at each time point and the cumulative number of events are also shown for each outcome. Graphs in which all data are shown with a y axis scale from 0 to 100 (along with the data shown, as here, on an expanded y axis) are provided in Figure S8 in the Supplementary Appendix.

to those of the main analysis. Table S6 details all analyses performed during the study, and Table S7 includes the life tables for the various outcomes.

DISCUSSION

This study evaluates the effectiveness of the novel BNT162b2 mRNA vaccine[1] against Covid-19 in a nationwide mass vaccination setting. Estimated vaccine effectiveness during the follow-up period starting 7 days after the second dose was 92% for documented infection, 94% for symptomatic Covid-19, 87% for hospitalization, and 92% for severe Covid-19. Estimated effectiveness during days 14 through 20 (after one dose) and days 21 through 27 (gradual shifting between the first and second vaccine doses) was 46% and 60% for documented infection, 57% and 66% for symptomatic Covid-19, 74% and 78% for hospitalization, 62% and 80% for severe Covid-19, and 72% and 84% for Covid-19–related death, respectively.

The first primary end point evaluated in the randomized trial of the BNT162b2 vaccine was symptomatic Covid-19. In both the randomized trial and our study, the cumulative incidence of symptomatic Covid-19 in the vaccinated and unvaccinated groups began to diverge around day 12 after the first dose.[1] The estimated vaccine efficacy for symptomatic Covid-19 starting at day 7 after the second dose was 95% in the randomized trial, as compared with 94% in our study. The estimated efficacy between the first dose and the second dose was 52% in the trial, as compared with 29% in our study. This difference may reflect the high level of transmission in Israel during the study period,[14] which affected both the vaccinated persons and the controls equally during the first 12 days after administration of the first dose. To eliminate this distortion, we estimated first-dose effectiveness of the vaccine against Covid-19 for the period from days 14 through 20; the estimated effectiveness was 57%.

The estimated effectiveness for documented infection during days 14 through 20 was 46% in our study. A relatively similar effectiveness of 51% was reported by Chodick et al.,[15] who evaluated a cohort from another health care organization in Israel and used a different study design that compared infection among vaccinated persons at days 13 through 24 after the first dose against infection during days 0 through 12.

In the randomized trial, the estimated vaccine efficacy for severe Covid-19 (89% over the entire study period) was based on only 10 cases. Our study recorded 229 cases of severe Covid-19 — 55 in the vaccinated group and 174 in the

Table 2. Estimated Vaccine Effectiveness against Covid-19 Outcomes during Three Time Periods.*

Period	Documented Infection		Symptomatic Illness		Hospitalization		Severe Disease		Death	
	1–RR	Risk Difference	1–RR	Risk Difference	1–RR	Risk Difference	1–RR	Risk Difference	1–RR	Risk Difference
	% (95% CI)	*no./1000 persons (95% CI)*	% (95% CI)	*no./1000 persons (95% CI)*	% (95% CI)	*no./1000 persons (95% CI)*	% (95% CI)	*no./1000 persons (95% CI)*	% (95% CI)	*no./1000 persons (95% CI)*
14 to 20 days after first dose	46 (40–51)	2.06 (1.70–2.40)	57 (50–63)	1.54 (1.28–1.80)	74 (56–86)	0.21 (0.13–0.29)	62 (39–80)	0.14 (0.07–0.21)	72 (19–100)	0.03 (0.01–0.07)
21 to 27 days after first dose	60 (53–66)	2.31 (1.96–2.69)	66 (57–73)	1.34 (1.09–1.62)	78 (61–91)	0.22 (0.13–0.31)	80 (59–94)	0.18 (0.10–0.27)	84 (44–100)	0.06 (0.02–0.11)
7 days after second dose to end of follow-up	92 (88–95)	8.58 (6.22–11.18)	94 (87–98)	4.61 (3.29–6.53)	87 (55–100)	0.22 (0.08–0.39)	92 (75–100)	0.32 (0.13–0.52)	NA	NA

* Confidence intervals were estimated using the percentile bootstrap method with 500 repetitions. Estimates were calculated only for cells with more than 10 instances of an outcome across the two groups. NA denotes not available, and RR risk ratio.

Table 3. Estimated Vaccine Effectiveness against Covid-19 Outcomes in Subpopulations According to Characteristics at Baseline.*

Characteristic and Period	Documented Infection		Symptomatic Illness	
	1–RR	Risk Difference	1–RR	Risk Difference
	% (95% CI)	no./1000 persons (95% CI)	% (95% CI)	no./1000 persons (95% CI)
Male sex				
14 to 20 days after first dose	41 (32 to 50)	1.71 (1.22 to 2.21)	52 (41 to 61)	1.26 (0.90 to 1.62)
21 to 27 days after first dose	57 (48 to 65)	2.25 (1.76 to 2.75)	62 (49 to 72)	1.30 (0.92 to 1.67)
7 days after second dose to end of follow-up	91 (80 to 96)	7.33 (4.48 to 10.84)	88 (71 to 98)	2.90 (1.87 to 4.02)
Female sex				
14 to 20 days after first dose	50 (41 to 57)	2.39 (1.84 to 2.86)	60 (52 to 68)	1.81 (1.43 to 2.19)
21 to 27 days after first dose	63 (55 to 71)	2.38 (1.91 to 2.91)	69 (58 to 78)	1.38 (1.02 to 1.71)
7 days after second dose to end of follow-up	93 (88 to 97)	9.75 (6.84 to 13.48)	96 (90 to 100)	6.22 (3.60 to 9.56)
Age, 16 to 39 yr				
14 to 20 days after first dose	49 (41 to 57)	2.29 (1.74 to 2.88)	57 (46 to 68)	1.38 (0.99 to 1.80)
21 to 27 days after first dose	64 (54 to 72)	2.80 (2.20 to 3.48)	67 (52 to 78)	1.27 (0.89 to 1.73)
7 days after second dose to end of follow-up	94 (87 to 97)	8.72 (5.72 to 12.69)	99 (96 to 100)	4.06 (2.76 to 5.66)
Age, 40 to 69 yr				
14 to 20 days after first dose	47 (40 to 55)	2.13 (1.69 to 2.66)	59 (50 to 67)	1.68 (1.32 to 2.05)
21 to 27 days after first dose	58 (49 to 67)	2.19 (1.67 to 2.70)	65 (53 to 74)	1.38 (1.03 to 1.80)
7 days after second dose to end of follow-up	90 (82 to 95)	8.96 (6.16 to 13.05)	90 (75 to 98)	5.01 (2.53 to 8.67)
Age, ≥70 yr				
14 to 20 days after first dose	22 (–9 to 44)	0.81 (–0.28 to 1.89)	44 (19 to 64)	1.36 (0.48 to 2.36)
21 to 27 days after first dose	50 (19 to 72)	1.40 (0.42 to 2.35)	64 (37 to 83)	1.35 (0.62 to 2.22)
7 days after second dose to end of follow-up	95 (87 to 100)	6.10 (3.43 to 9.61)	98 (90 to 100)	4.77 (2.14 to 7.70)
No coexisting conditions				
14 to 20 days after first dose	49 (42 to 56)	2.13 (1.69 to 2.59)	55 (45 to 63)	1.32 (0.98 to 1.67)
21 to 27 days after first dose	66 (58 to 73)	2.49 (1.99 to 2.98)	73 (62 to 82)	1.27 (0.92 to 1.64)
7 days after second dose to end of follow-up	91 (83 to 96)	7.67 (4.90 to 11.07)	93 (78 to 100)	3.54 (1.79 to 5.90)
One or two coexisting conditions				
14 to 20 days after first dose	43 (32 to 53)	2.05 (1.41 to 2.73)	57 (45 to 66)	1.74 (1.25 to 2.24)
21 to 27 days after first dose	56 (45 to 65)	2.43 (1.77 to 3.16)	62 (47 to 73)	1.56 (1.05 to 2.06)
7 days after second dose to end of follow-up	95 (88 to 98)	10.53 (6.73 to 14.40)	95 (88 to 100)	6.21 (3.82 to 8.95)
Three or more coexisting conditions				
14 to 20 days after first dose	37 (12 to 55)	1.60 (0.43 to 2.76)	62 (43 to 77)	2.19 (1.20 to 3.18)
21 to 27 days after first dose	37 (–1 to 62)	1.03 (–0.03 to 2.02)	47 (11 to 73)	0.97 (0.16 to 1.86)
7 days after second dose to end of follow-up	86 (72 to 95)	5.83 (3.16 to 9.03)	89 (68 to 98)	3.97 (1.41 to 6.68)
Obesity				
14 to 20 days after first dose	49 (32 to 65)	2.50 (1.40 to 3.75)	65 (48 to 79)	2.31 (1.32 to 3.33)
21 to 27 days after first dose	48 (19 to 66)	2.02 (0.69 to 3.25)	50 (11 to 73)	1.25 (0.18 to 2.27)
7 days after second dose to end of follow-up	95 (88 to 100)	12.43 (6.03 to 20.70)	98 (91 to 100)	9.60 (4.03 to 17.39)
Type 2 diabetes mellitus				
14 to 20 days after first dose	25 (–10 to 51)	1.17 (–0.36 to 2.74)	48 (14 to 68)	1.94 (0.49 to 3.28)

Table 3. (Continued.)

Characteristic and Period	Documented Infection		Symptomatic Illness	
	1–RR	Risk Difference	1–RR	Risk Difference
	% (95% CI)	*no./1000 persons (95% CI)*	*% (95% CI)*	*no./1000 persons (95% CI)*
21 to 27 days after first dose	49 (2 to 78)	1.29 (0.04 to 2.67)	60 (10 to 84)	1.18 (0.12 to 2.27)
7 days after second dose to end of follow-up	91 (71 to 100)	6.85 (3.31 to 11.33)	91 (68 to 100)	5.06 (1.84 to 8.96)
Hypertension				
14 to 20 days after first dose	28 (2 to 49)	1.12 (0.08 to 2.26)	45 (16 to 64)	1.33 (0.37 to 2.22)
21 to 27 days after first dose	45 (15 to 66)	1.49 (0.42 to 2.53)	59 (31 to 79)	1.47 (0.60 to 2.39)
7 days after second dose to end of follow-up	93 (85 to 99)	7.67 (4.35 to 11.72)	95 (84 to 100)	5.60 (2.97 to 8.92)

* Confidence intervals were estimated using the percentile bootstrap method with 500 repetitions. Estimates were calculated only for cells with more than 10 instances of an outcome across the two groups. RR denotes risk ratio.

unvaccinated group — resulting in an estimated effectiveness of 62% for days 14 through 20 after the first dose, 80% for days 21 through 27, and 92% for 7 or more days after the second dose.

The large sample size in our study also allowed us to estimate vaccine effectiveness for specific subpopulations that the randomized trial was not sufficiently powered to evaluate. In the trial, the estimated efficacy for Covid-19 among persons up to 55 years of age, older than 55 years, and 65 years or older 7 days after the second dose was 94 to 96%. We were able to study more granular age groups, and we estimated that the vaccine effectiveness was similar for adults 70 years of age or older and for younger age groups for the same time period.

The randomized trial estimated vaccine efficacy for patients with one or more coexisting conditions according to the Charlson comorbidity index[16] and specifically for patients with obesity or hypertension. These measures do not provide clarity regarding effectiveness in patients with multiple coexisting conditions. We estimated vaccine effectiveness in relation to various numbers of coexisting conditions and found indications that effectiveness may be slightly lower among persons with higher numbers of coexisting conditions.

Two factors make the present study uniquely suited to evaluating the effectiveness of the BNT162b2 vaccine in a practical application: first, a rare combination of rich medical background data, Covid-19 PCR test results (for the documented infection outcome), and patient follow-up data in both community (for the symptomatic Covid-19 outcome) and inpatient (for all other outcomes) settings — CHS has maintained such an integrated data repository for over half the Israeli population, and has updated it daily, for more than two decades; and second, the rapid pace and high uptake of Covid-19 vaccine in Israel and the high disease rates during the vaccination campaign. On the other hand, the rapid pace of the vaccination campaign contributed to the frequent censoring of data for matched unvaccinated controls, especially among those over the age of 60 years (often only a few days after matching) and the corresponding reduction in the average follow-up period of the study.

Concerns have emerged regarding the possible resistance of SARS-CoV-2 variants to Covid-19 vaccines[17,18] and neutralizing antibodies.[19,20] During the study period, an increasing share of SARS-CoV-2 isolates in Israel — up to 80% in the days before data extraction — were of the B.1.1.7 variant.[21] Thus, this study estimates an average effectiveness of the vaccine over multiple strains. Although we cannot provide a specific effectiveness estimate for the B.1.1.7 variant, the plateau observed during the later periods in the cumulative incidence curve for vaccinated persons suggests that the BNT162b2 vaccine is also effective for this variant, an observation consistent with previous reports that showed preserved neutralizing antibody titers.[22] The B.1.351 variant

was estimated to be rare in Israel at the time of data extraction.[23]

As with any observational study, our study may have been affected by residual confounding due to differences between vaccinated persons and unvaccinated controls, especially in terms of health-seeking behavior. We therefore performed rigorous matching on a wide range of factors that may be expected to confound the causal effect of the vaccine on the various outcomes. After the matching process, we found a consistent pattern of similarity between the groups in the days just before day 12 after the first dose (the anticipated onset of the vaccine effect), which thus serve as a "negative control"[24] period (Fig. 2, Fig. S6, and Table S7). This similarity occurred despite a temporary increase in events among unvaccinated controls during the very first days after the first vaccine dose, most likely stemming from the fact that persons who choose to be vaccinated on a specific day are feeling well at the time of vaccination. The similarity of the study groups in coexisting conditions and known risk factors for severe Covid-19 (Table 1 and Fig. S2) provides further evidence of exchangeability (i.e., absence of confounding). However, this rigorous matching process came at the cost of not including in the final cohort approximately 34% of the vaccinated persons who otherwise met the study's eligibility criteria. Limited matching on age and sex only would have been insufficient to eliminate the early confounding (Fig. S6).

We also excluded population groups with high internal variability in the probability of vaccination or outcome, such as health care workers, persons confined to the home for medical reasons, and nursing home residents, to avoid residual confounding. Although the randomized trial was also less likely to include persons who were not healthy enough to comply with the scheduled visits and vaccination plan, it did not exclude health care workers.

To assess a possible selection bias that could stem from informative censoring, whereby controls who are vaccinated feel well around the time of vaccination, we performed a sensitivity analysis in which they were kept in the unvaccinated group for a period of time that was set differently for each outcome (Fig. S7 and Table S5). This analysis showed results similar to those of the main analysis, which suggests that any such bias was small in our analysis.

Finally, the date of onset of symptoms was not available for the analysis. Instead, for infection outcomes, the date was set to the date of swab collection for the first positive PCR test. Given that there was likely to have been a time gap between the onset of symptoms and swab collection, the observed divergence of the cumulative incidence plots for the infection outcomes between the vaccinated persons and unvaccinated controls may be slightly delayed. In parallel, there might be an underestimation of the vaccine effectiveness at each time window, since the estimate actually reflects a narrower window for the vaccine to be active. Because SARS-CoV-2 PCR testing is highly accessible in Israel and can be done without referral in a matter of hours, we estimate this potential time gap and thus the vaccine effectiveness underestimation to be small. In interpreting the effectiveness estimates for more severe outcomes, longer median gaps should be kept in mind (Fig. S3): 1 day for hospitalization, 5 days for severe Covid-19, and 11 days for Covid-19 death.

This study estimates a high effectiveness of the BNT162b2 vaccine for preventing symptomatic Covid-19 in a noncontrolled setting, similar to the vaccine efficacy reported in the randomized trial. Our study also suggests that effectiveness is high for the more serious outcomes: hospitalization, severe illness, and death. Furthermore, the estimated benefit increases in magnitude as time passes. These results strengthen the expectation that newly approved vaccines can help to mitigate the profound global effects of the Covid-19 pandemic.

Dr. Lipsitch receives support from the Morris–Singer Fund.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

Owing to data privacy regulations, the raw data for this study cannot be shared.

We thank Daniel Nevo, for his advice on analysis, and Yatir Ben-Shlomo, Galit Shaham, and Uriah Finkel, for their assistance with the project.

REFERENCES

1. Polack FP, Thomas SJ, Kitchin N, et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. N Engl J Med 2020;383:2603-15.
2. Baden LR, El Sahly HM, Essink B, et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. N Engl J Med 2020; 384:403-16.
3. ClinicalTrials.gov. Study to describe the safety, tolerability, immunogenicity, and efficacy of RNA vaccine candidates against COVID-19 in healthy individuals. 2020 (https://clinicaltrials.gov/ct2/show/NCT04368728).
4. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. Am J Epidemiol 2016;183:758-64.
5. Zambrano LD, Ellington S, Strid P, et al. Update: characteristics of symptomatic women of reproductive age with laboratory-confirmed SARS-CoV-2 infection by pregnancy status — United States, January 22–October 3, 2020. MMWR Morb Mortal Wkly Rep 2020;69:1641-7.
6. Centers for Disease Control and Prevention. People with certain medical conditions. February 3, 2021 (https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html).
7. Centers for Disease Control and Prevention. Older adults at greater risk of requiring hospitalization or dying if diagnosed with COVID-19. December 13, 2020 (https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/older-adults.html).
8. National Institutes of Health. Coronavirus disease 2019 (COVID-19) treatment guidelines. 2021 (https://www.covid19treatmentguidelines.nih.gov/).
9. Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. Psychol Methods 2010;15:234-49.
10. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958;53:457-81.
11. Lipsitch M, Goldstein E, Ray GT, Fireman B. Depletion-of-susceptibles bias in influenza vaccine waning studies: how to ensure robust results. Epidemiol Infect 2019;147:e306.
12. Gallagher T, Lipsitch M. Postexposure effects of vaccines on infectious diseases. Epidemiol Rev 2019;41:13-27.
13. Hernán MA. The hazards of hazard ratios. Epidemiology 2010;21:13-5.
14. Kaslow DC. Force of infection: a determinant of vaccine efficacy? January 26, 2021 (https://www.medrxiv.org/content/10.1101/2021.01.21.21250235v1.full). preprint.
15. Chodick G, Tene L, Patalon T, et al. The effectiveness of the first dose of BNT162b2 vaccine in reducing SARS-CoV-2 infection 13-24 days after immunization: real-world evidence. January 29, 2021 (https://www.medrxiv.org/content/10.1101/2021.01.27.21250612v1). preprint.
16. Charlson M, Szatrowski TP, Peterson J, Gold J. Validation of a combined comorbidity index. J Clin Epidemiol 1994;47:1245-51.
17. Mahase E. Covid-19: Novavax vaccine efficacy is 86% against UK variant and 60% against South African variant. BMJ 2021;372:n296.
18. Johnson & Johnson announces single-shot Janssen COVID-19 vaccine candidate met primary endpoints in interim analysis of its phase 3 ENSEMBLE trial. New Brunswick, NJ: Johnson & Johnson, January 29, 2021 (https://www.jnj.com/johnson-johnson-announces-single-shot-janssen-covid-19-vaccine-candidate-met-primary-endpoints-in-interim-analysis-of-its-phase-3-ensemble-trial).
19. Wang P, Liu L, Iketani S, et al. Increased resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7 to antibody neutralization. bioRxiv 2021 January 26 (Preprint).
20. Xie X, Zou J, Fontes-Garfias CR, et al. Neutralization of N501Y mutant SARS-CoV-2 by BNT162b2 vaccine-elicited sera. bioRxiv 2021 January 7 (Preprint).
21. Efrati I. Israel to extend COVID vaccine drive to anyone over 16 starting Thursday. Haaretz. February 3, 2021.
22. Muik A, Wallisch A-K, Sänger B, et al. Neutralization of SARS-CoV-2 lineage B.1.1.7 pseudovirus by BNT162b2 vaccine–elicited human sera. Science 2021 January 29 (Epub ahead of print).
23. Hoffman M. Coronavirus: 80 cases of South African variant discovered in Israel. Jerusalem Post. February 2, 2021 (https://www.jpost.com/breaking-news/coronavirus-80-cases-of-south-african-variant-discovered-in-israel-657452).
24. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology 2010;21:383-8.

ORIGINAL ARTICLE

Dexmedetomidine or Propofol for Sedation in Mechanically Ventilated Adults with Sepsis

C.G. Hughes, P.T. Mailloux, J.W. Devlin, J.T. Swan, R.D. Sanders, A. Anzueto, J.C. Jackson, A.S. Hoskins, B.T. Pun, O.M. Orun, R. Raman, J.L. Stollings, A.L. Kiehl, M.S. Duprey, L.N. Bui, H.R. O'Neal, Jr., A. Snyder, M.A. Gropper, K.K. Guntupalli, G.J. Stashenko, M.B. Patel, N.E. Brummel, T.D. Girard, R.S. Dittus, G.R. Bernard, E.W. Ely, and P.P. Pandharipande, for the MENDS2 Study Investigators*

ABSTRACT

BACKGROUND

Guidelines currently recommend targeting light sedation with dexmedetomidine or propofol for adults receiving mechanical ventilation. Differences exist between these sedatives in arousability, immunity, and inflammation. Whether they affect outcomes differentially in mechanically ventilated adults with sepsis undergoing light sedation is unknown.

METHODS

In a multicenter, double-blind trial, we randomly assigned mechanically ventilated adults with sepsis to receive dexmedetomidine (0.2 to 1.5 μg per kilogram of body weight per hour) or propofol (5 to 50 μg per kilogram per minute), with doses adjusted by bedside nurses to achieve target sedation goals set by clinicians according to the Richmond Agitation–Sedation Scale (RASS, on which scores range from −5 [unresponsive] to +4 [combative]). The primary end point was days alive without delirium or coma during the 14-day intervention period. Secondary end points were ventilator-free days at 28 days, death at 90 days, and age-adjusted total score on the Telephone Interview for Cognitive Status questionnaire (TICS-T; scores range from 0 to 100, with a mean of 50±10 and lower scores indicating worse cognition) at 6 months.

RESULTS

Of 432 patients who underwent randomization, 422 were assigned to receive a trial drug and were included in the analyses — 214 patients received dexmedetomidine at a median dose of 0.27 μg per kilogram per hour, and 208 received propofol at a median dose of 10.21 μg per kilogram per minute. The median duration of receipt of the trial drugs was 3.0 days (interquartile range, 2.0 to 6.0), and the median RASS score was −2.0 (interquartile range, −3.0 to −1.0). We found no difference between dexmedetomidine and propofol in the number of days alive without delirium or coma (adjusted median, 10.7 vs. 10.8 days; odds ratio, 0.96; 95% confidence interval [CI], 0.74 to 1.26), ventilator-free days (adjusted median, 23.7 vs. 24.0 days; odds ratio, 0.98; 95% CI, 0.63 to 1.51), death at 90 days (38% vs. 39%; hazard ratio, 1.06; 95% CI, 0.74 to 1.52), or TICS-T score at 6 months (adjusted median score, 40.9 vs. 41.4; odds ratio, 0.94; 95% CI, 0.66 to 1.33). Safety end points were similar in the two groups.

CONCLUSIONS

Among mechanically ventilated adults with sepsis who were being treated with recommended light-sedation approaches, outcomes in patients who received dexmedetomidine did not differ from outcomes in those who received propofol. (Funded by the National Institutes of Health; ClinicalTrials.gov number, NCT01739933.)

The authors' full names, academic degrees, and affiliations are listed in the Appendix. Address reprint requests to Dr. Hughes at 1211 21st Ave. S., 422 MAB, Nashville, TN 37212, or at christopher.hughes@vumc.org.

*The MENDS2 Study Investigators are listed in the Supplementary Appendix, available at NEJM.org.

This article was published on February 2, 2021, at NEJM.org.

N Engl J Med 2021;384:1424-36.
DOI: 10.1056/NEJMoa2024922

Worldwide, at least 20 million patients each year have sepsis with severe organ dysfunction,[1] with over 20% receiving mechanical ventilation.[2,3] Sedative medications are frequently used for patient comfort and safety but may potentiate acute brain dysfunction (e.g., delirium or coma) and long-term cognitive impairment.[4-10] Basic and translational studies show that among the recommended sedatives, dexmedetomidine (an alpha$_2$ receptor agonist) has antiinflammatory and bacterial clearance properties that are superior to those of gamma-aminobutyric acid (GABA) agonists, such as benzodiazepines and propofol, and also reduces neuronal apoptosis and promotes biomimetic sleep — all of which could improve clinical outcomes.[11-17] Trials comparing dexmedetomidine with benzodiazepines in adults have shown that the use of dexmedetomidine results in improvement in outcomes such as delirium, coma, and time receiving mechanical ventilation.[18,19] Patients treated with dexmedetomidine had a lower incidence of subsequent infection,[19] and the beneficial effects of dexmedetomidine, including lower 28-day mortality, were more pronounced in patients with sepsis.[18,20]

A noninferiority trial comparing dexmedetomidine with propofol in critically ill patients, about half of whom had sepsis, showed that patients who received dexmedetomidine were more interactive, but the choice of sedation did not affect the duration of mechanical ventilation, the length of stay in the intensive care unit (ICU) or hospital, or short-term mortality.[21] The differences between the sedatives with respect to the risk of acute brain dysfunction or cognitive impairment and mortality months after critical illness were unclear. Subsequent open-label trials with dexmedetomidine as the primary sedative did not show a reduction in acute brain dysfunction, a greater number of ventilator-free days, or lower mortality at 180 days than was shown with control sedative regimens (primarily propofol), although concomitant nontrial sedatives were frequently used and few patients were maintained at light sedation.[22,23]

The Society of Critical Care Medicine[24] recommends sedation with either dexmedetomidine or propofol targeted to light levels of sedation for adults receiving mechanical ventilation and continuous sedation. Given the superior immunomodulatory effects of dexmedetomidine and its benefit in patients with sepsis as compared with benzodiazepines, we designed the MENDS2 trial (Maximizing the Efficacy of Sedation and Reducing Neurological Dysfunction and Mortality in Septic Patients with Acute Respiratory Failure) to test whether dexmedetomidine leads to better short-term and long-term outcomes than propofol in mechanically ventilated adults with sepsis.

Methods

Trial Design and Oversight

We conducted a double-blind, randomized, controlled trial at 13 medical centers in the United States. The institutional review board at each center approved the protocol (available with the full text of this article at NEJM.org). Patients or their surrogates provided written informed consent before enrollment. The trial was designed by the authors, who gathered and analyzed the data, attest to the accuracy and completeness of the data, vouch for the fidelity of the trial to the protocol, and wrote and agreed to submit the manuscript for publication. An independent data and safety monitoring board provided oversight of the trial. Pfizer supplied the dexmedetomidine trial drug but had no role in the design or conduct of the trial, analysis of the data, or writing of the manuscript. The Food and Drug Administration approved an Investigational New Drug application for dexmedetomidine administered for more than 24 hours and for doses up to 1.5 μg per kilogram per hour (see the Supplementary Appendix, available at NEJM.org). We registered the trial at ClinicalTrials.gov before enrollment began. Before group assignments were unmasked, we registered the statistical analysis plan at Open Science Framework (https://osf.io/dfyxh/) in January 2019 (with publication in March 2020).[25]

Patient Selection and Randomization

We included adults who were sequentially admitted to a medical or surgical ICU, had suspected or known infection, and were treated with continuous sedation for invasive mechanical ventilation. Patients were excluded if they had baseline severe cognitive impairment; were pregnant or breast-feeding; were blind, deaf, or unable to understand approved languages; had second-degree or third-degree heart block or persistent bradycardia requiring intervention; had an allergy to dexmedetomidine or propofol; had an indication for benzodiazepines; were anticipated to have immediate discontinuation of mechanical ventila-

tion; were expected to have neuromuscular blockade for more than 48 hours; were in a moribund state; or had received mechanical ventilation for more than 96 hours before meeting all inclusion criteria. Additional details on exclusion and inclusion criteria are provided in Section S1 in the Supplementary Appendix. We randomly assigned patients to receive dexmedetomidine or propofol in a 1:1 ratio using computer-generated permuted blocks stratified by enrollment site and age (<65 years vs. ≥65 years). Researchers, clinicians (except bedside nurses), patients, and families were unaware of the group assignments.

TRIAL INTERVENTIONS AND MEASUREMENTS

Investigational pharmacists prepared dexmedetomidine (5 μg per milliliter) and propofol (10 mg per milliliter) in identical intravenous fluid bags covered with opaque plastic bags to be administered in units of milliliters per hour to maintain study masking (Sections S2 and S3). Bedside nurses covered intravenous tubing with opaque coverings and verified that covers were in place before study personnel or clinicians entered patients' rooms. The trial drug was initially infused at a dose corresponding to the same sedative dosing that the patient was receiving immediately before randomization. Bedside nurses used a weight-based dosing guideline (0.15 to 1.5 μg per kilogram of actual body weight per hour for dexmedetomidine and 5 to 50 μg per kilogram of actual body weight per minute for propofol) to adjust the trial drug every 10 minutes to target sedation goals set by the clinical team and documented each adjustment and the rationale for it. The clinical team used the Richmond Agitation–Sedation Scale (RASS, on which scores range from −5 [unresponsive] to +4 [combative]),[26] to set the sedation goal, which was primarily light sedation (RASS score 0 to −2).

Administration of the trial drug was temporarily held in the event of hypotension, bradycardia, sedation deeper than the target level, spontaneous awakening trials, or surgery. The trial drug was permanently discontinued if the patient had persistent symptomatic bradycardia, new onset second- or third-degree heart block, serious allergic reactions, suspected propofol-related infusion syndrome (refractory shock, rhabdomyolysis, acidosis, and kidney failure related to high propofol exposure), or any serious adverse event related to the intervention. The trial drug was discontinued after the 14-day intervention period, extubation, or discharge from the ICU, whichever came first. Patients whose trachea was extubated and reintubated within the 14-day intervention period resumed the trial drug if sedation was indicated.

We treated pain with intermittent opioid boluses or fentanyl infusion (see Section S4 for details regarding rescue sedation, neuromuscular blockade, and treatment of agitated delirium). Additional patient care practices (e.g., administration of fluids, vasopressors, or antibiotics and extubation criteria) were based on international guideline recommendations.[24,27]

All centers performed, and investigators reinforced, the ABCDE (awakening and breathing coordination, choice of sedation, delirium monitoring and management, and early mobility) bundle,[28,29] with daily adherence recorded. In addition to care assessments made by nurses, trained research personnel assessed patients with the use of RASS for level of arousal,[26] Confusion Assessment Method for the ICU (CAM-ICU)[30] for delirium, and the Critical Care Pain Observation Tool[31] for pain; assessments were made twice daily in the ICU and then once daily after transfer from the ICU for up to 14 days or until discharge from the hospital or death. We strived to conduct delirium assessments when the patient was maximally awake. A RASS score of −4 or −5 indicated coma, and a positive CAM-ICU score indicated delirium.

Six months after randomization, research personnel assessed patients' cognition with the Telephone Interview for Cognitive Status (TICS) questionnaire[32] and a validated telephone cognitive battery,[33] functional status with the Katz Activities of Daily Living (ADL) scale[34] and the Functional Activities Questionnaire (FAQ),[35] and quality of life with the European Quality of Life–5 Dimensions (EQ-5D) survey (Section S5).[36]

TRIAL END POINTS

The primary efficacy end point was the number of calendar days alive without delirium or coma during the 14-day intervention period. Secondary efficacy end points included ventilator-free days at 28 days, death at 90 days, and global cognition at 6 months using the age-adjusted TICS total score (TICS-T score). Additional details regarding efficacy, adherence, and safety end points are provided in Section S5, along with a list of additional end points not reported here.

STATISTICAL ANALYSIS

Owing to enrollment that was slower than anticipated, the data and safety monitoring board and the National Institutes of Health approved a protocol amendment in March 2017 to lower the enrollment target from 530 patients to 420 patients receiving the trial drug to provide 85% power to detect a 1.5-day difference in days alive without delirium or coma between groups and 80% power to detect a 12 percentage-point absolute difference in mortality at 90 days, assuming an expected mortality of 30% in the propofol group. We had at least 80% power to detect a 3.9-point difference in age-adjusted TICS-T scores between groups, with a 5-point difference considered to be clinically important.

We analyzed data in the modified intention-to-treat population, which was prespecified as all patients who underwent randomization and received a trial drug. We analyzed primary and secondary end points using both univariate methods and multivariable regression models and considered adjusted analyses to be the primary analyses. We analyzed days alive without delirium and coma, ventilator-free days, and age-adjusted TICS-T scores at 6 months using proportional-odds logistic regression and analyzed death at 90 days using Cox proportional-hazards regression (adjusted for covariates listed in Section S5).

We adjusted the level of statistical significance for the primary end point analysis to $P<0.044$ to account for one prespecified planned interim analysis. The level of statistical significance for all other end points was $P<0.05$. Simple imputation was used for missing in-hospital variables and multiple imputation for partially available long-term end points to avoid bias owing to missing variables. We did not adjust for multiple comparisons in the analysis of secondary end points. We used Research Electronic Data Capture software (REDCap, Vanderbilt University) for data management and R, version 3.6.2 (R Foundation for Statistical Computing), for statistical analyses.

RESULTS

PATIENTS

From May 2013 through December 2018, we screened 4840 patients, 4402 (91%) of whom met at least one exclusion criterion (Fig. 1). Of 438 patients enrolled, 6 were subsequently found to be ineligible, 432 patients underwent randomization, and 422 began receiving dexmedetomidine (214 patients) or propofol (208 patients). The demographic and in-hospital characteristics of the patients are shown in Table 1 and Table S1.

TRIAL INTERVENTIONS

Details of trial drug dosing, dose adjustment, and sedation regimen are shown in Table 2 and Table S2. The median RASS score as assessed by the research team was −2 (interquartile range, −3.00 to −1.00) while patients were receiving a trial drug (median days of administration, 3.0 [interquartile range, 2.0 to 6.0]), indicating light sedation and ability of patients to make eye contact with only verbal stimulation (Fig. S1). The overall time spent at the target sedation was close to 60% in both groups (Fig. S2). Bedside nurses adjusted the trial drug infusion a median of 10 times (interquartile range, 5 to 21) over the duration of administration. Common reasons for adjustment of the infusion included undersedation, oversedation, and hypotension. The trial drug was temporarily held in approximately one quarter of all patients. Rescue midazolam was used in about half the patients, most often for procedural sedation or during neuromuscular blockade, and the median daily exposure on days it was administered was 4 mg (interquartile range, 2 to 11). The use of open-label propofol (received by 13% in the dexmedetomidine group and 8% in the propofol group) and dexmedetomidine (4% in the dexmedetomidine group and 3% in the propofol group) was infrequent and doses were low, indicating high adherence to the protocol. Overall, 42% of the patients received an antipsychotic medication. Soft wrist restraints during receipt of mechanical ventilation were the standard of care in our trial ICUs; thus, 96% of patients had restraints in place for a median of 3 days (interquartile range, 2 to 5). Neuromuscular blockade infusion was used in 17% of patients for a median of 1 day (interquartile range, 1 to 2) at some point while they were receiving the trial drug. Pain was well controlled in both groups according to Critical Care Pain Observation Tool scoring, and we noted high adherence to all components of the ABCDE bundle. We proactively assessed for unblinding among clinicians and research staff and found an episode of unblinding in 58 patients (14%), with a similar frequency in the two groups.

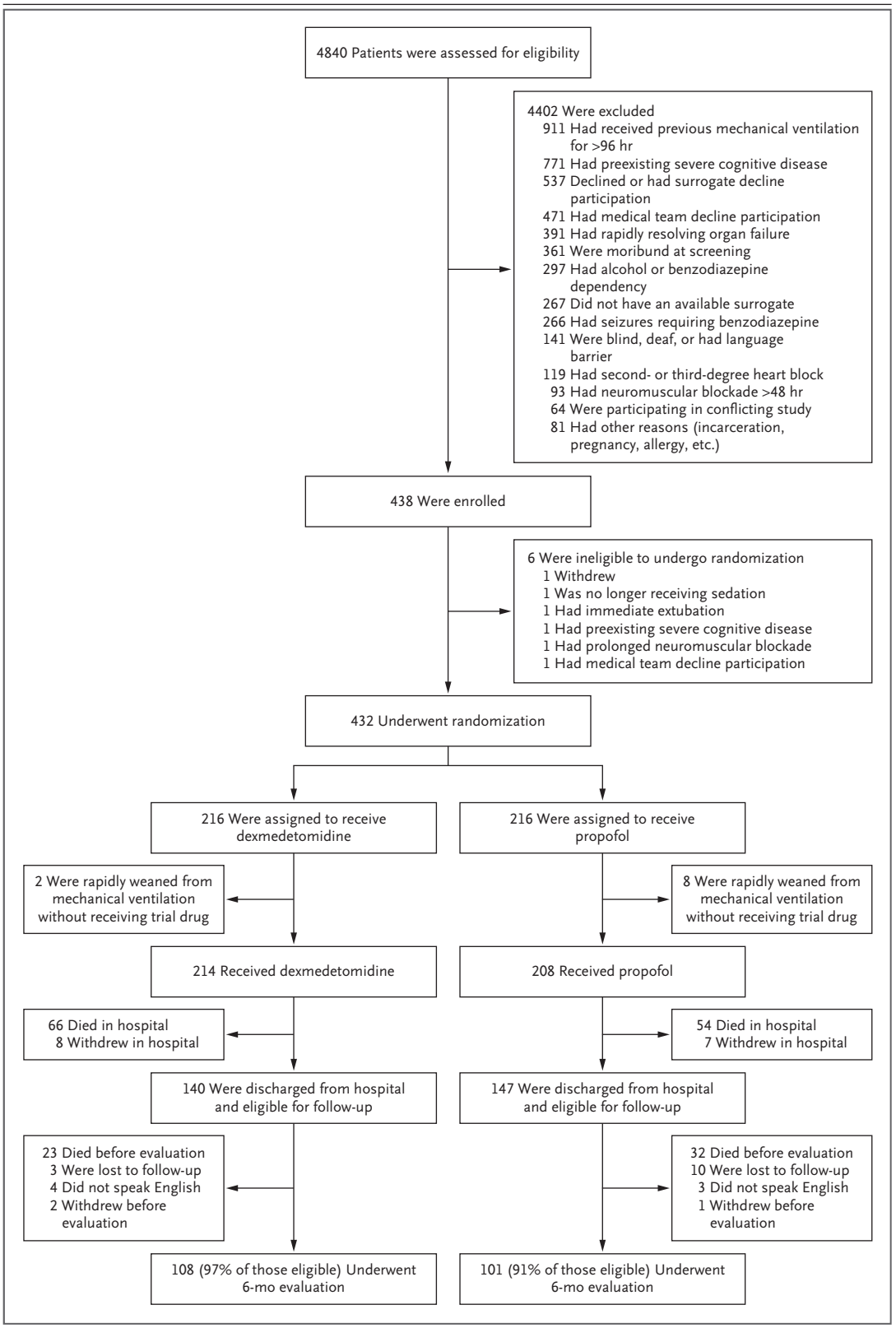

4840 Patients were assessed for eligibility
4402 Were excluded
911 Had received previous mechanical ventilation for >96 hr
771 Had preexisting severe cognitive disease
537 Declined or had surrogate decline participation
471 Had medical team decline participation
391 Had rapidly resolving organ failure
361 Were moribund at screening
297 Had alcohol or benzodiazepine dependency
267 Did not have an available surrogate
266 Had seizures requiring benzodiazepine
141 Were blind, deaf, or had language barrier
119 Had second- or third-degree heart block
93 Had neuromuscular blockade >48 hr
64 Were participating in conflicting study
81 Had other reasons (incarceration, pregnancy, allergy, etc.)
438 Were enrolled
6 Were ineligible to undergo randomization
1 Withdrew
1 Was no longer receiving sedation
1 Had immediate extubation
1 Had preexisting severe cognitive disease
1 Had prolonged neuromuscular blockade
1 Had medical team decline participation
432 Underwent randomization
216 Were assigned to receive dexmedetomidine
216 Were assigned to receive propofol
2 Were rapidly weaned from mechanical ventilation without receiving trial drug
8 Were rapidly weaned from mechanical ventilation without receiving trial drug
214 Received dexmedetomidine
208 Received propofol
66 Died in hospital
8 Withdrew in hospital
54 Died in hospital
7 Withdrew in hospital
140 Were discharged from hospital and eligible for follow-up
147 Were discharged from hospital and eligible for follow-up
23 Died before evaluation
3 Were lost to follow-up
4 Did not speak English
2 Withdrew before evaluation
32 Died before evaluation
10 Were lost to follow-up
3 Did not speak English
1 Withdrew before evaluation
108 (97% of those eligible) Underwent 6-mo evaluation
101 (91% of those eligible) Underwent 6-mo evaluation

Figure 1 (facing page). Screening, Randomization, Follow-up, and Analysis.

The number of patients excluded for each criterion total more than the total number of patients excluded because some patients met more than one exclusion criterion.

EFFICACY END POINTS

The adjusted number of days alive without delirium or coma over the 14-day intervention period was not significantly different between the dexmedetomidine group (adjusted median, 10.7 days; 95% confidence interval [CI], 8.5 to 12.5) and the propofol group (adjusted median, 10.8 days; 95% CI, 8.7 to 12.6) (odds ratio, 0.96; 95% CI, 0.74 to 1.26; P=0.79). Similarly, we found no significant differences between the dexmedetomidine and propofol groups in the number of ventilator-free days at 28 days (adjusted median, 23.7 vs. 24.0 days; odds ratio, 0.98; 95% CI, 0.63 to 1.51) or in death at 90 days (81 patients [38%] vs. 82 patients [39%]; hazard ratio, 1.06; 95% CI, 0.74 to 1.52). Results of primary and secondary efficacy end point analyses are shown in Table 3, Figure 2, and Figure S3.

We assessed more than 90% of eligible patients at 6 months after randomization (Fig. 1). Approximately 25% in each group had age-adjusted TICS-T scores that were 2 standard deviations below population norms (i.e., a score of ≤30, at a level consistent with impairment), which suggests clinically important cognitive dysfunction 6 months after critical illness. We observed no significant differences between the dexmedetomidine and propofol groups in age-adjusted TICS-T scores at 6 months (adjusted median score, 40.9 vs. 41.4; odds ratio, 0.94; 95% CI, 0.66 to 1.33). There were no clinically meaningful differences between groups in median cognitive, functional, and quality-of-life assessment scores at 6 months (Table S4).

Results of sensitivity analyses that included the 10 patients who underwent randomization but never received a trial drug (Table S5) were qualitatively similar to the results of the main analyses. Results of differential effects of the study treatment on end points according to age at enrollment, baseline cognition, and medical as compared with surgical hospitalization show that the clinical importance of these interactions appeared to be minimal (Figs. S4 through S9); however, the trial may not have been adequately powered to draw conclusions about these or other subgroups.

SAFETY END POINTS

Data on organ dysfunction and safety end points by group are shown in Tables S6 and S7, respectively. The proportions of patients who had organ dysfunction, hypotension, or severe lactic acidosis after randomization were similar in the two groups. Symptomatic bradycardia requiring discontinuation of the trial drug was similar in the two groups (Table S2). Fewer patients in the dexmedetomidine group had acute respiratory distress syndrome (ARDS) or signs of trial drug withdrawal, and fewer patients in the propofol group extubated themselves. Median plasma triglyceride levels and the proportion of patients with severely elevated levels of triglyceride (>500 mg per deciliter) were quantitatively higher in the propofol group than in the dexmedetomidine group on days 7 and 14, although these differences are unlikely to be clinically relevant. Similarly, median plasma cortisol levels at day 14 were slightly lower in the dexmedetomidine group than in the propofol group, including a higher proportion of patients with low cortisol (<20 μg per deciliter). Clinicians had access to these results without indication of group assignment and discontinued the trial drug in eight patients owing to hypertriglyceridemia. One patient had suspected propofol-related infusion syndrome (later disproved) and had the propofol discontinued.

DISCUSSION

In this multicenter, double-blind, randomized, controlled trial involving mechanically ventilated adults with sepsis who were being treated with recommended light-sedation approaches, we did not find evidence that sedation with dexmedetomidine led to more days alive without acute brain dysfunction than propofol. Furthermore, we found no difference in ventilator-free days at 28 days, death at 90 days, or global cognition (as assessed with the use of age-adjusted TICS-T scores) at 6 months between the dexmedetomidine and propofol groups. Safety end points were also similar in the two groups.

Although recent data suggest that many criti-

Table 1. Baseline Characteristics of the Patients.*

Characteristic	Dexmedetomidine (N=214)	Propofol (N=208)
Median age (IQR) — yr	59 (48–68)	60 (50–68)
Female sex — no. (%)	93 (43)	88 (42)
Median body-mass index (IQR)†	30 (25–38)	29 (25–37)
Race or ethnic group — no. (%)‡		
White	188 (88)	177 (85)
Black	15 (7)	23 (11)
Latinx	12 (6)	18 (9)
Multiple or other	11 (5)	8 (4)
Median IQCODE-SF score (IQR)§	3.06 (3.00–3.23)	3.00 (3.00–3.25)
Median Charlson Comorbidities Index score (IQR)¶	2 (1–4)	2 (1–4)
Admitted to surgical ICU — no. (%)	76 (36)	72 (35)
Median APACHE II score at ICU admission (IQR)‖	27 (21–32)	27 (22–32)
Median days from ICU admission to trial enrollment (IQR)	1.21 (0.67–1.95)	1.17 (0.68–1.94)
Median days of mechanical ventilation before trial enrollment (IQR)	0.98 (0.58–1.36)	0.97 (0.61–1.54)
Median total SOFA score at trial enrollment (IQR)**	10 (8–13)	10 (8–12)
Shock, receiving vasopressor, at enrollment — no. (%)	119 (56)	102 (49)
Known or suspected source of infection — no. (%)		
Blood	92 (43)	79 (38)
Lung	116 (54)	133 (64)
Abdomen	19 (9)	20 (10)
Urinary tract	46 (21)	55 (26)
Skin or wound	23 (11)	26 (12)
Stool	12 (6)	12 (6)
Other	24 (11)	21 (10)
Infection status — no. (%)		
Infection confirmed by culture	146 (68)	132 (63)
Infection suspected but not confirmed by culture	58 (27)	68 (33)
Infection ruled out	10 (5)	8 (4)
Dexmedetomidine before enrollment — no. (%)	35 (16)	25 (12)
Propofol before enrollment — no. (%)	131 (61)	129 (62)
Benzodiazepine before enrollment — no. (%)	62 (29)	73 (35)
Opioid before enrollment — no. (%)	144 (67)	147 (71)
Antipsychotic agent before enrollment — no. (%)	24 (11)	27 (13)
Delirium at enrollment — no. (%)††	75 (35)	91 (44)
Level of arousal closest to the time of randomization — no. (%)‡‡		
Coma: RASS –5 or –4	81 (38)	74 (36)
Deep sedation: RASS –3	29 (14)	38 (18)
Light sedation: RASS –2 or –1	85 (40)	75 (36)
Awake and calm: RASS 0	13 (6)	14 (7)
Agitated: RASS +1 to +4	6 (3)	7 (3)

* Percentages may not total 100 because of rounding. Summary statistics are reported for nonmissing values. ICU denotes intensive care unit, and IQR interquartile range.

† The body-mass index is the weight in kilograms divided by the square of the height in meters.

‡ Race or ethnic group was reported by the patient or determined by the treating physicians.

§ The Informant Questionnaire on Cognitive Decline in the Elderly short form (IQCODE-SF)[37] was used to determine preexisting dementia; scores range from 1.0 to 5.0, with higher scores indicating more severe cognitive impairment.

¶ Scores on the Charlson Comorbidity Index range from 0 to 33, with higher scores indicating a higher risk of death from a coexisting illness.

‖ The Acute Physiology and Chronic Health Evaluation (APACHE II) assesses the risk of death on a scale from 0 to 71, with higher scores indicating a higher risk of death.

** The Sequential Organ Failure Assessment (SOFA) is used to track organ failure in the ICU; scores range from 0 to 24, with higher scores indicating greater severity of illness.

†† Delirium was deemed to be present when the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU, which scores delirium as either present [positive] or not present [negative]), was positive.

‡‡ The Richmond Agitation–Sedation Scale (RASS) measures levels of consciousness on a scale from –5 (unresponsive) to +4 (combative).

Table 2. Adherence and Sedation Regimen.

Outcome	Dexmedetomidine N = 214	Propofol N = 208
Median hours from meeting inclusion criteria to drug initiation (IQR)	22.4 (13.4–31.3)	22.1 (12.8–33.7)
Median hours from randomization to drug initiation (IQR)	1.3 (0.9–2.2)	1.3 (0.8–2.1)
Trial drug administration		
Median days of receipt of drug (IQR)	3.0 (2.0–5.0)	4.0 (2.0–6.0)
Median days from first meeting trial criteria to initiation of drug (IQR)	1.00 (0.00–1.00)	1.00 (0.00–1.00)
Median daily volume on days administered (IQR) — ml	119 (46–243)	131 (67–229)
Median daily dose on days administered (IQR)	0.27 μg/kg/hr (0.11–0.61)	10.2 μg/kg/min (5.5–18.4)
Median total no. of drug adjustments per patient (IQR)	9 (5–15.8)	11.5 (5.8–25)
Drug temporarily held — no. (%)*	60 (28)	57 (27)
Median no. of times drug temporarily held per patient (IQR)	1 (1–1)	1 (1–2)
Drug permanently discontinued — no. (%)	25 (12)	23 (11)
Trial or clinical team aware of the drug used — no. (%)	27 (13)	31 (15)
Withdrawal from trial during hospitalization — no. (%)	10 (5)	9 (4)
Median RASS score while receiving drug (IQR)	–2.00 (–3.00 to –1.00)	–1.95 (–3.03 to –0.98)
Percent time at target sedation level while receiving drug	57	60
Median CPOT score while receiving drug (IQR)†	0.33 (0.00–0.83)	0.31 (0.00–0.87)
Percent of days with adherence to ABCDE bundle‡		
Spontaneous awakening trial	98	98
Spontaneous breathing trial	93	95
Coordination of awakening and breathing trials	86	84
Nondrug delirium interventions	99	99
Early mobilization	91	92
Median daily fentanyl dose on days administered (IQR) — μg/hr	68 (28–119)	56 (20–95)
Midazolam exposure		
Ever used — no. (%)	114 (53)	90 (43)
Median days among users (IQR)	2.0 (1.0–4.0)	1.0 (1.0–2.0)
Median daily dose on days administered (IQR) — mg per day	3.8 (2.0–10.9)	4.0 (2.0–10.8)
Antipsychotic exposure		
Ever used — no. (%)	90 (42)	87 (42)
Median days among users (IQR)	5.0 (2.0–7.8)	4.0 (2.0–8.0)
Median daily dose on days administered (IQR) — mg§	2.2 (1.0–6.4)	3.6 (1.0–6.3)
Open-label propofol exposure		
Ever used — no. (%)	27 (13)	16 (8)
Median days among users (IQR)	2.0 (1.0–3.0)	1.5 (1.0–2.0)
Median daily dose on days administered (IQR) — μg/kg/min	10.8 (4.9–17.4)	4.8 (3.4–6.6)
Open-label dexmedetomidine exposure		
Ever used — no. (%)	9 (4)	6 (3)
Median days among users (IQR)	1.0 (1.0–2.0)	1.0 (1.0–3.2)
Median daily dose on days administered (IQR) — μg/kg/hr	0.24 (0.04–0.30)	0.26 (0.07–0.7)

* The reasons for temporary holding of the drug included oversedation, hypotension, or bradycardia; spontaneous awakening trials or times during which patients were not being sedated, were not receiving mechanical ventilation, or were in the operating room are not included.
† The Critical Care Pain Observation Tool (CPOT) is used to assess for pain by evaluating facial expression, body movement, muscle tension, and adherence to use of the ventilator if intubated or vocalization if extubated. Total scores range from 0 to 8, with scores higher than 2 indicating the presence of pain.
‡ The ABCDE bundle includes evaluations for awakening and breathing coordination, choice of sedation, delirium monitoring and management, and early mobility.
§ Values shown are in intravenous haloperidol equivalents.

Table 3. Primary and Secondary Efficacy End Points.*

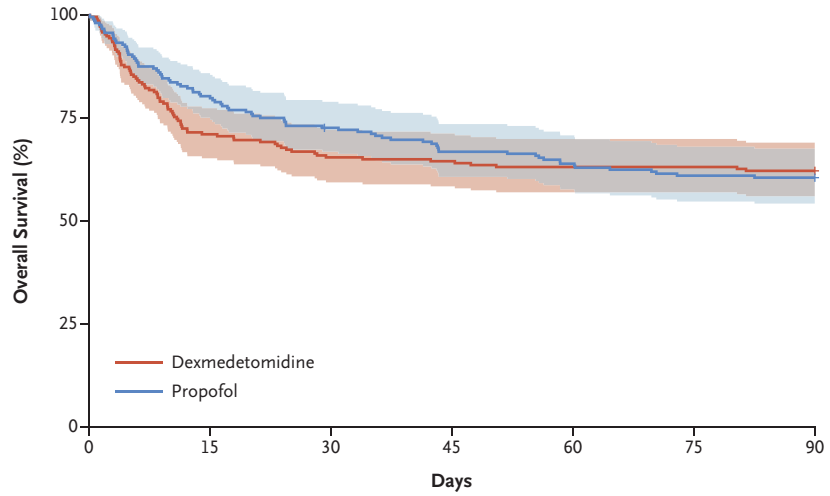
End Point	Dexmedetomidine (N=214)	Propofol (N=208)
Primary end point		
Days alive without delirium or coma at 14 days		
Unadjusted no. of days — median (IQR)	8.0 (1.0–12.8)	7.5 (1.8–11.2)
Adjusted no. of days — median (95% CI)	10.7 (8.5–12.5)	10.8 (8.7–12.6)
Adjusted odds ratio (95% CI)	0.96 (0.74–1.26)	Reference
Secondary end points		
Ventilator-free days at 28 days		
Unadjusted no. of days — median (IQR)	20.9 (0.0–26.1)	19.9 (4.2–24.9)
Adjusted no. days — median (95% CI)	23.7 (20.5–25.4)	24.0 (20.9–25.4)
Adjusted odds ratio (95% CI)	0.98 (0.63–1.51)	Reference
Death at 90 days		
Unadjusted no. of patients (%)	81 (38)	82 (39)
Adjusted hazard ratio (95% CI)	1.06 (0.74–1.52)	Reference
TICS-T score at 6 mo†		
Unadjusted score — median (IQR)	39 (28–48)	38 (30–46)
Adjusted score — median (95% CI)	40.9 (33.6–47.1)	41.4 (34.0–47.3)
Adjusted odds ratio (95% CI)	0.94 (0.66–1.33)	Reference

* Variables in adjusted analyses, except for analysis of death at 90 days, included the following: age at trial enrollment; education; baseline cognitive function as determined according to the IQCODE-SF; preexisting coexisting conditions according to the Charlson Comorbidities Index; SOFA assessment on the day of enrollment (excluding central nervous system component); level of arousal at randomization according to the RASS score closest to the time of randomization; exposure to propofol, dexmedetomidine, benzodiazepines, opioids, and antipsychotics between the time of ICU admission and midnight before enrollment; medical (vs. surgical) patient; and infection type. Variables in adjusted analyses for death at 90 days included the following: age at trial enrollment, baseline cognitive function as determined according to the IQCODE-SF, preexisting coexisting conditions according to the Charlson Comorbidities Index, SOFA assessment on the day of enrollment (excluding central nervous system component), medical (vs. surgical) patient, and infection type.

† Age-adjusted total scores on the Telephone Interview for Cognitive Status questionnaire (TICS-T) range from 0 to 100 with a mean of 50±10; lower scores indicate worse cognition, and a score of 35 or less indicates cognitive impairment.

cally ill adults receiving mechanical ventilation may not require sedative infusions,[38,39] our trial specifically enrolled adults with sepsis who had a high severity of illness, a greater risk for ARDS, and a higher requirement for continuous sedation. It was important to better characterize the effect of greater arousability, analgesic properties, and lack of respiratory depression observed with dexmedetomidine as compared with GABAergic sedatives in this population. Data indicate meaningful differences between dexmedetomidine and GABAergic sedatives with respect to innate immunity and risk of infection, including evidence that dexmedetomidine may offer superior anti-inflammatory effects.[11-17] Despite these theoretical benefits and studies supporting the use of dexmedetomidine, the choice between dexmedetomidine and propofol alone does not appear to substantially affect patient outcomes in the complex milieu of critical illness with sepsis. Our findings, therefore, strongly reinforce current guidelines[24] that recommend the use of either dexmedetomidine or propofol for light sedation when continuous sedation is needed for adults with or without sepsis who require mechanical ventilation.

Our trial builds on other trials that have compared dexmedetomidine with propofol,[21-23,40] with important methodologic advances that include a higher degree of sedative trial drug blinding, a better separation between groups with regard to sedative exposure, and stricter adherence to light



No. at Risk (Cumulative No. of Deaths)							
Dexmedetomidine	214 (0)	152 (62)	140 (74)	138 (76)	135 (79)	135 (79)	133 (81)
Propofol	208 (0)	167 (41)	150 (57)	138 (69)	132 (75)	126 (81)	125 (82)

Figure 2. Effects of Dexmedetomidine and Propofol on 90-Day Survival.

The Kaplan–Meier method was used to estimate the probability of survival. In the adjusted analyses, there was no significant difference between the trial groups with respect to death at 90 days (hazard ratio with dexmedetomidine vs. propofol, 1.06; 95% confidence interval [CI], 0.74 to 1.52). Results have not been adjusted for multiple comparisons. The shading indicates 95% confidence intervals.

sedation approaches, with high compliance with a standardized, multicomponent sedation management bundle (i.e., the ABCDE bundle)[28,29] that has been shown to reduce mortality and improve other important outcomes. One study by Kawazoe et al.[22] randomly assigned 201 patients with sepsis who required mechanical ventilation to open-label sedation with dexmedetomidine (up to 0.7 μg per kilogram per hour) or sedation without dexmedetomidine (infusions of propofol or midazolam or both) for up to 7 days. On 1 or more study days, 29% of the dexmedetomidine group received propofol (nearly three times the crossover rate of our study) and up to 21% received midazolam. The authors found no significant difference in the number of days without delirium or coma, the number of ventilator-free days, or mortality at 28 days with dexmedetomidine use, although the trial was probably underpowered to measure a difference in mortality.

More recently, Shehabi et al.[23] performed a landmark open-label, randomized trial of dexmedetomidine (up to 1.5 μg per kilogram per hour) as compared with usual care (infusions of propofol or midazolam or both) for up to 28 days in more than 3900 patients with critical illness. The authors did not find a significant difference between the groups in the number of days without delirium or coma at 28 days, the number of ventilator-free days at 28 days, death at 90 days (including in subgroup analyses of 806 patients with suspected or confirmed sepsis), or death at 180 days. Most patients (86%) in the dexmedetomidine group received concomitant propofol for a median of 2.0 days, and 23% received midazolam for a median of 0.5 days; this lack of separation between groups limits the interpretation of the results. Despite an unmasking episode in 14% of patients and crossover in about 10% of patients in the present study, we believe that our methodologic rigor allows a more definitive conclusion that dexmedetomidine and propofol have similar efficacy with regard to acute brain dysfunction, mechanical ventilation requirement, and mortality when light sedation goals and the ABCDE bundle are used to care for critically ill mechanically ventilated adults with sepsis. Biologically, patients with sepsis should derive im-

portant benefits from dexmedetomidine because of its immunomodulatory and antiinflammatory properties; thus, it is highly unlikely that patients without sepsis would have outcomes with dexmedetomidine substantially different from those we report.

An expanding area of interest in the care of critically ill patients is the prevention of cognitive impairment, functional impairment, and decline in quality of life after hospital discharge. The study by Shehabi et al.[23] showed similar scores in cognition (as assessed with the Informant Questionnaire on Cognitive Decline in the Elderly [IQCODE]) and quality of life (as assessed with the EQ-5D) at 180 days in the dexmedetomidine and control groups. Using a more robust cognitive assessment battery, we found clinically important cognitive dysfunction in approximately 25% of patients after sepsis and critical illness even with light sedation approaches, and the use of dexmedetomidine as compared with propofol did not alter this finding. Considering our high follow-up rates and use of a robust assessment battery, it appears that sedation choice does not affect survivorship outcomes when currently recommended sedation approaches are used.

Our trial has a number of strengths but also some notable limitations. We made every effort to mask the delivery of propofol and dexmedetomidine considering their different physical properties. Although an episode of unmasking of the group assignment to a clinician or research team member occurred in 14% of patients, adherence to blinding in our trial was higher than that reported in similar clinical trials of propofol and dexmedetomidine. We allowed clinicians to set sedation targets, achieved good separation between groups regarding sedative exposure, and had robust follow-up. In general, patients had light levels of sedation with low doses of sedative medications and concomitant opioid analgesia. This may reflect changing sedation strategies conforming to recommended practices or the need for lower sedative doses in patients with sepsis. We had some cross-contamination of sedative use, although substantially less than that in similar sedation studies, and had a rescue protocol that included the use of low-dose antipsychotic medications. The trial drug was started a median of 22 hours after the patient met all inclusion criteria, which may have limited our ability to affect outcomes. We had slower-than-anticipated enrollment, which required an adjustment of the sample size, yet had adequate power to study the questions of interest. Some exclusions were the result of clinicians not having equipoise regarding sedation for a given patient or were due to patients' (or their surrogates') decision not to agree to enrollment in the trial, factors that may affect generalizability. Overall, we believe that we studied a representative population of patients with sepsis in centers across the United States and provide more definitive evidence regarding the choice of sedation in critically ill patients with sepsis who require mechanical ventilation.

Our trial showed that among critically ill adults with sepsis who were receiving mechanical ventilation and for whom recommended light-sedation approaches were used, dexmedetomidine did not lead to better outcomes than propofol with respect to days alive without acute brain dysfunction, ventilator-free days, death at 90 days, or cognition at 6 months.

Supported by a grant (HL111111) from the National Institutes of Health (NIH), a grant (TR000445) from the NIH for the use of REDCap, grants from the NIH to Dr. Duprey (1F31AG066460) and to Dr. Girard (HL135144), and a grant from the National Heart Lung and Blood Institute to Dr. Pun (HL14678-01). Dexmedetomidine was provided by Pfizer (previously by Hospira).

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

A data sharing statement provided by the authors is available with the full text of this article at NEJM.org.

APPENDIX

The authors' full names and academic degrees are as follows: Christopher G. Hughes, M.D., M.S.C.R., Patrick T. Mailloux, D.O., John W. Devlin, Pharm.D., Joshua T. Swan, Pharm.D., M.P.H., Robert D. Sanders, M.D., Antonio Anzueto, M.D., James C. Jackson, Psy.D., Aimee S. Hoskins, B.S.N., R.N., Brenda T. Pun, D.N.P., R.N., Onur M. Orun, M.S., Rameela Raman, Ph.D., Joanna L. Stollings, Pharm.D., Amy L. Kiehl, M.A., Matthew S. Duprey, Pharm.D., Ph.D., Lan N. Bui, Pharm.D., M.P.H., Hollis R. O'Neal, Jr., M.D., Allison Snyder, M.S.N., A.P.R.N., Michael A. Gropper, M.D., Ph.D., Kalpalatha K. Guntupalli, M.D., Gregg J. Stashenko, M.D., Mayur B. Patel, M.D., M.P.H., Nathan E. Brummel, M.D., M.S.C.I., Timothy D. Girard, M.D., M.S.C.I., Robert S. Dittus, M.D., M.P.H., Gordon R. Bernard, M.D., E. Wesley Ely, M.D., M.P.H., and Pratik P. Pandharipande, M.D., M.S.C.I.

The authors' affiliations are as follows: the Critical Illness, Brain Dysfunction, and Survivorship Center (C.G.H., J.C.J., A.S.H., B.T.P., O.M.O., R.R., J.L.S., A.L.K, M.B.P., N.E.B., T.D.G., R.S.D., G.R.B., E.W.E., P.P.P.), the Center for Health Services Research (C.G.H., J.C.J., R.R., M.B.P., T.D.G., R.S.D., E.W.E., P.P.P.), the Division of Anesthesiology Critical Care Medicine, Department of Anesthesiol-

ogy (C.G.H., P.P.P.), the Division of Allergy, Pulmonary, and Critical Care Medicine (J.C.J., B.T.P., G.R.B., E.W.E.), and the Division of General Internal Medicine and Public Health (R.S.D.), Department of Medicine, the Departments of Biostatistics (O.M.O., R.R.) and Pharmaceutical Services (J.L.S.), and Division of Trauma and Surgical Critical Care, Department of Surgery (M.B.P.), Vanderbilt University Medical Center, and the Anesthesia Service (C.G.H., P.P.P.), Research Service (J.C.J.), Surgical Service (M.B.P.), and Geriatric Research, Education and Clinical Center (R.S.D., E.W.E.), Department of Veterans Affairs Medical Center, Tennessee Valley Healthcare System — both in Nashville; the Neuroscience Institute and Department of Critical Care Medicine, Maine Medical Center, Portland (P.T.M.); the Department of Pharmacy and Health Systems Sciences, Bouve College of Health Sciences, Northeastern University, Boston (J.W.D., M.S.D.); the Departments of Pharmacy (J.T.S., L.N.B.) and Surgery (J.T.S.) and the Center for Outcomes Research (J.T.S.), Houston Methodist, and the Pulmonary, Critical Care and Sleep Medicine Section, Ben Taub Hospital, Baylor College of Medicine (K.K.G.), Houston; the Division of Pulmonary/Critical Care Medicine, University of Texas Health, and the South Texas Veterans Health Care System, San Antonio (A.A.); and Texas Health Harris Methodist Hospital Fort Worth, Fort Worth (A.S.) — all in Texas; the University of Sydney, and the Department of Anaesthetics, Royal Prince Alfred Hospital, Sydney, and the Department of Anesthesiology, University of Wisconsin, Madison (R.D.S.); Pulmonary and Critical Care Medicine, Baton Rouge General Medical Center and Our Lady of the Lake Regional Medical Center, Baton Rouge, LA (H.R.O.); the Department of Anesthesia and Perioperative Care, University of California, San Francisco, San Francisco (M.A.G.); Pulmonary and Critical Care, Mission Hospital, Asheville, NC (G.J.S.); the Division of Pulmonary, Critical Care and Sleep Medicine, Department of Internal Medicine, Ohio State University Wexner Medical Center, Columbus (N.E.B.); and Clinical Research, Investigation, and Systems Modeling of Acute Illness Center, Department of Critical Care Medicine, University of Pittsburgh School of Medicine, Pittsburgh (T.D.G.).

REFERENCES

1. Fleischmann C, Scherag A, Adhikari NK, et al. Assessment of global incidence and mortality of hospital-treated sepsis: current estimates and limitations. Am J Respir Crit Care Med 2016;193:259-72.
2. Dhital R, Basnet S, Poudel DR. Predictors and outcome of invasive mechanical ventilation in hospitalized patients with sepsis: data from National Inpatient Sample. J Community Hosp Intern Med Perspect 2018;8:49-52.
3. Vincent J-L, Marshall JC, Namendys-Silva SA, et al. Assessment of the worldwide burden of critical illness: the Intensive Care Over Nations (ICON) audit. Lancet Respir Med 2014;2:380-6.
4. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. Lancet 2014; 383:911-22.
5. Ely EW, Shintani A, Truman B, et al. Delirium as a predictor of mortality in mechanically ventilated patients in the intensive care unit. JAMA 2004;291:1753-62.
6. Shehabi Y, Riker RR, Bokesch PM, Wisemandle W, Shintani A, Ely EW. Delirium duration and mortality in lightly sedated, mechanically ventilated intensive care patients. Crit Care Med 2010;38: 2311-8.
7. Pisani MA, Kong SYJ, Kasl SV, Murphy TE, Araujo KLB, Van Ness PH. Days of delirium are associated with 1-year mortality in an older intensive care unit population. Am J Respir Crit Care Med 2009;180:1092-7.
8. Schuler A, Wulf DA, Lu Y, et al. The impact of acute organ dysfunction on long-term survival in sepsis. Crit Care Med 2018; 46:843-9.
9. Pandharipande PP, Girard TD, Jackson JC, et al. Long-term cognitive impairment after critical illness. N Engl J Med 2013;369:1306-16.
10. Saczynski JS, Marcantonio ER, Quach L, et al. Cognitive trajectories after postoperative delirium. N Engl J Med 2012; 367:30-9.
11. Helmy SA, Al-Attiyah RJ. The immunomodulatory effects of prolonged intravenous infusion of propofol versus midazolam in critically ill surgical patients. Anaesthesia 2001;56:4-8.
12. Wang X, Cheng Y, Liu X, Yang J, Munoz D, Zhang C. Unexpected pro-injury effect of propofol on vascular smooth muscle cells with increased oxidative stress. Crit Care Med 2011;39:738-45.
13. Kelbel I, Koch T, Weber A, Schiefer HG, van Ackern K, Neuhof H. Alterations of bacterial clearance induced by propofol. Acta Anaesthesiol Scand 1999;43:71-6.
14. Memiş D, Hekimoğlu S, Vatan I, Yandim T, Yüksel M, Süt N. Effects of midazolam and dexmedetomidine on inflammatory responses and gastric intramucosal pH to sepsis, in critically ill patients. Br J Anaesth 2007;98:550-2.
15. Venn RM, Bryant A, Hall GM, Grounds RM. Effects of dexmedetomidine on adrenocortical function, and the cardiovascular, endocrine and inflammatory responses in post-operative patients needing sedation in the intensive care unit. Br J Anaesth 2001; 86:650-6.
16. Qiao H, Sanders RD, Ma D, Wu X, Maze M. Sedation improves early outcome in severely septic Sprague Dawley rats. Crit Care 2009;13:R136.
17. Nishina K, Akamatsu H, Mikawa K, et al. The effects of clonidine and dexmedetomidine on human neutrophil functions. Anesth Analg 1999;88:452-8.
18. Pandharipande PP, Pun BT, Herr DL, et al. Effect of sedation with dexmedetomidine vs lorazepam on acute brain dysfunction in mechanically ventilated patients: the MENDS randomized controlled trial. JAMA 2007;298:2644-53.
19. Riker RR, Shehabi Y, Bokesch PM, et al. Dexmedetomidine vs midazolam for sedation of critically ill patients: a randomized trial. JAMA 2009;301:489-99.
20. Pandharipande PP, Sanders RD, Girard TD, et al. Effect of dexmedetomidine versus lorazepam on outcome in patients with sepsis: an a priori-designed analysis of the MENDS randomized controlled trial. Crit Care 2010;14:R38.
21. Jakob SM, Ruokonen E, Grounds RM, et al. Dexmedetomidine vs midazolam or propofol for sedation during prolonged mechanical ventilation: two randomized controlled trials. JAMA 2012;307:1151-60.
22. Kawazoe Y, Miyamoto K, Morimoto T, et al. Effect of dexmedetomidine on mortality and ventilator-free days in patients requiring mechanical ventilation with sepsis: a randomized clinical trial. JAMA 2017;317:1321-8.
23. Shehabi Y, Howe BD, Bellomo R, et al. Early sedation with dexmedetomidine in critically ill patients. N Engl J Med 2019; 380:2506-17.
24. Devlin JW, Skrobik Y, Gélinas C, et al. Clinical practice guidelines for the prevention and management of pain, agitation/sedation, delirium, immobility, and sleep disruption in adult patients in the ICU. Crit Care Med 2018;46(9):e825-e873.
25. Chandrasekhar R, Hughes CG, Pun BT, Orun OM, Ely EW, Pandharipande PP. Statistical analysis plan for the Maximizing the Efficacy of Sedation and Reducing Neurological Dysfunction and Mortality in Septic Patients with Acute Respiratory Failure trial. Crit Care Resusc 2020;22:63-71.
26. Sessler CN, Gosnell MS, Grap MJ, et al. The Richmond Agitation-Sedation Scale: validity and reliability in adult intensive care unit patients. Am J Respir Crit Care Med 2002;166:1338-44.
27. Rhodes A, Evans LE, Alhazzani W, et al. Surviving Sepsis Campaign: international guidelines for management of sepsis and septic shock: 2016. Crit Care Med 2017;45: 486-552.
28. Barnes-Daly MA, Phillips G, Ely EW. Improving hospital survival and reducing brain dysfunction at seven California community hospitals: implementing PAD

guidelines via the ABCDEF bundle in 6,064 patients. Crit Care Med 2017;45:171-8.
29. Pun BT, Balas MC, Barnes-Daly MA, et al. Caring for critically ill patients with the ABCDEF bundle: results of the ICU Liberation Collaborative in over 15,000 adults. Crit Care Med 2019;47:3-14.
30. Ely EW, Inouye SK, Bernard GR, et al. Delirium in mechanically ventilated patients: validity and reliability of the confusion assessment method for the intensive care unit (CAM-ICU). JAMA 2001;286:2703-10.
31. Gélinas C, Fillion L, Puntillo KA, Viens C, Fortier M. Validation of the Critical-Care Pain Observation Tool in adult patients. Am J Crit Care 2006;15:420-7.
32. Brandt J, Spencer M, Folstein MF. The telephone interview for cognitive status. Neuropsychiatry Neuropsychol Behav Neurol 1988;1:111-7.
33. Christie JD, Plotkin Biester RC, Taichman DB, et al. Formation and validation of a telephone battery to assess cognitive function in acute respiratory distress syndrome survivors. J Crit Care 2006;21:125-32.
34. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged — the Index of Adl: a standardized measure of biological and psychosocial function. JAMA 1963;185:914-9.
35. Pfeffer RI, Kurosaki TT, Harrah CH Jr, Chance JM, Filos S. Measurement of functional activities in older adults in the community. J Gerontol 1982;37:323-9.
36. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001;33:337-43.
37. Jorm AF. A short form of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): development and cross-validation. Psychol Med 1994;24:145-53.
38. Strøm T, Martinussen T, Toft P. A protocol of no sedation for critically ill patients receiving mechanical ventilation: a randomised trial. Lancet 2010;375:475-80.
39. Olsen HT, Nedergaard HK, Strøm T, et al. Nonsedation or light sedation in critically ill, mechanically ventilated patients. N Engl J Med 2020;382:1103-11.
40. Djaiani G, Silverton N, Fedorko L, et al. Dexmedetomidine versus propofol sedation reduces delirium after cardiac surgery: a randomized controlled trial. Anesthesiology 2016;124:362-8.

REVIEW ARTICLE

Julie R. Ingelfinger, M.D., *Editor*

Insights into Glomerular Filtration and Albuminuria

Thomas Benzing, M.D., and David Salant, M.D.

Chronic kidney diseases affect more than 10% of the world's population, and most cases arise from disorders of the kidney's filtration barrier, which is located within a million microvascular units called glomeruli.[1] Although it has been known for many decades that, in the kidney, glomeruli are the site of plasma ultrafiltration and urine production, both the molecular design and function of the filtration barrier remained elusive until recently.[2,3] Moreover, several decades since the recognition that inhibitors of the renin–angiotensin system are beneficial in reducing proteinuria and slowing the progression of diabetic kidney disease, patients are still at risk for end-stage kidney failure from diabetes and other proteinuric kidney diseases.

Evidence is emerging about the added value of sodium–glucose cotransporter 2 (SGLT2) inhibitors, beyond their glucose-lowering effect, when they are used to treat patients with or without diabetes who have proteinuria and declining kidney function.[4-6] Various mechanisms have been proposed to explain the renoprotective effect of SGLT2 inhibitors,[7] including a reduction in pressure within the glomerular capillaries, with resulting protection of glomerular podocytes, which are the targets of injury in most, if not all, proteinuric kidney diseases. Reduction of the glomerular pressure appears to be mediated by constriction of the afferent arterioles, small vessels that supply the glomerular microcirculation with enormous amounts of blood from the circulation. As discussed below, such observations align closely with our current understanding of the respective roles of glomerular capillary pressure, the glomerular basement membrane (GBM), and podocytes in regulating glomerular permeability to albumin and other proteins.

Kidney function depends on the bulk filtration of large volumes of water and small solutes to clear potential toxins that are derived from intracellular metabolism and gastrointestinal microbial metabolism, as well as to maintain salt and water and acid–base homeostasis. The glomeruli produce as much as 180 liters of glomerular filtrate per day in healthy adults, yet only very small amounts of albumin leak into the urine, the end product, with its much smaller volume.[8] Although estimates of the fraction of albumin in the glomerular filtrate (as compared with in plasma) have varied according to the techniques used to measure it, and some filtered albumin is unquestionably retrieved by tubular reabsorption,[9-11] the amount of plasma proteins that escape with the glomerular filtrate is tiny and depends on the selective permeability of the glomerular filtration barrier.[12]

Diseases that reduce the glomerular capillary surface area available for filtration or that alter the intrinsic permeability of the capillary wall reduce the glomerular filtration rate (GFR). Although downstream compensatory mechanisms maintain the glomerular–tubular balance and regulate fluids, electrolytes, and the acid–base balance at physiologic levels, even small reductions in the GFR are associ-

From Department II of Internal Medicine and Center for Molecular Medicine Cologne, University of Cologne, Faculty of Medicine, and University Hospital Cologne, and the Excellence Cluster CECAD, University of Cologne, Cologne, Germany (T.B.); and the Section of Nephrology, Department of Medicine, Boston Medical Center and Boston University School of Medicine, Boston (D.S.). Address reprint requests to Dr. Benzing at Department II of Internal Medicine and Center for Molecular Medicine Cologne, University Hospital Cologne, Kerpener Str. 62, 50937 Cologne, Germany, or at thomas.benzing@uk-koeln.de.

N Engl J Med 2021;384:1437-46.
DOI: 10.1056/NEJMra1808786

ated with increased cardiovascular morbidity and mortality and reduced overall survival.[1,13,14] Albuminuria, another manifestation of diseases that affect the glomerular capillary wall by altering its selective permeability, is also associated with increased cardiovascular morbidity and mortality, even at levels of urine albumin not generally regarded as pathologic and even in the absence of hypertension and diabetes.[15,16] In this review, we discuss current insights, based on classic studies that defined the size- and charge-selective properties of the glomerular filter,[17] to help explain how the unique structure and composition of the glomerular capillary wall maintain highly selective filtration properties when healthy, and how that changes with kidney disease.

EFFECTS OF PODOCYTE DAMAGE

The capillaries in each of the million or so glomeruli in the human kidneys contain a filtration device. Each filtration device consists of three layers: specialized and fenestrated endothelium that lines the luminal side of the capillary wall; an extracellular matrix–based GBM that contains type IV collagen, laminin-521, and nidogen, as well as sulfated proteoglycans; and podocytes that cover the outer surface of the GBM, closely enveloping the glomerular capillaries through extensions (foot processes) that interdigitate with those of adjacent podocytes (Fig. 1).[18,19] The podocyte foot processes of neighboring cells are separated by filtration slits that are bridged by a membrane-like cell junction, called a slit diaphragm[20]; the foot processes are firmly attached to the GBM by various proteins that lead to cell–matrix adhesion.[21] The intricate structure of podocytes allows for ultrafiltration of the large volumes of fluid and small solutes that are necessary for normal clearance of toxic wastes; albumin and most other plasma protein components are retained in the bloodstream.

The identification of mutations in genes expressed by podocytes as the genetic cause of albuminuria in both familial and sporadic kidney disease has spurred research into podocyte pathobiology and furthered our understanding of the glomerular filtration barrier.[22-28] Such studies started about two decades ago with the identification of the genetic cause of congenital nephrotic syndrome of the Finnish type, a rare autosomal recessive disorder caused by mutations in *NPHS1*.

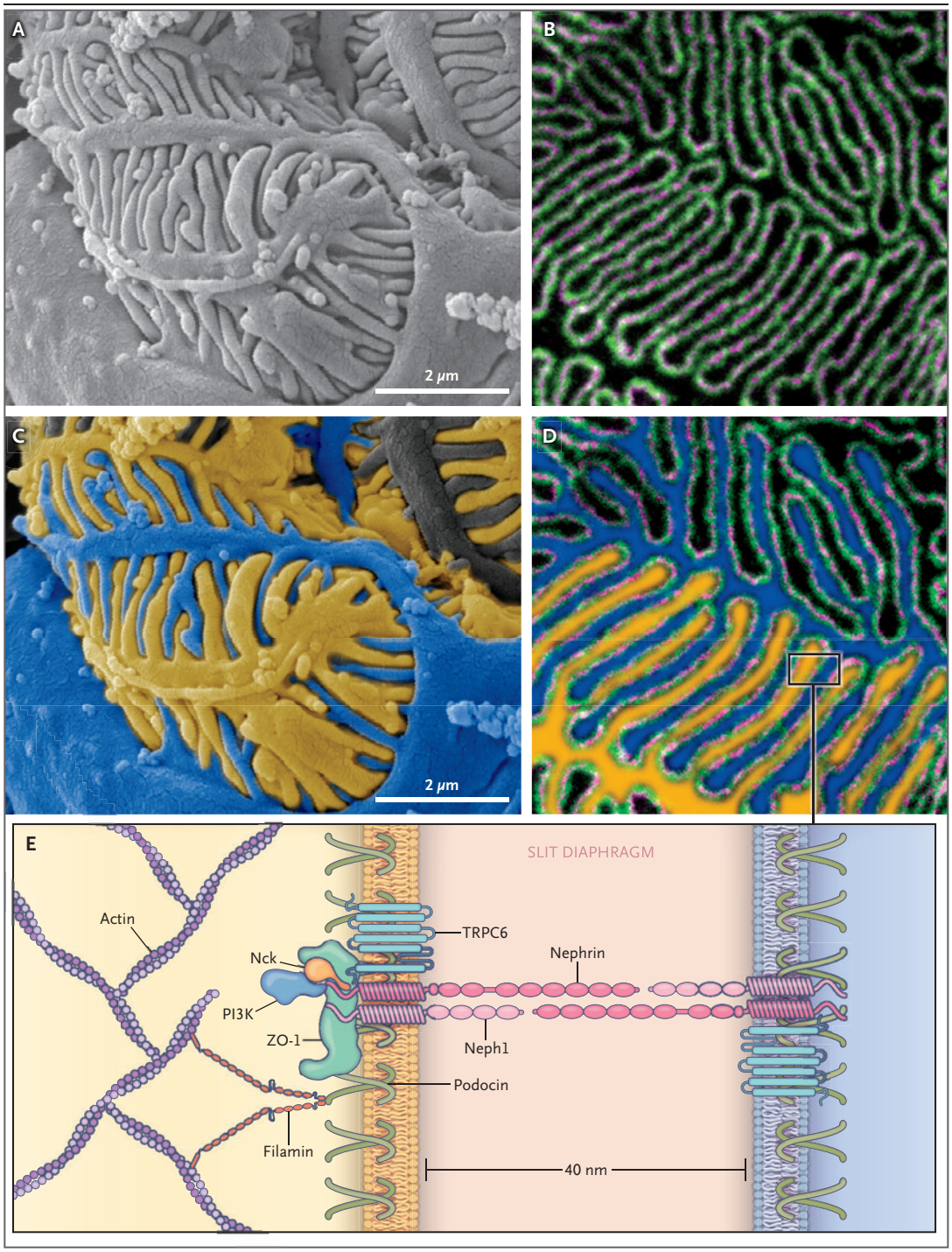
Figure 1 (facing page). Morphologic Features of Podocyte Foot Processes on Ultra-High-Resolution Imaging.

Scanning electron microscopy shows the outer aspect of glomerular capillaries, where plasma ultrafiltration occurs (Panels A and C). Stimulated emission depletion microscopy shows the slit diaphragm connecting adjacent foot processes (Panels B and D) (magenta indicates nephrin, and green, podocin). Color coding of adjacent interdigitating foot processes (Panels C and D) shows the interaction between neighboring podocytes. Also shown is a schematic representation of the slit-diaphragm protein complex that bridges the distance between neighboring foot processes and allows the formation of a filtration slit (Panel E). Nephrin and Neph1 are transmembrane proteins with extracellular domains that connect adjacent foot processes. The cytoplasmic tails of these proteins interact with scaffold proteins such as ZO-1 (zonula occludens 1), signaling adapters such as Nck, and kinases such as phosphatidylinositol 3-kinase (PI3K) to regulate the actin cytoskeleton. The membrane protein podocin clusters at the membrane, interacts with nephrin, and coordinates the protein and lipid environment at the slit-diaphragm protein complex, which renders TRPC6 (transient receptor potential cation channel 6) a mechanosensitive channel.

These mutations result in a severe albuminuria in infants and young children, along with progressive kidney failure.

NPHS1 encodes the immunoglobulin superfamily protein nephrin, a major constituent of the slit diaphragm (Fig. 1). Nephrin molecules bridge the distance between two adjacent foot processes to form a 40-nm membrane-like cell junction.[22,29,30] Part of a large multiprotein complex at the filtration slit (Fig. 1), nephrin recruits adaptor proteins to induce signaling to the podocyte cytoskeleton.[31-35] It is now clear that nephrin-based protein interactions, which are essential for shaping the unique podocyte ultrastructure, mediate signal transduction by responding to mechanical cues and controlling cytoskeletal rearrangements in podocytes (Fig. 1). Moreover, podocin, a product of *NPHS2*, has been shown to interact with nephrin at the slit diaphragm[31] and to organize the lipid environment of the slit-diaphragm complex as a mechanosensor at the filtration slit that also contains ion channels.[26,27,33] Podocin is the most commonly mutated protein in children and adolescents who have "steroid-resistant" nephrotic syndrome (nephrotic syndrome that does not remit with glucocorticoid therapy).

A number of additional podocyte-expressed genes have been identified that, when mutated, cause albuminuria, including the cytoskeletal



genes *ACTN4* and *INF2*[25,36]; these observations are consistent with the critical role of the actin cytoskeleton of podocytes in maintaining the foot-process architecture and the integrity of the glomerular filtration barrier. Studies of these mutations and the resultant mutant proteins have clearly indicated that podocyte injury can cause albuminuria. Moreover, numerous acquired diseases, including diabetic nephropathy, minimal change disease, focal and segmental glomerulosclerosis, membranous nephropathy, hypertensive kidney disease, human immunodeficiency virus–associated nephropathy, and lupus nephritis, also affect podocytes, causing dysfunction of

the filtration barrier and albuminuria. When podocytes are injured, the intercellular junctions and cytoskeletal structure of the foot processes are altered, and the cells are characterized by a simplified architecture, called foot-process effacement.[37,38] These changes are, in principle, reversible, underlining the dynamic structure of podocytes. However, podocytes are postmitotic cells and have a very limited capacity for self-renewal.[39-42] Thus, podocyte loss, whether due to detachment or cell death, results in irreversible damage and scarring of the renal filtration units.[43]

The hypothesis that podocyte loss is a culprit in the development of glomerulosclerosis was formulated more than 30 years ago[39,41] and has subsequently been proved both experimentally and clinically.[44-47] Among persons with steroid-resistant nephrotic syndrome, mutations have also been identified in genes encoding mitochondrial proteins, which lead to mitochondrial dysfunction and impaired respiratory enzyme activity.[48] Such mutations have similarly been observed in a mouse model of proteinuria in which oxygen free radical damage occurs in podocytes.[49] Although numerous mutations involving podocyte proteins have been identified — a list that keeps growing as technological advances are made and more genes are found to modulate the function of podocytes[50] — most forms of podocyte injury are acquired and of these, many are antibody-mediated.

EFFECTS OF PODOCYTE AUTOIMMUNITY

Some of the earliest examples of acquired podocyte autoimmunity were derived from studies in Heymann nephritis, a model of membranous nephropathy in rats in which circulating antibodies bind to the target antigen, megalin, in coated pits on the soles of podocyte foot processes, where they activate complement and cause morphologic changes that are characteristic of human membranous nephropathy. These changes include foot-process effacement, slit-diaphragm dislocation, severe proteinuria, and generation of reactive oxygen species, with disorganization of the GBM through new matrix production and lipid peroxidation of type IV collagen.[51,52] The antigen in most cases of human membranous nephropathy was subsequently identified and was shown to be the target of circulating autoantibodies to the M-type phospholipase A_2 receptor (PLA_2R). PLA_2R is expressed on human podocytes and is shed along with anti-PLA_2R autoantibodies to form subepithelial immune deposits.[53] A growing list of additional podocyte target antigens have subsequently been identified in anti-PLA_2R antibody–negative cases of membranous nephropathy.[54-57] Though much less common than anti-PLA_2R antibodies, these antibodies lead to the same or very similar pathological features and are manifested clinically as nephrotic syndrome or severe albuminuria.

In addition to autoantibodies to podocyte antigens as a cause of glomerulopathy, there are two unusual but highly informative examples of glomerulopathies caused by alloantibodies directed at podocyte proteins. In babies with a truncating mutation of *NPHS1* (Fin-major), the slit-diaphragm protein nephrin is absent and end-stage kidney failure develops early in life as a result. When such patients receive a kidney transplant, nephrotic syndrome sometimes recurs. However, the mechanism is different from that of congenital nephrotic syndrome. In patients in whom nephrin was never expressed, the syndrome is due to the development of antinephrin alloantibodies directed at a neoantigen in the transplanted kidney.[58,59] This observation was recapitulated in a rodent model by injecting antibodies directed at the extracellular region of nephrin.[60]

A second example of alloimmune nephropathy involving a podocyte antigen was described in babies born with nephrotic syndrome whose mothers had a deficiency of neutral endopeptidase (NEP) that was due to sensitization in previous pregnancies with a NEP-positive partner.[61,62] Transplacental passage of the maternal IgG anti-NEP antibodies bound NEP on the fetal podocytes and induced membranous nephropathy in the neonate, manifested as severe proteinuria. Podocyte injury with simplification of the foot processes and secondary changes in the GBM is common to all these conditions.

Although such studies clearly support the critical role of podocytes in maintaining a functional kidney filtration barrier, defects in the GBM, as well as injury to glomerular endothelial cells, can also cause albuminuria, reinforcing the concept that all three layers of the filtration barrier are required for permselective glomerular ultrafiltration. The contribution of the GBM may be exemplified by the fact that mutation of components of laminin-521 in Pierson's syndrome,

an inherited mutation in the laminin β_2 chain,[63] as well as mutations in the α_3, α_4, and α_5 chains of type IV collagen in Alport's syndrome,[64] results in albuminuria and progressive kidney disease. Moreover, damage to the glomerular endothelium can also cause albuminuria. For example, in preeclampsia, interference in vascular endothelial growth factor (VEGF) signaling to the glomerular endothelial cells causes albuminuria and nephrotic syndrome.[65] Preeclampsia, which affects 5 to 10% of pregnant women in the United States, is a complex hypertensive disease characterized by overexpression of soluble fms-like tyrosine kinase 1 (sFlt-1), a soluble VEGF receptor that binds and neutralizes VEGF. The resultant lack of VEGF leads to maternal vascular dysfunction and organ damage.[66,67] Similarly, anti-VEGF therapy with bevacizumab in patients with cancer can cause albuminuria, hypertension, and glomerular disease.[68,69]

A BIOPHYSICAL MODEL OF GLOMERULAR ULTRAFILTRATION

Despite decades of research on the glomerular filtration barrier, a biophysical model to explain how the kidney filter allows extensive fluid filtration while restricting the sieving of macromolecules was lacking until relatively recently.[12,70] Several decades ago, studies with electron microscopy that localized different tracers of the size of albumin or larger indicated an important role of the GBM in retaining proteins in plasma while allowing free filtration of water and solutes, since the tracers did not enter the GBM but instead were restricted to its subendothelial surface.[71,72] Damage to podocytes mediated by puromycin, an antibiotic that inhibits protein synthesis and is used to study models of proteinuria, resulted in consecutive penetration of the tracers into the GBM and uptake by podocytes.[73,74] In contrast, other injected tracers appeared to pass through the GBM but were impeded at the level of the podocyte slit diaphragm, an observation that led to the conclusion that slit diaphragms are the primary barrier of the selective filter.[75,76] For decades, the controversy over control of filtration could not be resolved, and the interpretations based on a coarse filter at the GBM followed by a fine filter at the slit diaphragm did not explain why the glomerulus does not clog with partially filtered proteins.[77]

Given the abundance of evidence that podocyte injury underlies most, if not all, proteinuric kidney diseases, new technologies, including ultrahigh-resolution imaging and genetically engineered mouse models of human disease, were used to examine the glomerular filtration barrier under conditions not previously possible with ultrastructural tracers and conventional light and fluorescence microscopy. These advances led to the development of an experimentally validated biophysical model of glomerular ultrafiltration.[78] Filtration across the glomerular capillary is driven by a net filtration pressure of roughly 20 mm Hg, derived from a hydrostatic pressure gradient of about 40 mm Hg minus the oncotic pressure of the plasma (about 24 mm Hg as blood enters the glomerular capillary), which acts to restrain filtration (Fig. 2).[80] The remarkable luminal pressure exerts physical forces on

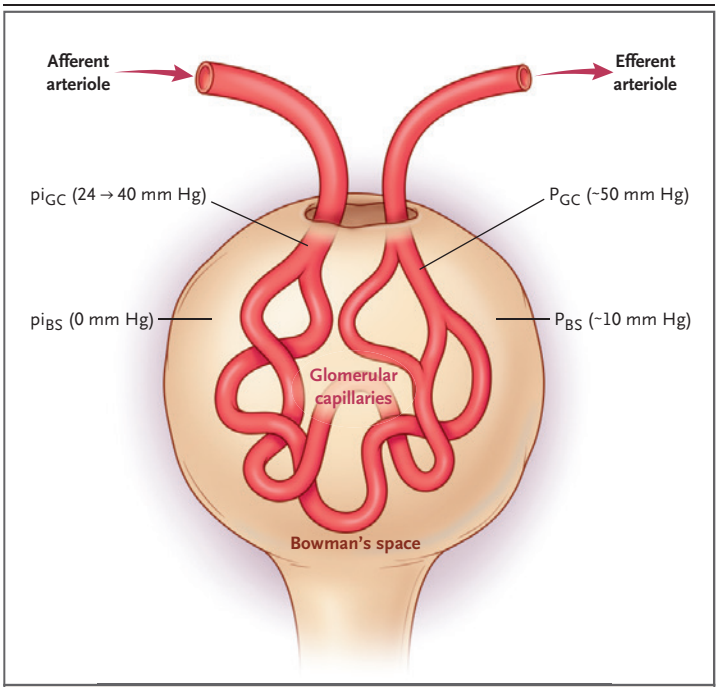


Figure 2. Pressure Gradients Driving and Inhibiting Kidney Filtration.
Filtration across the glomerular capillary is driven by a hydrostatic pressure gradient of about 40 mm Hg (the difference between glomerular capillary pressure [P_{GC}] of about 50 mm Hg and the Bowman's space hydrostatic pressure [P_{BS}] of 10 mm Hg), minus the oncotic pressure of the capillary plasma (pi_{GC}) (about 24 mm Hg as blood enters the glomerular capillary), which acts to restrain filtration. The luminal pressure exerts physical forces on the capillary wall that are counteracted by the glomerular basement membrane (GBM) and by podocytes. The pi_{GC} starts off at the value of normal arterial blood and rises as ultrafiltration removes fluid from the capillary. The oncotic pressure in Bowman's space (pi_{BS}) is constantly close to 0 mm Hg. Adapted from Giebisch and Windhager.[79]

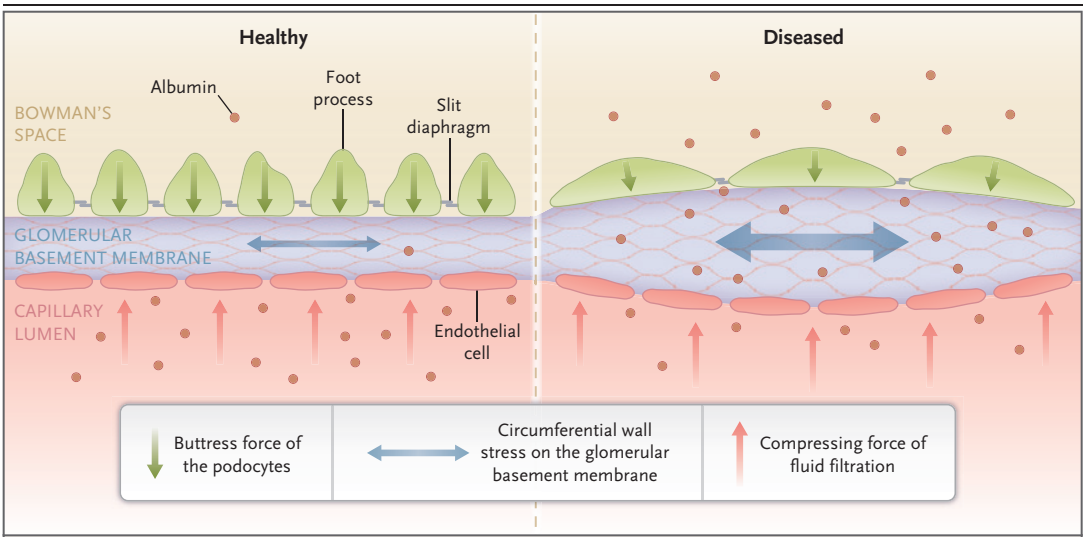


Figure 3. Gel Compression Model of Glomerular Ultrafiltration.

Interdigitating podocyte foot processes counteract the physical forces of hydrostatic pressure in the glomerular capillaries, compressing the gel-like structure of the GBM. The altered biophysical properties of the GBM act as a selectively permeable filter. When podocytes are injured, they take on a more simplified architecture and the slit-diaphragm length is greatly reduced. Concomitantly, the buttressing force of podocytes is lost and the compression force of the GBM is reduced, which increases the permeability to albumin. Adapted from Butt et al.[78]

the capillary wall that are counteracted by the GBM and by podocytes. Specifically, interdigitating podocyte foot processes serve as buttresses[81] against the physical forces of hydrostatic pressure in the glomerular capillaries, compressing the gel-like structure of the GBM (Fig. 3).[82,83] With these altered biophysical properties, the GBM acts as a permselective filter[78] and restricts the permeability to macromolecules transported by diffusion and bulk flow. Thus, the sophisticated foot-process architecture of podocytes not only maximizes the area available for the filtration of water and small solutes but also provides the mechanical resistance against blood pressure that compresses the GBM, preserving permselectivity and preventing loss of albumin and other macromolecules (Fig. 4).[78]

When podocytes are injured, they take on a more simplified architecture and the slit-diaphragm length is much reduced, resulting in a reduction in the filtration slit area and a decline in the glomerular filtration rate of water and small solutes (Fig. 4). Concomitantly, the buttressing force of podocytes is lost and the compressive force of the GBM is reduced, which increases the permeability to albumin (Fig. 3). This construct explains the conundrum of how the GFR may decline while permeability to albumin is increased, a phenomenon elegantly studied and documented in humans with proteinuric kidney disease.[84] Although the contribution of additional factors, such as electrokinetic forces at the GBM[85] and a repelling function of the charged glycocalyx of endothelial cells,[86] may also play a role, the biophysical model explains how the glomerular filter optimizes hydraulic conductivity for the filtration of enormous amounts of fluid by maximizing the filtration area (defined by the length of the filtration slit) while retarding passage of proteins through compression of the GBM.

These new data concerning glomerular filtration underscore the importance of regulated glomerular hemodynamics and have fundamental clinical implications beyond a better understanding of the beneficial effects of angiotensin-converting–enzyme inhibitors or angiotensin-receptor blockers. The length of the slit diaphragm is markedly reduced in early albuminuric disease.[78] Since the width of the filtration slit is thought to be fixed and determined by the interacting molecules that bridge the distance between adjacent foot processes, shortening the filtration slit appears to result in a reduction of the filtration area. In this scenario, the filtration rate is at least partially maintained by angioten-

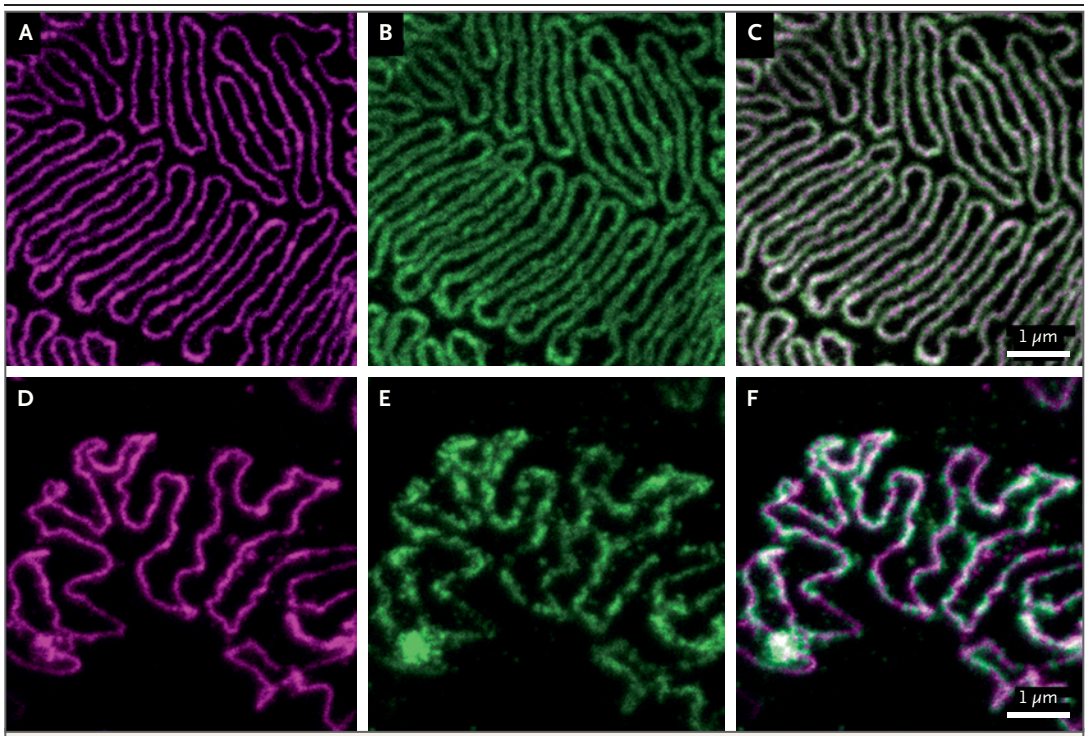


Figure 4. Damaged Podocytes Characterized by Rounded Processes and a Shortened Slit Diaphragm.

The sophisticated foot-process architecture of podocytes not only maximizes the area available for the filtration of water and small solutes but also provides the mechanical resistance against blood pressure that allows the compressed GBM to maintain selective permeability. The structure is lost in glomerular disease, resulting in a shortened slit diaphragm. Panels A, B, and C show the morphologic features of the slit diaphragm in a healthy state (in wild-type mice), and Panels D, E, and F show the altered morphologic features early in the course of the disease (with the *Nphs2*$^{R231Q/A286V}$ mutation). Magenta in Panels A and D indicates nephrin, and green in Panels B and E indicates podocin, with the overlaid colors shown in Panels C and F.

sin II–mediated contraction of the efferent arteriole, which has detrimental effects that offset the benefits of maintaining the GFR. First, the increased capillary pressure cannot be fully counteracted by the defective podocytes, which leads to an increase in proteinuria and, potentially, further injury. Second, maintaining the GFR while the filtration area is decreased drastically increases local fluid flow at the barrier, which exposes podocytes to considerable transverse shear stress and leads to loss of podocytes through detachment, as well as potential scarring of the glomeruli.[43,87] Preventing angiotensin II–mediated constriction of the efferent arteriole by blockade of the renin–angiotensin system is the cornerstone of antiproteinuric therapy to limit progressive podocyte injury and loss in diabetic and nondiabetic kidney disease.

However, hyperfiltration also occurs through loss of regulation of the afferent arteriole. Several studies have shown the mitigating effect of SGLT2 inhibitors on renal outcomes such as progression to end-stage kidney disease, doubling of the serum creatinine level, or death from renal causes in patients with diabetic (and potentially those with nondiabetic) kidney disease,[4,5,88] an effect that is thought to be primarily mediated through constriction of the afferent arteriole and prevention of hyperfiltration.[7] SGLT2 inhibition reduces reabsorption of glucose and sodium within the proximal tubule, which reestablishes sodium delivery to the macula densa and leads to a correction of hyperfiltration through tubuloglomerular feedback and afferent vasoconstriction.[89] Dysfunctional podocytes cannot sufficiently counteract elevated glomerular capillary pressure, suggesting that SGLT2-mediated afferent arteriole vasoconstriction may be beneficial (Fig. 2). The effect of SGLT2 inhibitors appears to be consistent across all levels of kid-

ney function, down to an estimated GFR of 30 ml per minute per 1.73 m^2 of body-surface area, whereas glucose-lowering effects are directly proportional to glomerular filtration and are substantially decreased when kidney function declines,[90] underscoring the importance of regulating glomerular hemodynamics in progressive renal disease.

CONCLUSIONS

Our understanding of the function of the glomerular capillary filter and the mechanisms underlying albuminuria has evolved during the past 20 years. After decades of research, there is now an opportunity to develop mechanism-based therapies that regulate glomerular hemodynamics, on the one hand, and protect mechanically sensitive podocytes, on the other hand, to prevent the progression of chronic kidney disease.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Martin Höhne for providing earlier versions of the figures and David Unnersjö-Jess and Linus Butt for providing the super-resolution images obtained with stimulated emission depletion microscopy.

REFERENCES

1. GBD Chronic Kidney Disease Collaboration. Global, regional, and national burden of chronic kidney disease, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2020;395:709-33.
2. Rhodin J. Electron microscopy of the glomerular capillary wall. Exp Cell Res 1955;8:572-4.
3. Deen WM, Lazzara MJ, Myers BD. Structural determinants of glomerular permeability. Am J Physiol Renal Physiol 2001;281:F579-F596.
4. Perkovic V, Jardine MJ, Neal B, et al. Canagliflozin and renal outcomes in type 2 diabetes and nephropathy. N Engl J Med 2019;380:2295-306.
5. Wanner C, Inzucchi SE, Lachin JM, et al. Empagliflozin and progression of kidney disease in type 2 diabetes. N Engl J Med 2016;375:323-34.
6. Heerspink HJL, Stefánsson BV, Correa-Rotter R, et al. Dapagliflozin in patients with chronic kidney disease. N Engl J Med 2020;383:1436-46.
7. Heerspink HJL, Perkins BA, Fitchett DH, Husain M, Cherney DZI. Sodium glucose cotransporter 2 inhibitors in the treatment of diabetes mellitus: cardiovascular and kidney effects, potential mechanisms, and clinical applications. Circulation 2016;134:752-72.
8. Brinkkoetter PT, Ising C, Benzing T. The role of the podocyte in albumin filtration. Nat Rev Nephrol 2013;9:328-36.
9. Oken DE, Flamenbaum W. Micropuncture studies of proximal tubule albumin concentrations in normal and nephrotic rats. J Clin Invest 1971;50:1498-505.
10. Tojo A, Endou H. Intrarenal handling of proteins in rats using fractional micropuncture technique. Am J Physiol 1992;263:F601-F606.
11. Russo LM, Sandoval RM, McKee M, et al. The normal kidney filters nephrotic levels of albumin retrieved by proximal tubule cells: retrieval is disrupted in nephrotic states. Kidney Int 2007;71:504-13.
12. Haraldsson B, Nyström J, Deen WM. Properties of the glomerular barrier and mechanisms of proteinuria. Physiol Rev 2008;88:451-87.
13. Hallan SI, Matsushita K, Sang Y, et al. Age and association of kidney measures with mortality and end-stage renal disease. JAMA 2012;308:2349-60.
14. Tonelli M, Muntner P, Lloyd A, et al. Risk of coronary events in people with chronic kidney disease compared with those with diabetes: a population-level cohort study. Lancet 2012;380:807-14.
15. Arnlöv J, Evans JC, Meigs JB, et al. Low-grade albuminuria and incidence of cardiovascular disease events in nonhypertensive and nondiabetic individuals: the Framingham Heart Study. Circulation 2005;112:969-75.
16. Fox CS, Matsushita K, Woodward M, et al. Associations of kidney disease measures with mortality and end-stage renal disease in individuals with and without diabetes: a meta-analysis. Lancet 2012;380:1662-73.
17. Chang RL, Deen WM, Robertson CR, et al. Permselectivity of the glomerular capillary wall: studies of experimental glomerulonephritis in the rat using neutral dextran. J Clin Invest 1976;57:1272-86.
18. Kriz W. Progressive renal failure — inability of podocytes to replicate and the consequences for development of glomerulosclerosis. Nephrol Dial Transplant 1996;11:1738-42.
19. Kriz W. Ontogenetic development of the filtration barrier. Nephron Exp Nephrol 2007;106(2):e44-e50.
20. Assady S, Benzing T, Kretzler M, Skorecki KL. Glomerular podocytes in kidney health and disease. Lancet 2019;393:856-8.
21. Lennon R, Randles MJ, Humphries MJ. The importance of podocyte adhesion for a healthy glomerulus. Front Endocrinol (Lausanne) 2014;5:160.
22. Kestilä M, Lenkkeri U, Männikkö M, et al. Positionally cloned gene for a novel glomerular protein — nephrin — is mutated in congenital nephrotic syndrome. Mol Cell 1998;1:575-82.
23. Boute N, Gribouval O, Roselli S, et al. NPHS2, encoding the glomerular protein podocin, is mutated in autosomal recessive steroid-resistant nephrotic syndrome. Nat Genet 2000;24:349-54.
24. Shih NY, Li J, Karpitskii V, et al. Congenital nephrotic syndrome in mice lacking CD2-associated protein. Science 1999;286:312-5.
25. Kaplan JM, Kim SH, North KN, et al. Mutations in ACTN4, encoding alpha-actinin-4, cause familial focal segmental glomerulosclerosis. Nat Genet 2000;24:251-6.
26. Reiser J, Polu KR, Möller CC, et al. TRPC6 is a glomerular slit diaphragm-associated channel required for normal renal function. Nat Genet 2005;37:739-44.
27. Winn MP, Conlon PJ, Lynn KL, et al. A mutation in the TRPC6 cation channel causes familial focal segmental glomerulosclerosis. Science 2005;308:1801-4.
28. Hinkes B, Wiggins RC, Gbadegesin R, et al. Positional cloning uncovers mutations in PLCE1 responsible for a nephrotic syndrome variant that may be reversible. Nat Genet 2006;38:1397-405.
29. Holthöfer H, Ahola H, Solin ML, et al. Nephrin localizes at the podocyte filtration slit area and is characteristically spliced in the human kidney. Am J Pathol 1999;155:1681-7.
30. Holzman LB, St John PL, Kovari IA, Verma R, Holthofer H, Abrahamson DR. Nephrin localizes to the slit pore of the glomerular epithelial cell. Kidney Int 1999;56:1481-91.
31. Huber TB, Kottgen M, Schilling B, Walz G, Benzing T. Interaction with podocin facilitates nephrin signaling. J Biol Chem 2001;276:41543-6.
32. Jones N, Blasutig IM, Eremina V, et al. Nck adaptor proteins link nephrin to the actin cytoskeleton of kidney podocytes. Nature 2006;440:818-23.
33. Huber TB, Schermer B, Müller RU, et al. Podocin and MEC-2 bind cholesterol

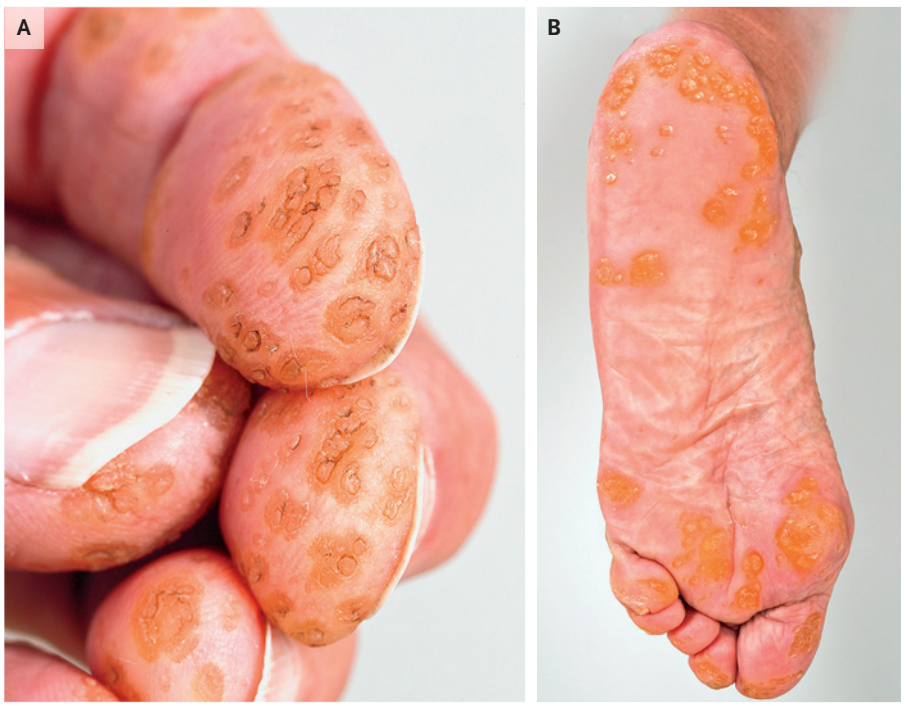
to regulate the activity of associated ion channels. Proc Natl Acad Sci U S A 2006; 103:17079-86.
34. Verma R, Wharram B, Kovari I, et al. Fyn binds to and phosphorylates the kidney slit diaphragm component Nephrin. J Biol Chem 2003;278:20716-23.
35. Fan X, Li Q, Pisarek-Horowitz A, et al. Inhibitory effects of Robo2 on nephrin: a crosstalk between positive and negative signals regulating podocyte structure. Cell Rep 2012;2:52-61.
36. Brown EJ, Schlöndorff JS, Becker DJ, et al. Mutations in the formin gene INF2 cause focal segmental glomerulosclerosis. Nat Genet 2010;42:72-6.
37. Kriz W, Endlich K. Podocytes and disease: introduction. Semin Nephrol 2012; 32:305-6.
38. Mundel P, Shankland SJ. Podocyte biology and response to injury. J Am Soc Nephrol 2002;13:3005-15.
39. Kriz W, Gretz N, Lemley KV. Progression of glomerular diseases: is the podocyte the culprit? Kidney Int 1998;54:687-97.
40. Kriz W, Lemley KV. The role of the podocyte in glomerulosclerosis. Curr Opin Nephrol Hypertens 1999;8:489-97.
41. Fries JW, Sandstrom DJ, Meyer TW, Rennke HG. Glomerular hypertrophy and epithelial cell injury modulate progressive glomerulosclerosis in the rat. Lab Invest 1989;60:205-18.
42. Pabst R, Sterzel RB. Cell renewal of glomerular cell types in normal rats: an autoradiographic analysis. Kidney Int 1983; 24:626-31.
43. Lemley KV. Mechanical challenges to the glomerulus and podocyte loss: evolution of a paradigm. Pflugers Arch 2017; 469:959-63.
44. Wharram BL, Goyal M, Wiggins JE, et al. Podocyte depletion causes glomerulosclerosis: diphtheria toxin-induced podocyte depletion in rats expressing human diphtheria toxin receptor transgene. J Am Soc Nephrol 2005;16:2941-52.
45. Ding F, Wickman L, Wang SQ, et al. Accelerated podocyte detachment and progressive podocyte loss from glomeruli with age in Alport Syndrome. Kidney Int 2017;92:1515-25.
46. Nishizono R, Kikuchi M, Wang SQ, et al. FSGS as an adaptive response to growth-induced podocyte stress. J Am Soc Nephrol 2017;28:2931-45.
47. Hara M, Yanagihara T, Kihara I. Cumulative excretion of urinary podocytes reflects disease progression in IgA nephropathy and Schönlein-Henoch purpura nephritis. Clin J Am Soc Nephrol 2007;2: 231-8.
48. Ashraf S, Gee HY, Woerner S, et al. ADCK4 mutations promote steroid-resistant nephrotic syndrome through CoQ10 biosynthesis disruption. J Clin Invest 2013; 123:5179-89.
49. Binder CJ, Weiher H, Exner M, Kerjaschki D. Glomerular overproduction of oxygen radicals in Mpv17 gene-inactivated mice causes podocyte foot process flattening and proteinuria: a model of steroid-resistant nephrosis sensitive to radical scavenger therapy. Am J Pathol 1999;154: 1067-75.
50. Warejko JK, Tan W, Daga A, et al. Whole exome sequencing of patients with steroid-resistant nephrotic syndrome. Clin J Am Soc Nephrol 2018;13:53-62.
51. Francis JM, Beck LH Jr, Salant DJ. Membranous nephropathy: a journey from bench to bedside. Am J Kidney Dis 2016; 68:138-47.
52. Neale TJ, Ojha PP, Exner M, et al. Proteinuria in passive Heymann nephritis is associated with lipid peroxidation and formation of adducts on type IV collagen. J Clin Invest 1994;94:1577-84.
53. Beck LH Jr, Bonegio RGB, Lambeau G, et al. M-type phospholipase A2 receptor as target antigen in idiopathic membranous nephropathy. N Engl J Med 2009; 361:11-21.
54. Tomas NM, Beck LH Jr, Meyer-Schwesinger C, et al. Thrombospondin type-1 domain-containing 7A in idiopathic membranous nephropathy. N Engl J Med 2014;371:2277-87.
55. Sethi S, Debiec H, Madden B, et al. Neural epidermal growth factor-like 1 protein (NELL-1) associated membranous nephropathy. Kidney Int 2020;97:163-74.
56. Sethi S, Debiec H, Madden B, et al. Semaphorin 3B-associated membranous nephropathy is a distinct type of disease predominantly present in pediatric patients. Kidney Int 2020;98:1253-64.
57. Caza T, Hassen S, Dvanajscak Z, et al. NELL1 is a target antigen in malignancy-associated membranous nephropathy. Kidney Int 2020 August 20 (Epub ahead of print).
58. Patrakka J, Ruotsalainen V, Reponen P, et al. Recurrence of nephrotic syndrome in kidney grafts of patients with congenital nephrotic syndrome of the Finnish type: role of nephrin. Transplantation 2002;73:394-403.
59. Holmberg C, Jalanko H. Congenital nephrotic syndrome and recurrence of proteinuria after renal transplantation. Pediatr Nephrol 2014;29:2309-17.
60. Topham PS, Kawachi H, Haydar SA, et al. Nephritogenic mAb 5-1-6 is directed at the extracellular domain of rat nephrin. J Clin Invest 1999;104:1559-66.
61. Debiec H, Guigonis V, Mougenot B, et al. Antenatal membranous glomerulonephritis due to anti–neutral endopeptidase antibodies. N Engl J Med 2002;346: 2053-60.
62. Debiec H, Nauta J, Coulet F, et al. Role of truncating mutations in MME gene in fetomaternal alloimmunisation and antenatal glomerulopathies. Lancet 2004;364: 1252-9.
63. Zenker M, Aigner T, Wendler O, et al. Human laminin beta2 deficiency causes congenital nephrosis with mesangial sclerosis and distinct eye abnormalities. Hum Mol Genet 2004;13:2625-32.
64. Barker DF, Hostikka SL, Zhou J, et al. Identification of mutations in the COL4A5 collagen gene in Alport syndrome. Science 1990;248:1224-7.
65. Maynard SE, Min JY, Merchan J, et al. Excess placental soluble fms-like tyrosine kinase 1 (sFlt1) may contribute to endothelial dysfunction, hypertension, and proteinuria in preeclampsia. J Clin Invest 2003;111:649-58.
66. Levine RJ, Maynard SE, Qian C, et al. Circulating angiogenic factors and the risk of preeclampsia. N Engl J Med 2004; 350:672-83.
67. Thadhani R, Hagmann H, Schaarschmidt W, et al. Removal of soluble Fms-like tyrosine kinase-1 by dextran sulfate apheresis in preeclampsia. J Am Soc Nephrol 2016;27:903-13.
68. Eremina V, Sood M, Haigh J, et al. Glomerular-specific alterations of VEGF-A expression lead to distinct congenital and acquired renal diseases. J Clin Invest 2003;111:707-16.
69. Eremina V, Jefferson JA, Kowalewska J, et al. VEGF inhibition and renal thrombotic microangiopathy. N Engl J Med 2008; 358:1129-36.
70. Jeansson M, Haraldsson B. Glomerular size and charge selectivity in the mouse after exposure to glucosaminoglycan-degrading enzymes. J Am Soc Nephrol 2003;14:1756-65.
71. Farquhar MG, Palade GE. Glomerular permeability. II. Ferritin transfer across the glomerular capillary wall in nephrotic rats. J Exp Med 1961;114:699-716.
72. Caulfield JP, Farquhar MG. The permeability of glomerular capillaries to graded dextrans: identification of the basement membrane as the primary filtration barrier. J Cell Biol 1974;63:883-903.
73. Caulfield JP, Farquhar MG. The permeability of glomerular capillaries of aminonuceoside nephrotic rats to graded dextrans. J Exp Med 1975;142:61-83.
74. Farquhar MG, Palade GE. Segregation of ferritin in glomerular protein absorption droplets. J Biophys Biochem Cytol 1960;7:297-304.
75. Graham RC Jr, Karnovsky MJ. Glomerular permeability: ultrastructural cytochemical studies using peroxidases as protein tracers. J Exp Med 1966;124:1123-34.
76. Venkatachalam MA, Karnovsky MJ, Cotran RS. Glomerular permeability: ultrastructural studies in experimental nephrosis using horseradish peroxidase as a tracer. J Exp Med 1969;130:381-99.
77. Smithies O. Why the kidney glomerulus does not clog: a gel permeation/diffusion hypothesis of renal function. Proc Natl Acad Sci U S A 2003;100:4108-13.
78. Butt L, Unnersjö-Jess D, Höhne M, et al. A molecular mechanism explaining

albuminuria in kidney disease. Nat Metab 2020;2:461-74.
79. Giebisch G, Windhager E. Glomerular filtration and renal blood flow. In: Boron WF, Boulpaep EL, eds. Medical physiology: a cellular and molecular approach. 2nd ed. Philadelphia: Elsevier Saunders, 2012:749-81.
80. Kriz W, Lemley KV. Mechanical challenges to the glomerular filtration barrier: adaptations and pathway to sclerosis. Pediatr Nephrol 2017;32:405-17.
81. Salant DJ. The structural biology of glomerular epithelial cells in proteinuric diseases. Curr Opin Nephrol Hypertens 1994;3:569-74.
82. Robinson GB, Walton HA. Glomerular basement membrane as a compressible ultrafilter. Microvasc Res 1989;38:36-48.
83. Fissell WH, Miner JH. What is the glomerular ultrafiltration barrier? J Am Soc Nephrol 2018;29:2262-4.
84. Shemesh O, Ross JC, Deen WM, Grant GW, Myers BD. Nature of the glomerular capillary injury in human membranous glomerulopathy. J Clin Invest 1986;77:868-77.
85. Moeller MJ, Tenten V. Renal albumin filtration: alternative models to the standard physical barriers. Nat Rev Nephrol 2013;9:266-77.
86. van den Berg BM, Wang G, Boels MGS, et al. Glomerular function and structural integrity depend on hyaluronan synthesis by glomerular endothelium. J Am Soc Nephrol 2019;30:1886-97.
87. Friedrich C, Endlich N, Kriz W, Endlich K. Podocytes are sensitive to fluid shear stress in vitro. Am J Physiol Renal Physiol 2006;291:F856-F865.
88. Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. N Engl J Med 2015;373:2117-28.
89. Pollock CA, Lawrence JR, Field MJ. Tubular sodium handling and tubuloglomerular feedback in experimental diabetes mellitus. Am J Physiol 1991;260:F946-F952.
90. Heerspink HJL, Desai M, Jardine M, Balis D, Meininger G, Perkovic V. Canagliflozin slows progression of renal function decline independently of glycemic effects. J Am Soc Nephrol 2017;28:368-75.

IMAGES IN CLINICAL MEDICINE

Chana A. Sacks, M.D., *Editor*

Palmoplantar Papules



A 68-year-old woman presented to the dermatology clinic with a 40-year history of slowly progressing, yellowish, hyperkeratotic papules and plaques on her hands (Panel A) and feet (Panel B). She did not have any pain, itching, nail changes, or skin fragility. Similar skin lesions were present in her mother, son, and granddaughter. Biopsy of the lesions was performed, and histopathological analysis showed orthohyperkeratosis with hypergranulosis. This presentation was consistent with a diagnosis of punctate palmoplantar keratoderma. Next-generation sequencing showed a nonsense mutation (c.370C→T) in *AAGAB*, which encodes alpha and gamma adaptin binding protein, in the patient and her son, and a diagnosis of autosomal dominant hereditary punctate palmoplantar keratoderma type 1 was made. This type of palmoplantar keratoderma type 1 may be associated with certain types of cancer. In this patient, the results of recent chest radiography, upper endoscopy, colonoscopy, Papanicolaou testing, and mammography were unremarkable. Treatment with topical ointments with 40% urea and 20% salicylic acid were initiated, which led to a mild reduction in hyperkeratosis. Regular cancer screening was recommended.

Benjamin Klein, M.D.
Regina Treudler, M.D.

Universitätsklinikum Leipzig
Leipzig, Germany
benjamin.klein@medizin.uni-leipzig.de

DOI: 10.1056/NEJMicm2032907

CASE RECORDS of the MASSACHUSETTS GENERAL HOSPITAL

Case 11-2021: A 39-Year-Old Woman with Fever, Flank Pain, and Inguinal Lymphadenopathy

Richelle C. Charles, M.D., Madeleine Sertic, M.B., Ch.B., Anne M. Neilan, M.D., M.P.H., and Aliyah R. Sohani, M.D.

From the Departments of Medicine (R.C.C., A.M.N.), Radiology (M.S.), Pediatrics (A.M.N.), and Pathology (A.R.S.), Massachusetts General Hospital, and the Departments of Medicine (R.C.C., A.M.N.), Radiology (M.S.), Pediatrics (A.M.N.), and Pathology (A.R.S.), Harvard Medical School — both in Boston.

N Engl J Med 2021;384:1448-56.
DOI: 10.1056/NEJMcpc2100273

PRESENTATION OF CASE

Dr. P. Connor Johnson (Medicine): A 39-year-old woman presented to this hospital with fever, flank pain, and tender inguinal lymphadenopathy.

The patient had been in her usual state of health until approximately 4 weeks before admission, when she noted tender bilateral inguinal swelling. Two days later, she began to have sharp, intermittent flank pain on the right side; she rated the pain at 5 on a scale of 0 to 10, with 10 indicating the most severe pain. During the next 3 days, she had nausea and a poor appetite. She noted foul-smelling urine but had no dysuria, urinary frequency, or hematuria.

Three weeks before admission, the patient was evaluated by her primary care physician. On examination, the temperature was 36.6°C, and she appeared well. There was costovertebral and abdominal tenderness on the right side, without rebound or guarding. The external genitalia were normal. Copious thin, white vaginal discharge was present; no cervical motion tenderness was noted, and the ovaries were normal on palpation. Multiple bilateral tender inguinal lymph nodes, measuring up to 2 cm in diameter, were noted on palpation. The remainder of the physical examination was normal. On microscopic examination of the vaginal discharge, clue cells were observed, but there were no fungal elements or trichomonads. Nucleic acid amplification tests of a cervical swab for *Chlamydia trachomatis* and *Neisseria gonorrhoeae* were negative. A blood test for IgM, IgG, and IgA antibodies to *C. trachomatis* serovars L1 and D through K was negative. Urinalysis revealed leukocyte esterase and nitrates; urine was obtained for culture. Vaginal metronidazole gel was prescribed.

Two days later, the urine culture grew more than 100,000 colony-forming units (CFU) of *Escherichia coli* per milliliter and 10,000 to 100,000 CFU of *Klebsiella pneumoniae* per milliliter. A 14-day course of trimethoprim–sulfamethoxazole was prescribed.

The inguinal swelling reportedly abated, but the flank pain and nausea persisted. Two days after the patient completed the prescribed course of trimethoprim–sulfamethoxazole, fever and night sweats developed. After 3 days of fever with a temperature of up to 38.5°C, she began vomiting and subsequently presented to the emergency department of this hospital for evaluation.

In the emergency department, the patient reported ongoing flank pain, malaise, nausea, and poor appetite with weight loss of 2.3 kg during the past month. There was no sore throat, cough, or diarrhea. She had a history of hypothyroidism, asthma, bipolar disorder, dysmenorrhea, and migraines. She had had multiple urinary tract infections; pyelonephritis had been diagnosed 4 years before this presentation. A copper intrauterine device had been inserted 6 months before this presentation. Medications included albuterol, budesonide–formoterol, divalproex, levothyroxine, loratadine, and sumatriptan as needed for migraines. The patient did not smoke tobacco, drink alcohol, or use illicit drugs. She had immigrated to the United States from Brazil 1 year earlier; she had been living in New England during the past year and had been working in a laboratory. She had recently spent 3 months in Brazil visiting her husband and had returned to New England shortly before she noted inguinal swelling. While in Brazil, she had remained in a large city and had not visited any remote locations. In New England, she lived alone with one kitten and one full-grown cat. There was no family history of cancer. An aunt had died from tuberculosis, but the patient had had only minimal contact with her aunt.

On examination, the temperature was 36.6°C, the blood pressure 100/63 mm Hg, the heart rate 77 beats per minute, the respiratory rate 16 breaths per minute, and the oxygen saturation 100% while the patient was breathing ambient air. The weight was 67.5 kg and the body-mass index (the weight in kilograms divided by the square of the height in meters) 24.4. The patient appeared ill. The right side of the abdomen was tender on palpation, without rebound or guarding. Hepatosplenomegaly was not present. Tender bilateral inguinal lymph nodes were noted. No palpable cervical, supraclavicular, infraclavicular, or axillary lymphadenopathy was noted. The remainder of the physical examination was normal. A urine test for human chorionic gonadotropin was negative. Blood levels of glucose, electrolytes, creatine kinase, and lactate dehydrogenase were normal, as were the results of liver-function and kidney-function tests. Blood tests for human immunodeficiency virus (HIV) type 1 p24 antigen and antibodies to HIV types 1 and 2 were negative, as was an interferon-γ release assay for *Mycobacterium tuberculosis*. Other laboratory test results are shown in Table 1. Imaging studies were obtained.

Table 1. Laboratory Data.

Variable	Reference Range, Adults*	On Admission, This Hospital
Hematocrit (%)	36–46	36
Hemoglobin (g/dl)	12–16	11.6
White-cell count (per μl)	4500–11,000	10,500
Differential count (per μl)		
Neutrophils	1800–7700	5910
Lymphocytes	1000–4800	3190
Monocytes	200–1200	960
Eosinophils	0–900	33
Basophils	0–300	5
Platelet count (per μl)	150,000–400,000	409,000
Mean corpuscular volume (fl)	80–100	85.6
Erythrocyte sedimentation rate (mm/hr)	0–20	41
C-reactive protein (mg/liter)	<8	100.7

* Reference values are affected by many variables, including the patient population and the laboratory methods used. The ranges used at Massachusetts General Hospital are for adults who are not pregnant and do not have medical conditions that could affect the results. They may therefore not be appropriate for all patients.

Dr. Madeleine Sertic: Computed tomography (CT) of the abdomen and pelvis (Fig. 1), performed without the administration of intravenous contrast material, revealed bilateral inguinal lymphadenopathy, with the largest node, on the right side, measuring 2.6 cm in greatest dimension. A right pelvic sidewall lymph node measuring 1.3 cm in greatest dimension, a gastrohepatic lymph node measuring 0.9 cm in greatest dimension, and lymphadenopathy in the porta hepatis were also seen. Transvaginal ultrasonography revealed an intrauterine device in an appropriate position, normal ovaries, and trace free fluid in the pelvis.

Dr. Johnson: Intravenous fluids and intramuscular ketorolac were administered, and the patient was admitted to the hospital. A diagnostic test was performed.

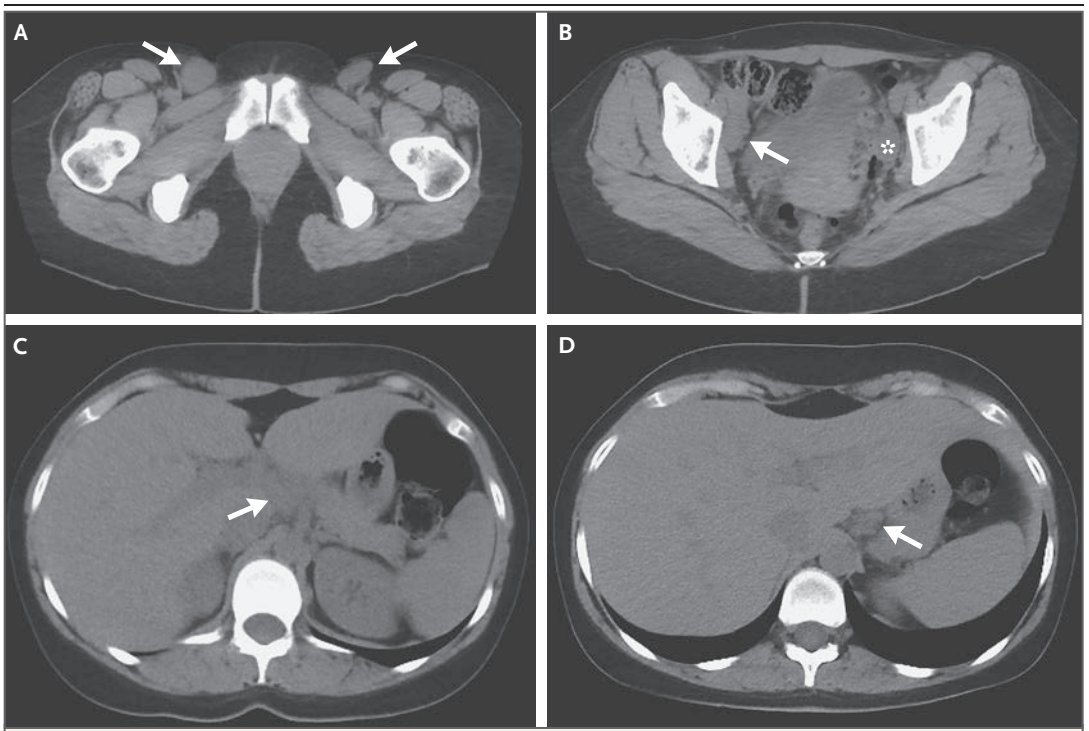


Figure 1. CT of the Abdomen and Pelvis.

Axial images of the abdomen and pelvis (Panels A and B), obtained without the administration of contrast material, show bilateral inguinal lymphadenopathy (Panel A, arrows), which is more prominent on the right side than on the left side, and right external iliac lymphadenopathy (Panel B, arrow) with a normal left ovary (asterisk). Axial images through the level of the upper abdomen (Panels C and D) show lymphadenopathy in the porta hepatis (Panel C, arrow) and a borderline-enlarged lymph node in the gastrohepatic ligament (Panel D, arrow).

DIFFERENTIAL DIAGNOSIS

Dr. Richelle C. Charles: This 39-year-old woman presented to this hospital with tender inguinal lymphadenopathy, flank pain on the right side, and abdominal pain, with systemic symptoms of fever, night sweats, anorexia, and weight loss after a 3-month stay in Brazil.

The causes of lymphadenopathy can be considered in broad disease-based categories, including infection, cancer, autoimmune disease, and other various disorders. The approach to evaluating a patient with lymphadenopathy typically includes determining whether the lymphadenopathy is localized, involving only one region, or whether it is generalized, affecting more than one region. This initial assessment, together with thorough history taking and physical examination to identify epidemiologic clues (e.g., exposure to cats, recent travel, insect bites, or high-risk sexual behavior) or pathologic features distal to the involved lymph nodes, can substantially narrow down the possibilities and, in most cases, lead to a diagnosis.

This patient had inguinal lymphadenopathy that was present predominantly on the right side, with contiguous pelvic wall lymphadenopathy seen on abdominal CT, findings that suggested a regional process. She had inguinal lymphadenopathy on the left side, as well, but the development of lymphadenopathy on both sides of the inguinal region can occur if the underlying process originates from a central site, such as from a sexually transmitted infection or from the unlikely event of inoculation of a pathogen that affects both legs. The abdominal CT also revealed an enlarged gastrohepatic lymph node and lymphadenopathy in the porta hepatis, findings that raise the possibility of a generalized process or early disseminated disease. Because fever and inguinal lymphadenopathy are the predominant features in this case, I will construct my differential diagnosis around this symptom complex (Table 2).

MYCOBACTERIAL INFECTION

Given that this patient had immigrated to the United States from Brazil 1 year earlier, chronic infections or reactivation of latent infections such as tuberculosis or endemic mycosis should be considered. Isolated peripheral tuberculous lymphadenitis is most often due to reactivation disease.[1] Typically, tuberculous lymphadenitis is painless, and the most common site of involvement is the cervical lymph nodes, but other peripheral lymph nodes can be involved.[1] This patient had lymphadenopathy in the porta hepatis; however, if she had abdominal tuberculosis, I would expect mesenteric lymphadenopathy, intestinal wall thickening, or ascites, all of which were absent in this case. Her fever, nausea, vomiting, weight loss, and abdominal pain would be consistent with this diagnosis, as would her elevated erythrocyte sedimentation rate and C-reactive protein level but otherwise normal laboratory test results. Although the patient had a negative interferon-γ release assay for *M. tuberculosis*, a negative test does not necessarily rule out tuberculosis.

Infection due to nontuberculous mycobacteria is also a possible diagnosis in this patient, but nontuberculous mycobacterial lymphadenitis is seen more often in children than in adults and would be rare in an immunocompetent adult.[2] Infection with *M. marinum* can cause a nodular skin infection with accompanying lymphadenopathy and is associated with skin trauma and exposure to a contaminated water source, but isolated lymphadenopathy without associated skin findings would be unusual.

Table 2. Causes of Inguinal Lymphadenopathy.

Infection
Infection of the lower limbs or other localized infection
Bacterial adenitis (staphylococcal or streptococcal)
Tuberculous lymphadenitis
Cat scratch disease (caused by *Bartonella henselae*)
Nocardiosis
Sporotrichosis
Tularemia
Sexually transmitted infection
Lymphogranuloma venereum
Chancroid
Syphilis
Genital herpes
Cancer
Leukemia or lymphoma
Metastatic carcinoma originating from cancer in the leg (e.g., melanoma), vulva, cervix, penis, anus, or rectum
Autoimmune disorder
Systemic lupus erythematosus
Dermatomyositis
Rheumatoid arthritis
Sarcoidosis

BARTONELLA HENSELAE INFECTION

Because the patient has a kitten and a cat, infection with bartonella species — specifically *Bartonella henselae*, the primary causative agent of cat scratch disease — is a consideration in this case.[3] Cats, particularly kittens, are the main reservoir of *B. henselae*; they transmit the bacteria through their saliva or through scratching (approximately 50% of cats are seropositive for *B. henselae*).[4,5] This patient did not have any evidence of scratch marks on examination, and information about previous scratches was not provided in the case presentation. Typically, a skin lesion (i.e., a papule, vesicle, or pustule) forms at the scratch site 3 to 10 days after inoculation; however, such a lesion is not always seen or properly identified.[6] Approximately 2 weeks after inoculation, regional lymphadenopathy proximal to the inoculation site, a characteristic feature of cat scratch disease, develops.[6] The most commonly involved sites of lymphadenopathy are the axillary, epitrochlear, cervical, supraclavicular, and submandibular regions.[6] Generalized lymphadenopathy is rare, but up to a third of cases involve more than one anatomical site.[6] Although lymphadenopathy associated with cat scratch disease usually resolves within 1 to 6 months without therapy,[7] the current recommendation is treatment with a short course of azithromycin.[8] Trimethoprim–sulfamethoxazole, which this patient received, does have in vitro activity against bartonella species and has been shown to be clinically effective in some cases.[7] This could explain why this patient's condition improved initially but later worsened after cessation of these antibiotic agents. The typical duration of treatment is 5 days for uncomplicated cat scratch disease. This patient received 2 weeks of therapy, but given the extent of her disease and involvement of gastrohepatic nodes and nodes in

the porta hepatis — findings that could suggest early dissemination of *B. hensalae* — she probably would have needed a longer course of treatment.

TOXOPLASMOSIS

Toxoplasmosis and cat scratch disease are both associated with adenopathy and exposure to cats. However, acute toxoplasmosis is manifested by a mononucleosis-like syndrome characterized by cervical lymphadenopathy or generalized, diffuse lymphadenopathy[9] that is similar to that seen in patients with acute Epstein–Barr virus infection, acute cytomegalovirus infection, and primary HIV infection. Because this patient had many features that were consistent with acute toxoplasmosis, it is difficult to rule out this diagnosis without additional testing for toxoplasmosis.

BACTERIAL INFECTIONS OF THE SKIN AND SOFT TISSUES

Bacterial adenitis, which is usually caused by *Staphylococcus aureus* and *Streptococcus pyogenes*, is a common type of acute unilateral lymphadenitis; however, the absence of signs of a leg infection makes these diagnoses unlikely. Less common causes of lymphadenitis include tularemia, nocardiosis, and sporotrichosis, all of which can cause regional lymphadenopathy distal to a site of inoculation. Typically, associated skin findings would be expected with tularemia, nocardiosis, and sporotrichosis; however, tularemia also has a glandular form without associated skin findings.[10] However, the epidemiologic features of tularemia (e.g., exposure to an infected animal or an animal bite) do not fit with this patient's presentation.

SEXUALLY TRANSMITTED INFECTIONS

The sexually transmitted infections that are commonly associated with inguinal lymphadenopathy are lymphogranuloma venereum (LGV), syphilis, and chancroid. Of these, LGV is the most likely diagnosis for this patient, given her symptoms, the timing of onset of symptoms, and the absence of skin findings on pelvic examination. LGV is a genital ulcer disease caused by the L1, L2, and L3 serovars of *C. trachomatis*.[11] Historically, it has been most prevalent in the tropics and subtropics, but its prevalence has been increasing in high-income countries, primarily among men who have sex with men.[11] The primary stage of infection is characterized by the presence of a genital ulcer at the site of inoculation that spontaneously heals within a few days. The secondary stage, occurring 2 to 6 weeks later, can be manifested by an enlarged, painful reactive inguinal node and is related to direct extension of the infection to the regional nodes.[11] The secondary stage can also be associated with anorectal symptoms including proctocolitis.[11] The diagnosis is based on the clinical presentation and epidemiologic findings and is confirmed by nucleic acid testing.[11] Serologic testing can also be performed to support the diagnosis. This patient underwent testing for LGV; both a serologic test and a nucleic acid test for *C. trachomatis* were negative.

The absence of genital ulcers in this patient makes chancroid, primary herpes simplex virus infection, and syphilis unlikely.[12] The secondary stage of syphilis is typically manifested by rash and generalized lymphadenopathy; regional lymphadenopathy, as seen in this case, would be atypical of syphilis.

TRAVEL-RELATED INFECTIONS

This patient had recently returned from a 3-month stay in Brazil visiting her husband. Many travel-related illnesses can be associated with fever and lymphadenopathy, including dengue virus infection, chikungunya virus infection, Zika virus infection, brucellosis, leptospirosis, enteric fever (typhoid or paratyphoid fever), and bartonellosis (caused by *B. quintana* and *B. bacilliformis*).[13] With the exception of enteric fever, which would be manifested by mesenteric lymphadenopathy, these illnesses are usually associated with diffuse lymphadenopathy rather than the regional lymphadenopathy seen in this patient. Infection with dengue, chikungunya, and Zika viruses can be ruled out in this case, since patients typically present with fever within 2 weeks after exposure.

Endemic mycoses, such as paracoccidioidomycosis, histoplasmosis, blastomycosis, and coccidioidomycosis, can cause widespread lymphadenopathy in patients with disseminated disease and are usually associated with hepatosplenomegaly and bone marrow dysfunction.[14] This patient had no recent exposures (she stayed in a city) or pulmonary symptoms that would suggest primary infection, and she was immunocompetent, which makes reactivation disease or disseminated disease unlikely.

CANCER, AUTOIMMUNE DISEASES, AND OTHER CONDITIONS

Given the presence of fever, lymphadenopathy, night sweats, and weight loss, lymphoma must be considered in this case. Common tumors that metastasize to the inguinal lymph nodes include squamous-cell carcinoma of the vulva, penis, and anus[15]; however, the patient's clinical examination and symptoms were not consistent with these diagnoses. A history of rash, weakness, or arthralgias is absent, making autoimmune disorders an unlikely cause of her presentation. Sarcoidosis characteristically manifests as pulmonary hilar lymphadenopathy (in 90% of cases) but can also be associated with peripheral lymphadenopathy.[16]

In summary, this patient's presentation with regional tender lymphadenopathy and history of exposure to cats makes cat scratch disease the most likely diagnosis in this case. An atypical presentation of a systemic infection, a systemic granulomatous process such as sarcoidosis, and lymphoproliferative disorder cannot be ruled out. To establish the diagnosis, I would recommend a right inguinal lymph-node biopsy and serologic testing for *B. henselae* and other infectious causes.

DR. RICHELLE C. CHARLES'S DIAGNOSIS

Bartonella henselae infection.

INITIAL DIAGNOSTIC TESTING

Dr. Aliyah R. Sohani: A right inguinal lymph-node biopsy was performed on an outpatient basis after the patient was discharged from the hospital. Microscopic examination of the biopsy specimen (Fig. 2) revealed architectural distortion characterized by marked capsular fibrosis, hypervascularity with increased plasma cells, florid reactive follicular hyperplasia, focal clusters of monocytoid B cells, and prominent granulomatous inflammation without neutrophils or necrosis. Immunohistochemical staining showed polytypic plasma cells and an appropriate number and distribution of B cells and T cells, and concurrent flow cytometry showed no evidence of a monoclonal B-cell population or an unusual T-cell population. Although an infectious cause appeared most likely, the combination of morphologic features did not point to a specific cause.

In addition, numerous stains for microorganisms were negative; these included acid-fast staining for mycobacteria, Grocott methenamine–silver staining for fungi, and Steiner staining and Warthin–Starry silver staining for spirochetes and other bacteria, as well as immunohistochemical staining for antibodies against *Treponema pallidum*, *Toxoplasma gondii*, and cytomegalovirus. It was difficult to reconcile these lymph node findings with serologic testing for bartonella that had been performed at the time of the patient's admission; the testing had shown a markedly elevated *B. henselae* IgG titer of 1:1024 or higher (reference titer, <1:128) and a moderately elevated *B. quintana* IgG titer of 1:512 (reference titer, <1:128), with undetectable *B. henselae* and *B. quintana* IgM titers of less than 1:20 (reference range for both, <1:20). In patients with lymphadenitis related to *B. henselae* infection (i.e., cat scratch disease), suppurative granulomas associated with neutrophilic clusters and central necrosis would be expected to be present, as would organisms on Warthin–Starry silver stain[17]; however, these features were not seen on histologic evaluation.

DISCUSSION OF MANAGEMENT

Dr. Anne M. Neilan: Because of a high suspicion of early disseminated bartonella infection, we recommended treatment with doxycycline. Before the patient had started doxycycline therapy, her condition had begun to improve; the palpable lymphadenopathy, fever, and anemia had all resolved, and the levels of inflammatory markers were declining. I met her in follow-up, 2 months after the onset of symptoms and 2 weeks after she had begun treatment with doxycycline. She had persistent epigastric pain and a metallic taste, which she attributed to doxycycline. Her examination was notable for mild flank pain on the right side with palpation, which she described as similar to her initial presentation, and which was distinct from her epigastric pain. She had no palpable lymphadenopathy. The result of a serologic test for paracoccidioides was not yet available, and lymph-node biopsy culture revealed no growth.

The patient was thinking about discontinuing treatment with doxycycline, and we considered

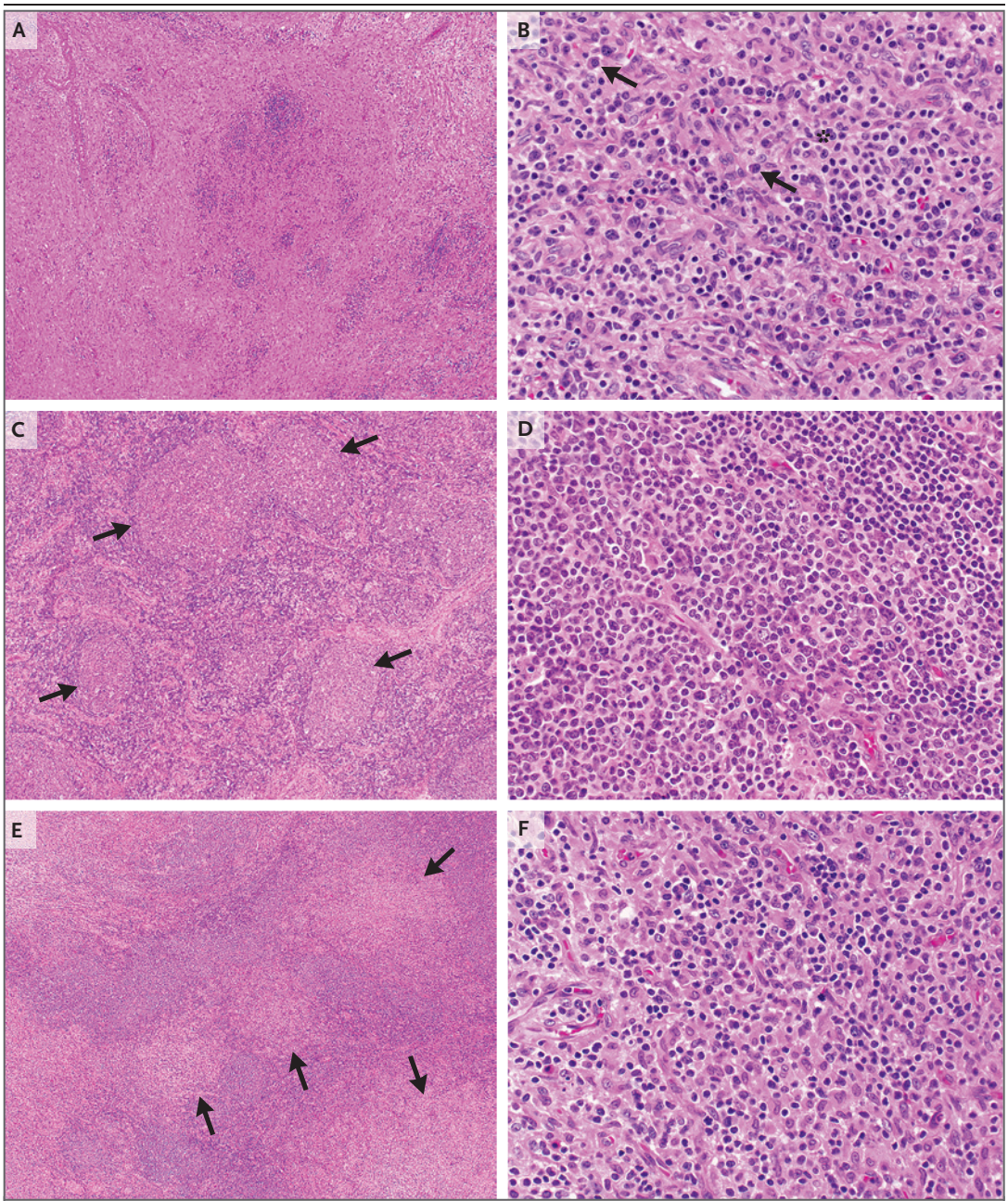


Figure 2. Biopsy Specimen of a Right Inguinal Lymph Node, Initial Diagnostic Evaluation.

Hematoxylin and eosin staining shows marked capsular thickening and fibrosis (Panel A). At the interface of the capsule and lymphoid tissue, there are prominent small vessels associated with increased plasma cells (Panel B, arrows), which were polytypic on RNA in situ hybridization for kappa and lambda immunoglobulin light chains (not shown). Some areas of the underlying lymphoid tissue contain increased numbers of reactive follicles (Panel C, arrows) with interspersed aggregates of reactive monocytoid B cells (Panel D), and other areas contain conspicuous histiocytic infiltrates that are consistent with ill-formed granulomas (Panel E, arrows). At higher magnification, histiocytes that can be seen within the granulomas are associated with small lymphocytes and plasma cells; giant cells and neutrophils are not present, and there is no central necrosis (Panel F).

the following questions: Was an alternative diagnosis, such as infection with paracoccidioides, possible? Was it likely that her serologic test results reflected previous exposure to bartonella rather than a new infection? Were there alternative antibiotic options? First, we discussed that

paracoccidioides was a possible cause, even though her presentation would be atypical. Although she reported that she had remained in a city setting in Brazil, and paracoccidioides is more often associated with rural settings, infection may occur throughout Brazil.[18] Furthermore, the patient's chest CT had shown tiny nondescript lung nodules, which are not typically associated with bartonella infections in immunocompetent hosts.[19] Second, we reasoned that, although the Centers for Disease Control and Prevention estimates that 30 to 40% of adopted shelter cats have bacteremia associated with bartonella, most people who care for cats do not have seroconversion.[20] Third, we discussed potential alternative antibiotics. Azithromycin or clarithromycin may be associated with ototoxicity with long-term use, erythromycin is associated with gastrointestinal side effects, and trimethoprim–sulfamethoxazole (which the patient had received previously) has variable in vitro activity against bartonella and has been associated with an inconsistent clinical response in case reports.[21]

Data that could inform both the choice of antibiotic and the duration of treatment are limited and are mostly based on expert opinion.[22] We ultimately decided on a 2-to-3-month course of antimicrobial therapy. Although the patient had persistent mild abdominal pain of unclear cause, she had had no obvious ophthalmologic, neurologic, hepatosplenic, or cardiac involvement at any point.

Given the patient's lingering symptoms and the fact that a definitive diagnosis might affect her decision to continue antibiotic therapy, a lymph-node specimen was sent for polymerase-chain-reaction (PCR) testing for 16S ribosomal RNA (rRNA) gene sequencing, which uses primers for the conserved 16S rRNA gene to identify the unique signature of most species of bacteria. The patient continued to receive doxycycline while the results were pending. We anticipated that the test would have a low negative predictive value, given that the bartonella bacterial loads had been so low that they were not seen on pathological examination and given that the patient had been treated before biopsy.

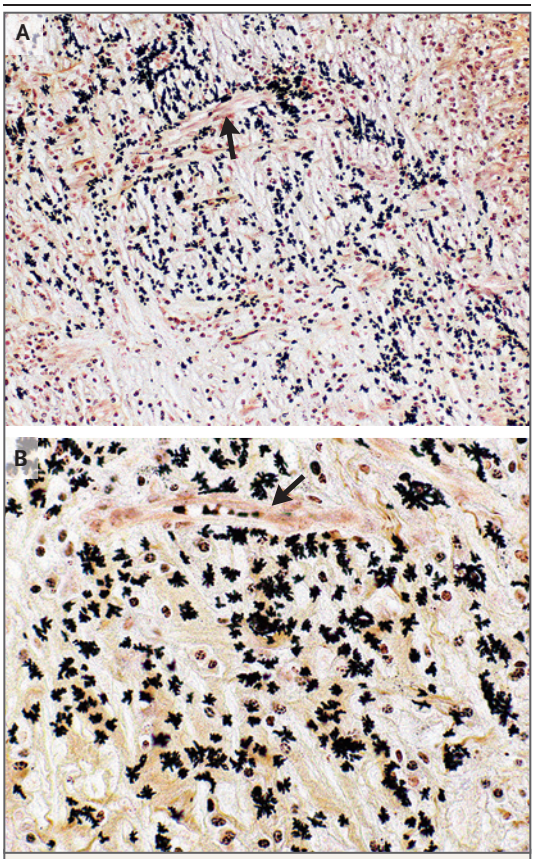


Figure 3. Biopsy Specimen of a Right Inguinal Lymph Node, Final Diagnostic Evaluation.

Warthin–Starry silver staining shows focal areas of the lymph-node specimen that contain clumps of bacilli that appear black against a light brown counterstain and are associated with small vessels (Panels A and B, arrows). At higher magnification, the bacilli are pleomorphic and appear in small clusters (Panel B).

FINAL DIAGNOSTIC TESTING

Dr. Sohani: PCR testing for the 16S rRNA gene showed the presence of *B. henselae* DNA. Reevaluation of the original Warthin–Starry silver stain from the patient's right inguinal lymph-node biopsy confirmed the absence of detectable organisms. Warthin–Starry silver staining was repeated on the paraffin-embedded block of lymph-node tissue that had been evaluated initially and was performed on all three remaining paraffin-embedded blocks of lymph-node tissue that had been submitted for microscopic evaluation. In a single block — one that was different from the one originally stained — Warthin–Starry silver staining highlighted scattered small foci of pleomorphic bacilli in clumps and clusters associated with small vessels, findings that were consistent with *B. henselae* infection (Fig. 3A and 3B). We hypothesized that the initiation of therapy before the lymph-node biopsy had altered the

typical histologic appearance of cat scratch disease and resulted in a paucity of detectable organisms.[23]

FOLLOW-UP

Dr. Neilan: After I shared the results of the PCR testing with the patient, she provided additional history. She had rescued her kitten 1 month before her hospitalization. Subsequently, her adult cat began to have lethargy, anorexia, and abdominal pain, and she noted a large lymph node for which she sought veterinary care. A lymph-node aspiration was unrevealing, and the older cat's condition improved after 5 to 7 days. In retrospect, we suspect that the kitten was the source of the patient's illness. Symptomatic kittens are more likely to have high-grade bacteremia than young adult cats and are more likely to scratch. The putative mechanism of human infection is that the feces of infected fleas contaminates cat claws and inoculates human wounds with viable bacteria. The elimination of fleas rather than treatment of the cat is the key to prevention.[24,25]

The patient continued treatment with doxycycline. After 3 months, *B. henselae* titers had declined from greater than the upper limit of quantitation to 1:128, *B. quintana* titers had normalized, and her abdominal pain had resolved. In addition, the paracoccidioides test had returned a negative result.

FINAL DIAGNOSIS

Bartonella henselae infection (cat scratch disease).

This case was presented at the Medicine Case Conference Interview Series.

No potential conflict of interest relevant to this article was reported.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

REFERENCES

1. Fontanilla J-M, Barnes A, von Reyn CF. Current diagnosis and management of peripheral tuberculous lymphadenitis. Clin Infect Dis 2011;53:555-62.
2. Medical Section of the American Lung Association. Diagnosis and treatment of disease caused by nontuberculous mycobacteria: this official statement of the American Thoracic Society was approved by the board of directors, March 1997. Am J Respir Crit Care Med 1997;156:S1-S25.
3. English CK, Wear DJ, Margileth AM, Lissner CR, Walsh GP. Cat-scratch disease. Isolation and culture of the bacterial agent. JAMA 1988;259:1347-52.
4. Zangwill KM, Hamilton DH, Perkins BA, et al. Cat scratch disease in Connecticut: epidemiology, risk factors, and evaluation of a new diagnostic test. N Engl J Med 1993;329:8-13.
5. Jameson P, Greene C, Regnery R, et al. Prevalence of Bartonella henselae antibodies in pet cats throughout regions of North America. J Infect Dis 1995;172:1145-9.
6. Carithers HA. Cat-scratch disease: an overview based on a study of 1,200 patients. Am J Dis Child 1985;139:1124-33.
7. Margileth AM. Antibiotic therapy for cat-scratch disease: clinical study of therapeutic outcome in 268 patients and a review of the literature. Pediatr Infect Dis J 1992;11:474-8.
8. Bass JW, Freitas BC, Freitas AD, et al. Prospective randomized double blind placebo-controlled evaluation of azithromycin for treatment of cat-scratch disease. Pediatr Infect Dis J 1998;17:447-52.
9. McCabe RE, Brooks RG, Dorfman RF, Remington JS. Clinical spectrum in 107 cases of toxoplasmic lymphadenopathy. Rev Infect Dis 1987;9:754-74.
10. Tularemia — Missouri, 2000–2007. MMWR Morb Mortal Wkly Rep 2009;58:744-8.
11. Stoner BP, Cohen SE. Lymphogranuloma venereum 2015: clinical presentation, diagnosis, and treatment. Clin Infect Dis 2015;61:Suppl 8:S865-S873.
12. Brown TJ, Yen-Moore A, Tyring SK. An overview of sexually transmitted diseases. J Am Acad Dermatol 1999;41:511-32.
13. Fairley JK. General approach to the returned traveler. In: Brunette GW, Nemhauser JB, eds. CDC yellow book 2020: health information for international travel. New York: Oxford University Press, 2019:603-10 (https://wwwnc.cdc.gov/travel/page/yellowbook-home).
14. Salzer HJF, Burchard G, Cornely OA, et al. Diagnosis and management of systemic endemic mycoses causing pulmonary disease. Respiration 2018;96:283-301.
15. Swan MC, Furniss D, Cassell OCS. Surgical management of metastatic inguinal lymphadenopathy. BMJ 2004;329:1272-6.
16. Iannuzzi MC, Rybicki BA, Teirstein AS. Sarcoidosis. N Engl J Med 2007;357:2153-65.
17. Hsi E, Schnitzer B. Reactive Lymphadenopathies. In: Jaffe ES, Arber DA, Campo E, et al., eds. Hematopathology. 2nd ed. Philadelphia: Elsevier, 2017:153-78 (https://www.us.elsevierhealth.com/hematopathology-9780323296137.html).
18. Leading International Fungal Education. Paracoccidioidomycosis (http://www.life-worldwide.org/fungal-diseases/paracoccidioidomycosis).
19. Bandyopadhyay A, Burrage LC, Gonzalez BE. Pulmonary nodules in an immunocompetent child with cat scratch disease. Pediatr Infect Dis J 2013;32:1390-2.
20. Centers for Disease Control and Prevention. Bartonella infection (cat scratch disease, trench fever, and Carrion's disease): for veterinarians. 2020 (https://www.cdc.gov/bartonella/veterinarians/index.html).
21. Koehler JE, Sanchez MA, Garrido CS, et al. Molecular epidemiology of bartonella infections in patients with bacillary angiomatosis-peliosis. N Engl J Med 1997;337:1876-83.
22. Case Records of the Massachusetts General Hospital (Case 30-2005). N Engl J Med 2005;353:1387-94.
23. Scott MA, McCurley TL, Vnencak-Jones CL, et al. Cat scratch disease: detection of Bartonella henselae DNA in archival biopsies from patients with clinically, serologically, and histologically defined disease. Am J Pathol 1996;149:2161-7.
24. Koehler JE, Glaser CA, Tappero JW. Rochalimaea henselae infection: a new zoonosis with the domestic cat as reservoir. JAMA 1994;271:531-5.
25. Chomel BB, Kasten RW, Floyd-Hawkins K, et al. Experimental transmission of Bartonella henselae by the cat flea. J Clin Microbiol 1996;34:1952-6.

EDITORIALS

Warming Up to Cold Perfusion

Winfred W. Williams, M.D., and James F. Markmann, M.D., Ph.D.

Each year, thousands of people in the United States die while awaiting lifesaving liver transplantation because there is an insufficient supply of donor organs. This crisis has prompted increased interest in expanding the organ supply through the promotion of both deceased and live liver donation and has led to the encouragement of use of organs from so-called expanded-criteria donation — grafts from older donors, grafts with suboptimal organ quality (e.g., with varying degrees of fatty liver deposits or steatosis), or grafts that had sustained prolonged ischemia times after they were obtained. Livers from deceased donors that are obtained after circulatory death (i.e., DCD [donation after circulatory death] livers) are perhaps the riskiest organs in this category. The challenge in using DCD livers lies in their obligatory exposure to warm ischemia while circulatory death is awaited. The current standard preservation method of ischemic cold storage may lead to profound ischemia–reperfusion injury at the time of transplantation, causing poor outcomes. These may include inferior graft survival and serious short-term complications, the most troublesome of which is ischemic cholangiopathy, which involves scarring of the biliary tree with resultant biliary obstruction and recurrent episodes of cholangitis. Because of these complications, the use of DCD livers is discouraged. Thus, many potentially transplantable organs may be declined; the utilization rate of livers obtained from a donor after circulatory death is only approximately 25%.[1]

Over the past two decades, numerous investigators have sought to develop new methods of preservation that might ameliorate ischemic organ damage. Machine perfusion of the liver before transplantation has emerged as a promising technology that allows for the continuous infusion of oxygen- and nutrient-rich perfusate in retrieved organs. There are two competing conditions under which machine perfusion is most often used — normothermic machine perfusion (at 35.5 to 37.5°C) and hypothermic machine perfusion (at 1 to 8°C). Each method has its strong proponents.

Normothermic machine perfusion is initiated either immediately after organ recovery at the donor hospital or after a period of ischemic cold storage during transit of the organ to the recipient center. Evidence suggests that so-called end-ischemic normothermic machine perfusion triggers ischemia–reperfusion injury; such injury is caused by reactive oxygen species and the putative activation of an inflammatory cascade, which compromises liver-allograft performance and, most critically, results in a lack of protection from biliary injury.[2] A recent study attempting to recondition marginal organs that had initially been discarded for transplantation used this approach.[3] Unfortunately, in that study, ischemic biliary strictures in DCD grafts were not prevented.[3] A large, prospective, randomized trial initiating normothermic machine perfusion soon after organ recovery showed that the incidence of ischemia–reperfusion injury was mitigated but that the incidence of ischemic biliary complications was not.[4] In contrast, hypothermic machine perfusion that is performed after ischemic

cold storage and just before implantation has been shown in nonrandomized studies to potentiate protective changes in mitochondrial metabolism and electron transfer, which reportedly leads to enhanced availability of critical energy substrates, such as ATP, thus providing an organ-protective metabolic milieu.[5]

Van Rijn et al. now report in the *Journal* the results of the DHOPE-DCD (Dual Hypothermic Oxygenated Perfusion of DCD Liver Grafts in Preventing Nonanastomotic Biliary Strictures after Transplantation) trial,[6] which was a prospective, randomized, controlled trial of hypothermic, oxygenated machine perfusion by means of simultaneous portal-vein and hepatic-artery cannulation in patients with end-stage liver disease; the trial was conducted in six European centers. The investigators found that the incidence of symptomatic nonanastomotic biliary stricture within 6 months after transplantation was approximately two thirds lower among patients in the machine-perfusion group than among patients in the control group, who received grafts that had undergone conventional ischemic cold storage. Furthermore, the investigators found that the incidence of early allograft dysfunction in the machine-perfusion group was almost half that in the control group.

This trial is noteworthy for several reasons. In this large, randomized trial of hypothermic machine perfusion of livers that was focused on livers obtained from donors after circulatory death, the reduced incidence of ischemic biliary damage bodes well for improving patient outcomes and avoiding costly retransplantations. Most importantly, this approach could make more organs available for transplantation if some of the many DCD organs currently discarded could now be transplanted safely. What remains to be determined, however, is whether dual hypothermic, oxygenated low-temperature perfusion can discriminate organ quality as well as machine perfusion at physiologic temperatures. Such discrimination could be critical in determining whether warm or cold perfusion will gain broad acceptance, since the greatest benefit of machine perfusion would be to decrease the current discard rate of 3000 livers annually. The pending results of two large, randomized trials of normothermic machine perfusion that were recently completed in the United States (ClinicalTrials.gov numbers, NCT02522871 and NCT02775162) may help to answer this question.

Machine perfusion has the potential to improve the function of organs obtained from deceased donors and to allow the assessment of organs of uncertain quality, which might dramatically increase the pool of livers available for transplantation. Moreover, this technology will also almost certainly yield exciting new applications, such as ex vivo defatting of steatotic livers, the induction of liver regeneration ex vivo, and the modification of organs by means of gene editing to improve post-transplantation outcomes.

Disclosure forms provided by the authors are available with the full text of this editorial at NEJM.org.

From the Division of Transplant Surgery and Center for Transplantation Sciences, Massachusetts General Hospital and Harvard Medical School, Boston (J.F.M.).

This editorial was published on February 24, 2021, at NEJM.org.

1. U.S. Department of Health and Human Services. Organ procurement and transplantation network: national data (https://optn.transplant.hrsa.gov/data/view-data-reports/national-data/).
2. Watson CJE, Kosmoliaptsis V, Pley C, et al. Observations on the ex situ perfusion of livers for transplantation. Am J Transplant 2018;18:2005-20.
3. Mergental H, Laing RW, Kirkham AJ, et al. Transplantation of discarded livers following viability testing with normothermic machine perfusion. Nat Commun 2020;11:2939.
4. Nasralla D, Coussios CC, Mergental H, et al. A randomized trial of normothermic preservation in liver transplantation. Nature 2018;557:50-6.
5. Dutkowski P, Guarrera JV, de Jonge J, Martins PN, Porte RJ, Clavien P-A. Evolving trends in machine perfusion for liver transplantation. Gastroenterology 2019;156:1542-7.
6. van Rijn R, Schurink IJ, de Vries Y, et al. Hypothermic machine perfusion in liver transplantation — a randomized trial. N Engl J Med 2021;384:1391-401.

DOI: 10.1056/NEJMe2102056

Back to the Future — The Therapeutic Potential of Psychedelic Drugs

Jeffrey A. Lieberman, M.D.

In *The Doors of Perception*, Aldous Huxley described his trial of mescaline as "the most extraordinary and significant experience available to human beings this side of the Beatific Vision." His exegesis was preceded by the synthesis of the hallucinogen lysergic acid diethylamide (LSD) by Sandoz chemist Albert Hoffman in 1938 and was followed by Hoffman's extraction of psilocybin from *Psilocybe mexicana* in 1959.[1] The convergence of scientific research and natural substances historically used by Indigenous peoples in healing and religious rituals sparked interest in what the British psychiatrist Sir Humphrey Osmond termed psychedelic (Greek for "mind manifesting") drugs.

Excitement over psychedelic drugs led to extravagant claims about their vast potential to expand human consciousness, elucidate the psychological architecture of the brain, and treat mental disorders. By the mid-1960s, LSD had been prescribed to approximately 40,000 patients in the United States and spawned over 1000 scientific papers, dozens of books, and cover stories in *Time* and *LIFE* magazines. Meanwhile, recreational use of these drugs, encouraged by countercultural icons Timothy Leary and Ken Kesey, spread. Appeals to "tune in, turn on, and drop-out" propelled unsupervised use to leap-frog medical research. People experiencing "bad trips" filled emergency departments, psychedelics were linked to notorious figures like Charles Manson, and nefarious U.S. Central Intelligence Agency–funded MK-Ultra programs were using psychedelics as a tool for interrogations.

By 1973, the perceived dangers and corrosive effects of psychedelics on American society led the government to reclassify them as Schedule I drugs (i.e., those with a "high potential for abuse and no currently accepted medical use") under the U.S. Controlled Substances Act.[2] For the next 40 years, research and clinical activity regarding psychedelics virtually ceased. But recent clinical applications of drugs such as cannabis and ketamine to treat medical conditions revived interest in psychedelics. Small companies and nonprofit organizations eager to invest in the next psychopharmaceutical revolution launched a "gold rush" to develop psychedelic drugs. However, synthesizing new psychedelic drugs and testing existing ones to gain approval from the Food and Drug Administration for clinical indications, much less for nonprescription recreational use, would be a formidable challenge in view of the cultural baggage attached to psychedelics because of their use as mystical sacraments, interrogation tools, agents for social change, and psychic-reboot mechanisms and because of how they are perceived as both a societal threat and a psychopharmacologic breakthrough.

Initially, drugs were classified as psychedelic on the basis of similar pharmacologic properties and clinical effects (e.g., LSD, psilocybin, mescaline, and dimethyltryptamine). However, the classification has since been expanded to include psychoactive drugs that have different pharmacological targets, such as MDMA (3,4-methylenedioxymethamphetamine ["ecstasy"]) and dissociatives (phencyclidine ["angel dust"] and ketamine). Encouraging results with respect to depression, anxiety, substance use disorders, and palliative care have been reported with these drugs over the past decade.[1,3] Of interest, the therapeutic effects were tied to the subjective report of the user's mystical experience.[1] However, these studies had methodologic limitations (the lack of comparator treatments, functional unblinding, expectancy effects, short follow-up periods, imprecise dosing, and variability in treatment settings).[1,3]

A series of studies by Carhart-Harris and colleagues,[3] culminating in the phase 2, randomized trial, published in this issue of the *Journal*,[4] that compared psilocybin (25 mg at baseline and week 3 plus daily placebo) with escitalopram (10 mg daily for the first 3 weeks, then 20 mg daily, plus 1 mg psilocybin at each of the two dosing sessions) over a 6-week period, provide tantalizing evidence for the efficacy of psilocybin in the treatment of major depressive disorder. However, although the psilocybin-treated patients showed a pattern of improvement, the

between-group differences did not reach statistical significance with respect to the primary outcome (change in score on the 16-item Quick Inventory of Depressive Symptomatology–Self-Report), and the analyses of the numerous secondary outcomes were not adjusted for multiple comparisons. Heterogeneity among patients, who volunteered for the trial in response to advertisements, and uncertainty regarding the appropriate therapeutic dose range and frequency of administration of psilocybin may have influenced the results.

Although this trial is an evidentiary milestone in the development of psychedelic drugs, it also reveals major knowledge gaps. It is unknown how these drugs produce their mind-altering effects. Psychedelics are partial agonists at the 5-hydroxytryptamine type 2A (5-HT_{2A}) receptor, but so are drugs such as lisuride that do not produce subjective effects.[5] Possible explanations of the different effects of psychedelics include functional selective biophysical engagement of the 5-HT_{2A} receptor and activation of alternative intracellular signaling pathways such as β-arrestin[6] and the ability to alter neural circuitry by stimulating proliferation of dendritic spines and synapse formation.[7] The latter explanation is challenged by the temporal dissociation of the immediate subjective experience and the subsequent neurobiologic alterations induced by psychedelic drugs. Another hypothesis is that psychedelics inactivate the prediction-error minimization function of the brain, thereby weakening the mental mechanisms that maintain one's sense of self.[8]

A fundamental question is whether the putative therapeutic effects of psychedelics would require a patient to have a mystical experience or would occur in its absence through the pharmacologic effects on the serotonergic system or remodeling of neural circuitry. To answer this question, compounds are being engineered that have the pharmacologic properties of psychedelics but that do not cause mind-altering effects.[9]

We should not ignore the unusual process by which psychedelics are being developed. This process markedly deviates from conventional models of drug development in which candidate compounds are screened against validated biologic targets and the most promising is selected to test in humans. The unconventional nature of psychedelic drug development is highlighted by the outsized investments they have attracted despite the limited patent protection of existing compounds.

The Carhart-Harris study notwithstanding, we are still awaiting definitive proof of the therapeutic efficacy of psychedelics and their capacity to improve the human condition. Should the mind-bending properties of psychedelics prove to be the panacea their proponents professed, informed consent and safety standards must be established. How do we explain mystical, ineffable, and potentially transformative experiences to patients, particularly if they are in a vulnerable state of mind?[10] What is their potential for addiction?

Given the controversial history, unique properties, and ambitious claims surrounding psychedelic drugs, their development must be guided by the most enlightened science and with the utmost methodologic rigor. However, if psychedelic drugs can indeed map the interface between the brain and mind, illuminate the path to personal growth, and offer therapeutic benefits to those with specific medical conditions, all the tumultuous and subversive efforts to exploit their potential will have been warranted.

Disclosure forms provided by the author are available with the full text of this editorial at NEJM.org.

From the Department of Psychiatry, Columbia University Vagelos College of Physicians and Surgeons, and New York State Psychiatric Institute — both in New York.

This editorial was updated on April 15, 2021, at NEJM.org.

1. Chi T, Gold JA. A review of emerging therapeutic potential of psychedelic drugs in the treatment of psychiatric illnesses. J Neurol Sci 2020;411:116715.
2. Nichols DE. Psychedelics. Pharmacol Rev 2016;68:264-355.
3. Nutt D, Carhart-Harris R. The current status of psychedelics in psychiatry. JAMA Psychiatry 2021;78:121-2.
4. Carhart-Harris R, Giribaldi B, Watts R, et al. Trial of psilocybin versus escitalopram for depression. N Engl J Med 2021; 384:1402-11.
5. González-Maeso J, Ang RL, Yuen T, et al. Identification of a serotonin/glutamate receptor complex implicated in psychosis. Nature 2008;452:93-7.
6. Kim K, Che T, Panova O, et al. Structure of a hallucinogen-activated Gq-coupled 5-HT_{2A} serotonin receptor. Cell 2020; 182(6):1574.e19-1588.e19.
7. Ly C, Greb AC, Cameron LP, et al. Psychedelics promote structural and functional neural plasticity. Cell Rep 2018;23: 3170-82.
8. Swanson LR. Unifying theories of psychedelic drug effects. Front Pharmacol 2018;9:172.
9. Cameron LP, Tombari RJ, Lu J, et al. A non-hallucinogenic psychedelic analogue with therapeutic potential. Nature 2021; 589:474-9.
10. Smith WR, Sisti D. Ethics and ego dissolution: the case of psilocybin. J Med Ethics 2020 May 27 (Epub ahead of print).

DOI: 10.1056/NEJMe2102835

MEDICINE AND SOCIETY

Debra Malina, Ph.D., *Editor*

No Cure without Care — Soothing Science Skepticism

Lisa Rosenbaum, M.D.

A few weeks ago, I cared for a patient, Mr. C., who had a history of bleeding from antiplatelet therapy but remained at high risk for thrombosis. At the time we met, he had just been extubated and, still a bit loopy from the sedation, couldn't give many details about the previous bleeds or thromboembolic events. But on one point he was completely lucid: "I don't like medications," he told me. "It's an unnatural process."

Alan Levinovitz, a professor of religious studies at James Madison University and author of the book *Natural*,[1] explained to me that from an evolutionary standpoint, the instinctive nature of Mr. C.'s aversion is exactly the point. Because there is way too much information for our minds to process as we orient ourselves in the world, our brains have evolved shortcuts, or heuristics, to lighten our cognitive load. Although behavioral economics has focused attention on how heuristics can compromise decision making, we also can't function without them. "You can't be constantly figuring out what is dangerous and safe or who is trustworthy or who isn't by crunching tons of data," Levinovitz said.

A preference for the natural is one of countless heuristics people may use to make health-related decisions — about everything from birthing to taking medication to getting vaccinated. And though some heuristics lead to health-compromising decisions, Levinovitz cautions against dismissing all decision makers who use mental shortcuts as irrational. "It makes it seem like the defect when people disagree with you is an intellectual deficiency, or an absence of information," he says. Instead, discordant health care decisions made by different people often reflect reliance on different heuristics to determine whom and what evidence to trust. Levinovitz's mother, for instance, was initially hesitant to get vaccinated because of her distrust of Trump. "My mom was wrong," he said, "but not because she is ignorant, immoral, or irrational. It's because she is operating with a heuristic that in this context leads to a bad conclusion."

Whereas many people's fundamental heuristic for health-related decisions is to trust medical and scientific experts, vaccine hesitancy reminds us of the many competing forces informing people's intuitions about health, be they religious, political, historical, or identity-based. To be clear, some of these forces are identifiable and should be addressed; the contribution of historical abuses and ongoing systemic racism to vaccine hesitancy in minority communities is a notable example. But in understanding people who simply have a feeling that Covid vaccines should be avoided, identifying specific heuristics matters less than simply recognizing the limits of data in shaping perceptions of truth. "We don't make our decisions about what's true based on an analysis of evidence," Levinovitz emphasized. "It's a profound misconception of how people figure out reality."

Though Covid hasn't changed human nature, its devastating consequences have highlighted the gap between what is true and what people believe. One memorable low for me was reading a South Dakota nurse's description of patients who were critically ill with Covid but continued to insist the virus was a hoax until the moment they were intubated.[2] If you can be denying the existence of a disease while you're dying from it, what hope is there for science to persuade people unaffected by that disease to take it seriously enough to get vaccinated?

For some subset of the population, not much. But although people who are aggressively denying science and disregarding others' health loom large in our minds, there are probably many more who are simply bewildered and no longer

know whom or what to trust. Undoubtedly, current vaccine skepticism is partly rooted in factors specific to this moment and these particular vaccines. But to the extent that hesitancy also reflects deeper, longer-standing fractures in our relationship with the public, its exploration provides an opportunity to improve patient care in ways that go far beyond the pandemic.

DOING MY OWN RESEARCH

Ms. A., a woman in her 40s who previously worked as an anesthesia technician, opposes undergoing Covid vaccination. Referring generally to the contents of vaccines — and using a common heuristic about profit-driven actors — she asked, "Are those good for the human body or just good for someone's pocket?" She described the corruption of the pharmaceutical industry and her related belief that physicians are all pawns, beholden to large corporations, unable to speak the truth even if we wanted to. Accordingly, she expressed both pity and disdain for people who blindly accept the recommendations of the scientific community. The essence of her approach to medical decisions echoed a refrain I hear often: "I need to do my own research."

The first patient who said that to me was a relatively young man for whom I had recommended an implantable cardioverter–defibrillator. I thought, "Why do you need to do your own research when there have been well-conducted randomized, controlled trials, incorporated into guidelines, suggesting that this intervention will prolong your life?" But in the spirit of shared decision making and patient empowerment, I respected his decision and his right to make it. I have often wondered since, however, what doing your own research actually means in a world where being informed can so readily degenerate into being misinformed. Ms. A.'s research on vaccines, for instance, confirmed her suspicion that they contain impurities, including "human DNA from aborted babies" and antifreeze. Though these claims are patently false, they crystallize the startling discrepancy between the time, money, and effort behind a scientific recommendation and the ease with which it can be discredited. How has science become so vulnerable to such undoing?

Vaccine-confidence expert Heidi Larson frames the problem historically. Whereas during the Enlightenment, science was perceived as a way to liberate people from religious dogma, she says, "Today, science has become the new dogma."[3] Larson emphasizes that science can't separate itself from culture, values, inequities, and power struggles.[4] "People are craving a bit more emotion, a bit more religion, something they can put passion into." Science has become devoid of feeling.[5]

Though science with feeling may seem entirely unscientific, Larson isn't advocating a diminution in scientific rigor as much as a contextualization of science to make it feel more relevant to people's lives. Rather than simply telling people they should get a Covid vaccine, for instance, Larson recommends beginning conversations by asking people how they've been coping or what they miss most. "We're fraying at the edges," she says, "and you want people to know we are in this together."[4] A narrow focus on getting the shot in their arms may overlook the fact that people's lives have been undone in ways that vaccines can't fix. In that sense, I think Larson is alluding to a more fundamental tension in the relationship between science and society that the pandemic has magnified: science may tell us what's true, but it can't tell us what's meaningful.

We exhort the public to "follow the science" because, for instance, hundreds of thousands of people are dying and science has found effective ways to mitigate viral spread. But if you're a restaurant owner facing bankruptcy because of closures, a mother whose career is on hold because your children's schools have closed, a man who was prevented from holding the hand of your wife of 50 years as she died, what does following the science mean to you? If we want people to follow science, we also need to acknowledge where science ends and values begin. The destructive forces of science denialism — magnified by the pandemic — have made it difficult to maintain this distinction. But many people who hesitate to follow scientific recommendations may not be rejecting science as much as they are responding to different values and priorities. For this group, what might make science feel more compelling?

Maybe it's time to focus as much on the messenger as we do on the message. In an essay published in December 2019, Harvard history-of-science professor Steven Shapin captures the

crucial difference between knowing science and believing people who know science.[6] Focusing on the three most salient examples of scientific contention — climate change, vaccine safety, and evolution — Shapin argues that what we've been calling a "Crisis of Truth" is really a crisis of trust. He admits, for instance, that his own understanding of climate change is less about knowing the details of the science than about knowing where "science lives." So what if he doesn't know the statistical means of determining global temperature and establishing its rate of change, as long as he knows how to vet the institutions and people who purport to have figured it out? Being a "knowledgeable person," Shapin writes, "may mean knowing a lot of stuff, but it certainly means *knowing who knows* and *who does not know*."

Medical science, which once seemed to live only with physicians, now seems to live everywhere and nowhere at once. Conducting robust scientific research is as critical as ever; the rapid development of highly effective Covid vaccines — possible only because of the decades of sound science that preceded it — speaks to the sanctity of the scientific process. Yet the unwillingness of a substantial proportion of the population to undergo vaccination reminds us of modern medicine's paradox: as science's capacity to improve population health has rapidly increased, so has its fragility. With a few clicks online, what we know can be rendered meaningless. Once the purveyors of knowledge, we now must learn to be its curators as well.

Because so many social, political, and historical forces feed this fragility, when I fail to make science compelling to patients, I often find myself blaming factors beyond my control. But reading Shapin made me wonder whether this sense of futility is an excuse for avoiding a deeper responsibility. When my patient said he wanted to do his own research on defibrillators, for instance, I assumed he meant he'd do a Google search. Because I'd already described the relevant trials, I felt my job was done. In retrospect, I suspect his skepticism was less about the evidence than about whether I, and the institutions I represent, could be trusted to look out for his best interests. Science alone can't overcome a lack of trust. And in that sense, my job was only beginning.

MISSING THE POINT

Just before the pandemic, Heidi Larson was invited to attend a lunch, where she was seated next to a woman who, Larson was warned, was "not into vaccines." Accustomed to contentious interactions about vaccines, Larson braced herself for a potentially difficult discussion. But the woman simply asked her several perfectly reasonable questions, mostly about the influenza vaccine. In listening to her, Larson seemed to make the generic guidance feel more personal. Indeed, as they left the lunch together, the woman said she was going to get a flu shot — "because you made it feel less anonymous." Larson was moved, struck by how easily we often give up on people. "It hit me in a way that she didn't realize," Larson told me. "We undervalue the power of talk."

Of course, very few physicians would have time for such a conversation during an office visit. And if documentation demands shape physician behavior, the requirements related to influenza vaccination (assess eligibility and document whether given or refused[7]) typify an incentive structure that values box checking over actual doctoring. But medicine's undervaluing of the power of talk has implications extending far beyond vaccines. For many physicians, listening has become a luxury, squeezed out by time constraints, the demands of the electronic health record, and the countless metrics demanding our attention. If you see a patient in the clinic with a newly diagnosed cardiomyopathy and don't prescribe a beta-blocker, the institute's "quality team" may alert you until you either prescribe the drug or justify its omission. If you spend 30 minutes listening to the same patient explain why he doesn't see the need to take a beta-blocker, no one cares — except, of course, the patient.

In some ways, we are victims of our own success. A physician friend recently told me a story about one of his mentors, a cardiologist who graduated from medical school in the 1930s. During the cardiologist's training — before antibiotics were in widespread use — he cared for a patient who was dying of bacterial endocarditis. The cardiologist, who would go on to earn renown for both clinical and scientific contributions, knew he had no medical treatments to offer the patient. So he offered himself instead: every night, he would sleep next to the patient in

the adjacent bed, accompanying him until he died. Now, of course, science has given us countless tools for preventing and treating disease. But somehow, in our efforts to systematize all we know (and make it profitable), the centrality of the doctor–patient relationship got lost. Is there some inevitable trade-off between the capacity to care and the capacity to cure?

Surely not for everyone. About a year ago, I was talking to one of our cardiology fellows about our admiration for many people in our division. He mentioned a revered senior clinician, and as we tried to analyze the essence of his gift — which seemed to extend beyond brilliance, or judgment, or even devotion — the fellow said, "You know, what he's really doing is therapy." He didn't mean the kind of therapy that tries to probe the depths of your unconscious to reveal your most primitive thoughts. He meant the seemingly simpler act of giving someone the space to be known. Having cared for this clinician's patients, many of whom will not proceed with any recommended intervention until he has offered his blessing, I can see that it is only because he knows them that they trust that he knows what's best for them.

But in this capacity, he is hardly alone. Though survey data suggest that less than a quarter of the U.S. public trusts the health care system at large, about 60% think that doctors can be trusted.[8] I suspect, then, that the many physicians who continue to earn their patients' trust do so despite the system, not because of it.

Before she named it *Stuck*, Larson was tentatively calling her book on vaccine hesitancy *Missing the Point*. We are so focused on changing people's minds to get vaccines into their bodies, she explained, that we've been ignoring the factors contributing to vaccine hesitancy in the first place. To be clear, some subset of the hesitant have been misinformed and simply need to hear accurate information; believing an antibiotic allergy is a contraindication to vaccination, for instance, is an easily corrected misunderstanding. Moreover, the vaccines absolutely represent a miracle of science; their many potential benefits — from preventing Covid deaths to restoring normalcy to society — should be shouted from the rooftops. But to the extent that vaccine hesitancy reflects a loss of faith in our health care system, this moment should force us to examine the ways in which our system is no longer deserving of that faith. The path forward, then, isn't to compromise science by turning it into an art; rather, it's to stop trying to turn the art of medicine into a science.

After Mr. C. — my patient who prefers natural remedies to medications — left the hospital, he agreed to be interviewed for this article, to elaborate on the origins of his preference. He told me lots of stories — about a log cabin in Oregon, a pond he'd created in his backyard, a moment at a bus stop 60 years ago when a severely disabled man, hunched under the weight of a huge backpack, told him that "Life was but a dream between two sleeps." When I finally asked him how he felt about getting a Covid vaccine, he said that because his body had already avoided the infection, he assumed he had a natural ability to fend it off. But he added, "You're going to tell me I have to, right?" No, I said, "my job right now isn't to be your doctor. I'm just here to listen." Only later did I realize that in making this distinction, I was missing the point.

This article is Part 2 in a two-part series.

Identifying details have been changed to protect people's privacy.

Disclosure forms provided by the author are available at NEJM.org.

Dr. Rosenbaum is a national correspondent for the *Journal*.

This article was published on February 19, 2021, at NEJM.org.

1. Levinovitz A. Natural: how faith in nature's goodness leads to harmful fads, unjust laws, and flawed science. Boston: Beacon Press, 2020.
2. Villegas P. South Dakota nurse says many patients deny the coronavirus exists — right up until death. The Washington Post. November 16, 2020.
3. Larson HJ. Stuck: how vaccine rumors start — and why they don't go away. New York: Oxford University Press, 2020.
4. Brown School of Public Health. A conversation with Heidi J. Larson, Ph.D. October 1, 2020 (https://www.brown.edu/academics/public-health/events/larson).
5. Center for Strategic and International Studies. CSIS book launch with Heidi J. Larson — Stuck. September 16, 2020 (https://www.csis.org/events/online-event-csis-book-launch-heidi-j-larson-%E2%80%93-stuck).
6. Shapin S. Is there a crisis of truth? Los Angeles Review of Books. December 2, 2019 (https://lareviewofbooks.org/article/is-there-a-crisis-of-truth/).
7. Centers for Medicare and Medicaid Services. Preventive care and screening: influenza immunization. Version 4.0. Quality ID no. 110 (NQF 0041). Chicago: American Medical Association, November 2019 (https://qpp.cms.gov/docs/QPP_quality_measure_specifications/CQM-Measures/2020_Measure_110_MIPSCQM.pdf).
8. Blendon RJ, Benson JM, Hero JO. Public trust in physicians — U.S. medicine in international perspective. N Engl J Med 2014; 371:1570-2.

DOI: 10.1056/NEJMms2101989

CORRESPONDENCE

Neutralizing Activity of BNT162b2-Elicited Serum

TO THE EDITOR: BNT162b2 is a nucleoside-modified RNA vaccine expressing the full-length prefusion spike glycoprotein (S) of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). In a randomized, placebo-controlled clinical trial involving approximately 44,000 participants, immunization conferred 95% efficacy against coronavirus disease 2019 (Covid-19).[1]

New, highly transmissible SARS-CoV-2 variants that were first detected in the United Kingdom (B.1.1.7 lineage), South Africa (B.1.351 lineage), and Brazil (P.1 lineage) with mutations in the *S* gene are spreading globally. To analyze effects on neutralization elicited by BNT162b2, we engineered *S* mutations from each of the three new lineages into USA-WA1/2020, a relatively early isolate of the virus from January 2020 (Fig. S1 in the Supplementary Appendix, available with the full text of this letter at NEJM.org). We thereby produced three recombinant viruses representing each of these lineages and two additional ones in which we engineered subsets of mutations of the B.1.351 lineage. Thus, the first recombinant virus had all the mutations found in the *S* gene in the B.1.1.7 lineage (B.1.1.7-spike), the second had all the mutations found in the *S* gene in the P.1 lineage (P.1-spike), the third had all the mutations found in the *S* gene in the B.1.351 lineage (B.1.351-spike), the fourth had an N-terminal domain deletion found in the B.1.351 lineage and the globally dominant D614G substitution (B.1.351-Δ242-244+D614G), and the fifth had the three mutations from the B.1.351 lineage affecting amino acids in the receptor-binding site (K417N, E484K, and N501Y) and a D614G substitution (B.1.351-RBD+D614G). The mutant amino acid residues in the B.1.351-RBD+D614G recombinant virus are also among those in the P.1 lineage virus, although in the P.1 lineage virus, K417 is mutated to threonine rather than asparagine. All the mutant viruses yielded infectious viral titers exceeding 10^7 plaque-forming units per milliliter. The B.1.1.7-spike and B.1.351-spike viruses formed plaques that were smaller than those formed by the other viruses (Fig. S2).

We performed 50% plaque reduction neutralization testing ($PRNT_{50}$) using 20 serum samples that had been obtained from 15 participants in the pivotal trial[1,2] 2 or 4 weeks after the administration of the second dose of 30 μg of BNT162b2 (which occurred 3 weeks after the first immunization) (Fig. S3). All the serum samples efficiently neutralized USA-WA1/2020 and all the viruses with variant spikes. Almost all of them did so at titers higher than 1:40. Geometric mean neutralizing titers against USA-WA1/2020, B.1.1.7-spike, P.1-spike, B.1.351-spike, B.1.351-Δ242-244+D614G, and B.1.351-RBD+D614G viruses were 532, 663, 437, 194, 485, and 331, respectively (Fig. 1 and Table S1). Thus, as compared with neutralization of USA-WA1/2020, neutralization of B.1.1.7-spike and P.1-spike viruses was roughly equivalent, and neutralization of B.1.351-spike virus was robust but lower. Our data are also consistent

THIS WEEK'S LETTERS

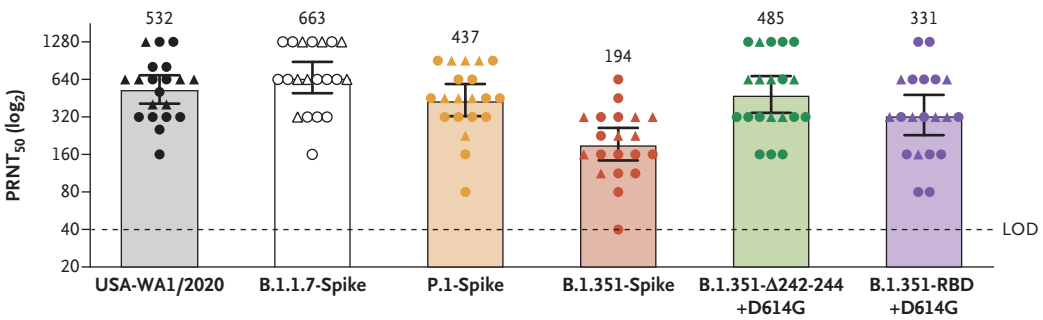


Figure 1. Serum Neutralization of Variant Strains of SARS-CoV-2 after the Second Dose of BNT162b2 Vaccine.

Shown are the results of 50% plaque reduction neutralization testing ($PRNT_{50}$) with the use of 20 samples obtained from 15 trial participants 2 weeks (circles) or 4 weeks (triangles) after the administration of the second dose of the BNT162b2 vaccine. The mutant viruses were obtained by engineering the full set of mutations in the B.1.1.7, P.1., or B.1.351 lineage or subsets of the *S* gene mutations in the B.1.351 lineage (B.1.351-Δ242-244+D614G and B.1.351-RBD+D614G) into USA-WA1/2020. Each data point represents the geometric mean $PRNT_{50}$ obtained with a serum sample against the indicated virus, including data from repeat experiments, as detailed in Table S1 in the Supplementary Appendix. The data for USA-WA1/2020 are from three experiments; for B.1.1.7-spike, B.1.351-Δ242-244+D614G, and B.1.351-RBD-D614G viruses from one experiment each; and for P.1-spike and B.1.351-spike viruses from two experiments each. In each experiment, the neutralization titer was determined in duplicate assays, and the geometric mean was taken. The heights of bars and the numbers over the bars indicate geometric mean titers. The I bars indicate 95% confidence intervals. Statistical analysis was performed with the use of the Wilcoxon signed-rank test. The statistical significance of the difference between geometric mean titers in the USA-WA1/2020 neutralization assay and in each variant virus neutralization assay with the same serum samples are as follows: P=0.02 for B.1.1.7-spike; P=0.06 for P.1-spike; P<0.001 for B.1.351-spike; P=0.99 for B.1.351-Δ242-244+D614G; and P=0.005 for B.1.351-RBD+D614G. LOD denotes limit of detection.

with lower neutralization titers against the virus with the full set of B.1.351-spike mutations than against virus with either subset of mutations. Our findings also suggest that mutations that result in amino acid substitutions K417N, E484K, and N501Y in the receptor-binding site have a greater effect on neutralization than the 242–244 deletion affecting the N-terminal domain of the spike protein.

Limitations of the study include the potential for mutations to alter neutralization by affecting spike function rather than antigenicity. Therefore, each neutralization assay with a different target virus is unique, and comparisons between neutralization titers from different assays should be interpreted with caution. Neutralizing activity against the B.1.351 lineage virus was robust at a geometric mean titer that was much higher than that obtained after one dose of BNT162b2, when strong efficacy was already observed in the C4591001 efficacy trial.[1-3] T-cell immunity may also be involved in protection,[4] and BNT162b2 immunization elicits CD8+ T-cell responses that recognize multiple variants.[5] Ultimately, conclusions about vaccine-mediated protection that are extrapolated from neutralization or T-cell data must be validated by real-world evidence collected in regions where the SARS-CoV-2 variants are circulating.

Yang Liu, Ph.D.
Jianying Liu, Ph.D.
Hongjie Xia, Ph.D.
Xianwen Zhang, B.S.
Camila R. Fontes-Garfias, Ph.D.
University of Texas Medical Branch
Galveston, TX

Kena A. Swanson, Ph.D.
Hui Cai, Ph.D.
Ritu Sarkar, M.A.
Wei Chen, M.S.
Mark Cutler, Ph.D.
David Cooper, Ph.D.
Pfizer Vaccine Research and Development
Pearl River, NY

Scott C. Weaver, Ph.D.
University of Texas Medical Branch
Galveston, TX

Alexander Muik, Ph.D.
Ugur Sahin, M.D.
BioNTech
Mainz, Germany

Kathrin U. Jansen, Ph.D.
Pfizer Vaccine Research and Development
Pearl River, NY

Xuping Xie, Ph.D.
University of Texas Medical Branch
Galveston, TX
xuxie@utmb.edu

Philip R. Dormitzer, M.D., Ph.D.
Pfizer Vaccine Research and Development
Pearl River, NY
philip.dormitzer@pfizer.com

Pei-Yong Shi, Ph.D.
University of Texas Medical Branch
Galveston, TX
peshi@utmb.edu

Drs. Y. Liu and J. Liu contributed equally to this letter.
Supported by Pfizer and BioNTech.
Disclosure forms provided by the authors are available with the full text of this letter at NEJM.org.

A preliminary version of this letter was published on February 17, 2021, and was updated on March 8, 2021, at NEJM.org.

1. Polack FP, Thomas SJ, Kitchin N, et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. N Engl J Med 2020;383:2603-15.
2. Walsh EE, Frenck RW Jr, Falsey AR, et al. Safety and immunogenicity of two RNA-based Covid-19 vaccine candidates. N Engl J Med 2020;383:2439-50.
3. Sahin U, Muik A, Vogler I, et al. BNT162b2 induces SARS-CoV-2-neutralising antibodies and T cells in humans. December 11, 2020 (https://www.medrxiv.org/content/10.1101/2020.12.09.20245175v1). preprint.
4. Liao M, Liu Y, Yuan J, et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. Nat Med 2020;26:842-4.
5. Skelly DT, Harding AC, Gilbert-Jaramillo J, et al. Vaccine-induced immunity provides more robust heterotypic immunity than natural infection to emerging SARS-CoV-2 variants of concern. February 9, 2021 (https://www.researchsquare.com/article/rs-226857/v1). preprint.

DOI: 10.1056/NEJMc2102017

Serum Neutralizing Activity Elicited by mRNA-1273 Vaccine

TO THE EDITOR: The mRNA-1273 vaccine against the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) elicited high neutralizing-antibody titers in phase 1 trial participants[1,2] and has been shown to be highly efficacious in preventing symptomatic Covid-19 disease and severe disease.[3] The emergence of SARS-CoV-2 variants in the United Kingdom (the B.1.1.7 lineage), South Africa (the B.1.351 lineage), Brazil (the P.1 lineage), and California (the B.1.427/B.1.429 lineage) has led to concerns about increased transmission and the potential of these variants to circumvent immunity elicited by natural infection or vaccination. The recent identification in the United Kingdom of a B.1.1.7 variant that includes the E484K mutation (B.1.1.7+E484K) furthers these concerns.

We assayed the neutralizing activity against recombinant vesicular stomatitis virus (rVSV)–based SARS-CoV-2 (a pseudovirus-based model) in serum samples obtained from eight participants in the phase 1 trial. The samples were

Figure 1 (facing page). Neutralization of SARS-CoV-2 Pseudoviruses in Serum Samples.
Serum samples obtained from participants who received the mRNA-1273 vaccine in a phase 1 trial were collected on day 36 (7 days after the participants received the second dose of the vaccine). Neutralization was measured with the use of a recombinant vesicular stomatitis virus (rVSV)–based pseudovirus neutralization assay that incorporated D614G or the indicated spike mutations present in the B.1.1.7 variant (Panels A and B), the B.1.351 variant (Panels C and D), or the P.1 variant, the B.1.427/B.1.429 (versions 1 and 2) variants, and the B.1.1.7+E484K variant (Panels E through I). The red dots indicate the results from serum samples of the individual participants; the white dots, white diamonds, and white triangles the same samples tested against the variants shown on the x axis; and the horizontal dashed lines the lower limit of quantification. The reciprocal neutralizing titers on the pseudovirus neutralization assay at a 50% inhibitory dilution (ID_{50}) are shown. In Panels A, C, and E, boxes and horizontal bars denote the interquartile range (IQR) and the median neutralizing titer, respectively. Whisker end points are equal to the maximum and minimum values below or above the median at 1.5 times the IQR. In Panels B, D, F, G, H, and I, the lines connect the D614G and variant neutralization titers in matched samples. We detected reductions by a factor of 1.2 in titers of neutralizing antibodies against the B.1.1.7 variant (Panel B), a factor of 6.4 against the B.1.351 variant (Panel D), a factor of 3.5 against the P.1 variant (Panel F), a factor of 2.3 against the B.1.427/B.1.429-v1 variant (Panel G), a factor of 2.8 against the B.1.427/B.1.429-v2 variant (Panel H), and a factor of 3.1 against the B.1.1.7+E484K variant (Panel I). Statistical analysis of matched pairs was performed with the use of the Wilcoxon signed-rank test.

1. Polack FP, Thomas SJ, Kitchin N, et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. N Engl J Med 2020;383:2603-15.
2. Walsh EE, Frenck RW Jr, Falsey AR, et al. Safety and immunogenicity of two RNA-based Covid-19 vaccine candidates. N Engl J Med 2020;383:2439-50.
3. Sahin U, Muik A, Vogler I, et al. BNT162b2 induces SARS-CoV-2-neutralising antibodies and T cells in humans. December 11, 2020 (https://www.medrxiv.org/content/10.1101/2020.12.09.20245175v1). preprint.
4. Liao M, Liu Y, Yuan J, et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. Nat Med 2020;26:842-4.
5. Skelly DT, Harding AC, Gilbert-Jaramillo J, et al. Vaccine-induced immunity provides more robust heterotypic immunity than natural infection to emerging SARS-CoV-2 variants of concern. February 9, 2021 (https://www.researchsquare.com/article/rs-226857/v1). preprint.

DOI: 10.1056/NEJMc2102017

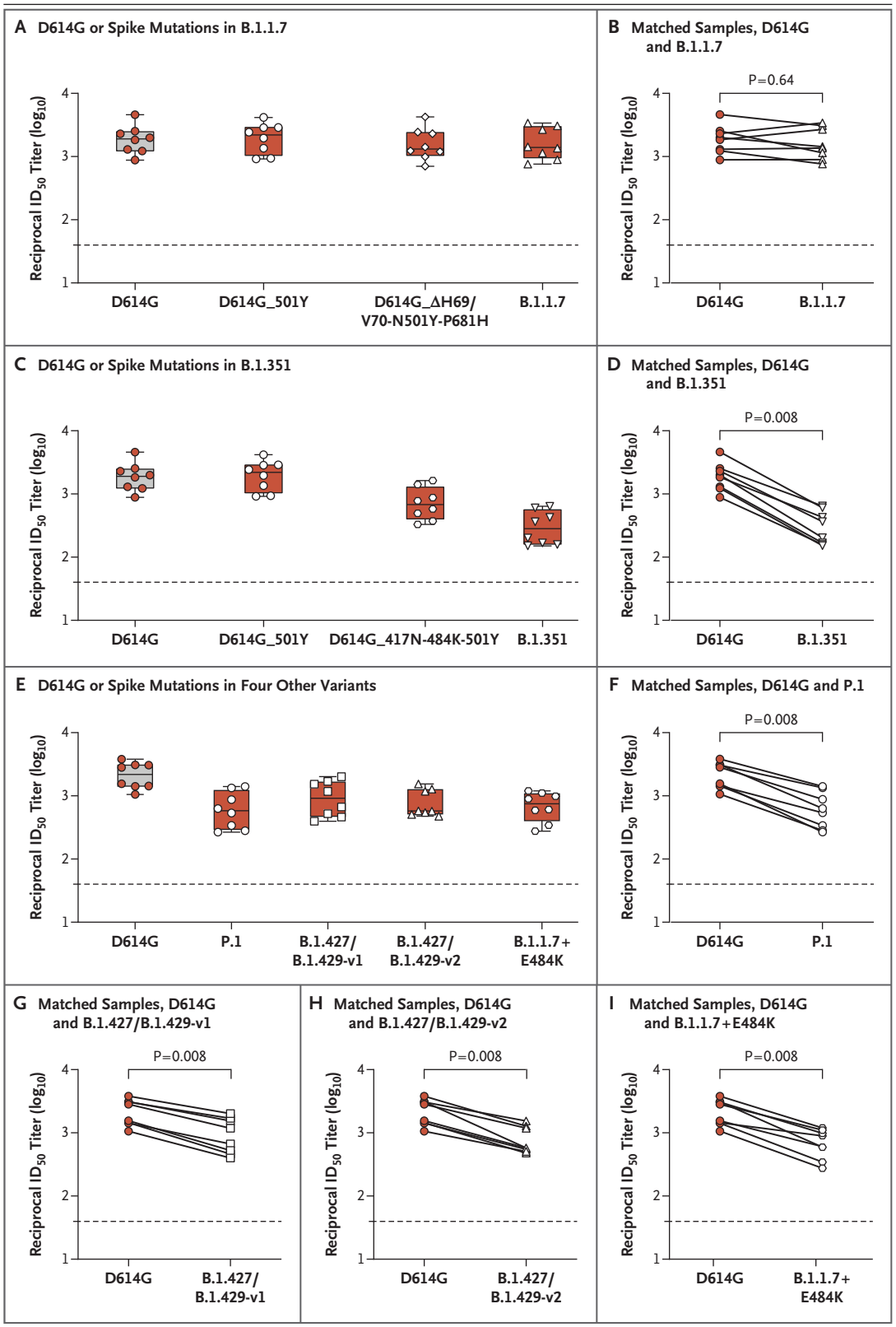
Serum Neutralizing Activity Elicited by mRNA-1273 Vaccine

TO THE EDITOR: The mRNA-1273 vaccine against the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) elicited high neutralizing-antibody titers in phase 1 trial participants[1,2] and has been shown to be highly efficacious in preventing symptomatic Covid-19 disease and severe disease.[3] The emergence of SARS-CoV-2 variants in the United Kingdom (the B.1.1.7 lineage), South Africa (the B.1.351 lineage), Brazil (the P.1 lineage), and California (the B.1.427/B.1.429 lineage) has led to concerns about increased transmission and the potential of these variants to circumvent immunity elicited by natural infection or vaccination. The recent identification in the United Kingdom of a B.1.1.7 variant that includes the E484K mutation (B.1.1.7+E484K) furthers these concerns.

We assayed the neutralizing activity against recombinant vesicular stomatitis virus (rVSV)–based SARS-CoV-2 (a pseudovirus-based model) in serum samples obtained from eight participants in the phase 1 trial. The samples were obtained 1 week after the participants had received the second dose of mRNA-1273 vaccine. We tested pseudoviruses bearing the spike proteins from the original Wuhan-Hu-1 isolate, the D614G variant, and the B.1.1.7, B.1.351, P.1, B.1.427/B.1.429, B.1.1.7+E484K, and other variants (20E [EU1], 20A.EU2, N439K-D614G, and the mink cluster 5 variant that was first identified in Denmark).

Both the full panel of mutations in *S* and a subset of mutations affecting the receptor-binding domain (RBD) region of the B.1.1.7 variant had no significant effect on neutralization by serum obtained from participants who had received the mRNA-1273 vaccine in the phase 1 trial (Fig. 1A and 1B). In contrast, we observed a decrease in titers of neutralizing antibodies against the P.1 variant, the B.1.427/B.1.429 variant (versions 1 and 2), the B.1.1.7+E484K variant, and the B.1.351 variant as well as a subset of its

Figure 1 (facing page). Neutralization of SARS-CoV-2 Pseudoviruses in Serum Samples.
Serum samples obtained from participants who received the mRNA-1273 vaccine in a phase 1 trial were collected on day 36 (7 days after the participants received the second dose of the vaccine). Neutralization was measured with the use of a recombinant vesicular stomatitis virus (rVSV)–based pseudovirus neutralization assay that incorporated D614G or the indicated spike mutations present in the B.1.1.7 variant (Panels A and B), the B.1.351 variant (Panels C and D), or the P.1 variant, the B.1.427/B.1.429 (versions 1 and 2) variants, and the B.1.1.7+E484K variant (Panels E through I). The red dots indicate the results from serum samples of the individual participants; the white dots, white diamonds, and white triangles the same samples tested against the variants shown on the x axis; and the horizontal dashed lines the lower limit of quantification. The reciprocal neutralizing titers on the pseudovirus neutralization assay at a 50% inhibitory dilution (ID_{50}) are shown. In Panels A, C, and E, boxes and horizontal bars denote the interquartile range (IQR) and the median neutralizing titer, respectively. Whisker end points are equal to the maximum and minimum values below or above the median at 1.5 times the IQR. In Panels B, D, F, G, H, and I, the lines connect the D614G and variant neutralization titers in matched samples. We detected reductions by a factor of 1.2 in titers of neutralizing antibodies against the B.1.1.7 variant (Panel B), a factor of 6.4 against the B.1.351 variant (Panel D), a factor of 3.5 against the P.1 variant (Panel F), a factor of 2.3 against the B.1.427/B.1.429-v1 variant (Panel G), a factor of 2.8 against the B.1.427/B.1.429-v2 variant (Panel H), and a factor of 3.1 against the B.1.1.7+E484K variant (Panel I). Statistical analysis of matched pairs was performed with the use of the Wilcoxon signed-rank test.


A D614G or Spike Mutations in B.1.1.7
Reciprocal ID50 Titer (log10)
4
3
2
1
D614G
D614G_501Y
D614G_ΔH69/
V70-N501Y-P681H
B.1.1.7
B Matched Samples, D614G
and B.1.1.7
P=0.64
Reciprocal ID50 Titer (log10)
D614G
B.1.1.7
C D614G or Spike Mutations in B.1.351
Reciprocal ID50 Titer (log10)
D614G
D614G_501Y
D614G_417N-484K-501Y
B.1.351
D Matched Samples, D614G
and B.1.351
P=0.008
Reciprocal ID50 Titer (log10)
D614G
B.1.351
E D614G or Spike Mutations in Four Other Variants
Reciprocal ID50 Titer (log10)
D614G
P.1
B.1.427/
B.1.429-v1
B.1.427/
B.1.429-v2
B.1.1.7+
E484K
F Matched Samples, D614G and P.1
P=0.008
Reciprocal ID50 Titer (log10)
D614G
P.1
G Matched Samples, D614G
and B.1.427/B.1.429-v1
P=0.008
Reciprocal ID50 Titer (log10)
D614G
B.1.427/
B.1.429-v1
H Matched Samples, D614G
and B.1.427/B.1.429-v2
P=0.008
Reciprocal ID50 Titer (log10)
D614G
B.1.427/
B.1.429-v2
I Matched Samples, D614G
and B.1.1.7+E484K
P=0.008
Reciprocal ID50 Titer (log10)
D614G
B.1.1.7+
E484K

mutations in the RBD. We detected reductions by a factor of between 2.3 and 6.4 in titers of neutralizing antibodies against this panel of variants (Fig. 1C through 1I). The largest effect on neutralization, reduction by a factor of 6.4, was measured against the B.1.351 variant (Fig. 1C and 1D). However, the geometric mean neutralizing titer against B.1.351 was 1:290, and all the serum samples fully neutralized the rVSV pseudovirus, albeit at relatively low dilutions (Fig. S1 in the Supplementary Appendix, available with the full text of this letter at NEJM.org). The effect of the E484K mutation was observed by comparing neutralizing activity against the B.1.1.7 variant with neutralizing activity against the B.1.1.7+E484K variant. We found a significant reduction in neutralizing titers when the E484K mutation was present (Fig. 1B and 1I). Using both rVSV and lentiviral neutralization assays, we observed a similar trend in serum samples obtained from macaque monkeys (Figs. S2 and S3).

The rVSV-based pseudovirus neutralization assay was also used to assess the neutralizing activity of serum obtained from participants who had received the mRNA-1273 vaccine in the phase 1 trial against the full-length spike protein of the dominant strain in 2020 (D614G), as well as against 20E (EU1), 20A.EU2, N439K-D614G, and mink cluster 5 variants (Table S1). We observed levels of neutralization against these variants that were similar to those against the Wuhan-Hu-1 (D614) isolate (Fig. S4).

Protection conferred by the mRNA-1273 vaccine against the P.1, B.1.427/B.1.429, and B.1.351 variants remains to be determined. Our findings underscore the importance of continued viral surveillance and evaluation of vaccine efficacy against new variants and may help to facilitate the establishment of correlates of protection in both nonhuman primates and humans.

Kai Wu, Ph.D.
Moderna
Cambridge, MA

Anne P. Werner, B.S.
National Institute of Allergy and Infectious Diseases
Bethesda, MD

Matthew Koch, B.S.
Angela Choi, Ph.D.
Elisabeth Narayanan, Ph.D.
Guillaume B.E. Stewart-Jones, Ph.D.
Tonya Colpitts, Ph.D.
Hamilton Bennett, M.S.
Moderna
Cambridge, MA

Seyhan Boyoglu-Barnum, Ph.D.
Wei Shi, Ph.D.
Juan I. Moliva, Ph.D.
Nancy J. Sullivan, Ph.D.
Barney S. Graham, M.D.
National Institute of Allergy and Infectious Diseases
Bethesda, MD

Andrea Carfi, Ph.D.
Moderna
Cambridge, MA

Kizzmekia S. Corbett, Ph.D.
Robert A. Seder, M.D.
National Institute of Allergy and Infectious Diseases
Bethesda, MD

Darin K. Edwards, Ph.D.
Moderna
Cambridge, MA
darin.edwards@modernatx.com

Dr. Wu and Ms. Werner and Drs. Seder and Edwards contributed equally to this letter.

Supported by the Intramural Research Program of the Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH); the Office of the Assistant Secretary for Preparedness and Response, Biomedical Advanced Research and Development Authority, Department of Health and Human Services (contract 75A50120C00034); and a research fellowship (to Dr. Corbett) that was partially funded by the Undergraduate Scholarship Program, Office of Intramural Training and Education, Office of the Director, NIH.

Disclosure forms provided by the authors are available with the full text of this letter at NEJM.org.

A preliminary version of this letter was published on February 17, 2021, at NEJM.org. This letter was published on March 17, 2021, at NEJM.org.

1. Jackson LA, Anderson EJ, Rouphael NG, et al. An mRNA vaccine against SARS-CoV-2 — preliminary report. N Engl J Med 2020;383:1920-31.
2. Anderson EJ, Rouphael NG, Widge AT, et al. Safety and immunogenicity of SARS-CoV-2 mRNA-1273 vaccine in older adults. N Engl J Med 2020;383:2427-38.
3. Baden LR, El Sahly HM, Essink B, et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. N Engl J Med 2021;384:403-16.

DOI: 10.1056/NEJMc2102179

Cardiovascular Outcomes with Sotagliflozin

TO THE EDITOR: In the SOLOIST-WHF trial (Effect of Sotagliflozin on Cardiovascular Events in Patients with Type 2 Diabetes Post Worsening Heart Failure), Bhatt et al. (Jan. 14 issue)[1] report that death from cardiovascular causes and hospitalizations and urgent visits for heart failure (the

ceived the mRNA-1273 vaccine in the phase 1 trial (Fig. 1A and 1B). In contrast, we observed a decrease in titers of neutralizing antibodies against the P.1 variant, the B.1.427/B.1.429 variant (versions 1 and 2), the B.1.1.7+E484K variant, and the B.1.351 variant as well as a subset of its mutations in the RBD. We detected reductions by a factor of between 2.3 and 6.4 in titers of neutralizing antibodies against this panel of variants (Fig. 1C through 1I). The largest effect on neutralization, reduction by a factor of 6.4, was measured against the B.1.351 variant (Fig. 1C and 1D). However, the geometric mean neutralizing titer against B.1.351 was 1:290, and all the serum samples fully neutralized the rVSV pseudovirus, albeit at relatively low dilutions (Fig. S1 in the Supplementary Appendix, available with the full text of this letter at NEJM.org). The effect of the E484K mutation was observed by comparing neutralizing activity against the B.1.1.7 variant with neutralizing activity against the B.1.1.7+E484K variant. We found a significant reduction in neutralizing titers when the E484K mutation was present (Fig. 1B and 1I). Using both rVSV and lentiviral neutralization assays, we observed a similar trend in serum samples obtained from macaque monkeys (Figs. S2 and S3).

The rVSV-based pseudovirus neutralization assay was also used to assess the neutralizing activity of serum obtained from participants who had received the mRNA-1273 vaccine in the phase 1 trial against the full-length spike protein of the dominant strain in 2020 (D614G), as well as against 20E (EU1), 20A.EU2, N439K-D614G, and mink cluster 5 variants (Table S1). We observed levels of neutralization against these variants that were similar to those against the Wuhan-Hu-1 (D614) isolate (Fig. S4).

Protection conferred by the mRNA-1273 vaccine against the P.1, B.1.427/B.1.429, and B.1.351 variants remains to be determined. Our findings underscore the importance of continued viral surveillance and evaluation of vaccine efficacy against new variants and may help to facilitate the establishment of correlates of protection in both nonhuman primates and humans.

Kai Wu, Ph.D.

Moderna
Cambridge, MA

Anne P. Werner, B.S.

National Institute of Allergy and Infectious Diseases
Bethesda, MD

Darin K. Edwards, Ph.D.

Moderna
Cambridge, MA
darin.edwards@modernatx.com

and Others

Dr. Wu and Ms. Werner and Drs. Seder and Edwards contributed equally to this letter.

A complete list of authors is available with the full text of this letter at NEJM.org.

Supported by the Intramural Research Program of the Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH); the Office of the Assistant Secretary for Preparedness and Response, Biomedical Advanced Research and Development Authority, Department of Health and Human Services (contract 75A50120C00034); and a research fellowship (to Dr. Corbett) that was partially funded by the Undergraduate Scholarship Program, Office of Intramural Training and Education, Office of the Director, NIH.

Disclosure forms provided by the authors are available with the full text of this letter at NEJM.org.

A preliminary version of this letter was published on February 17, 2021, at NEJM.org. This letter was published on March 17, 2021, at NEJM.org.

1. Jackson LA, Anderson EJ, Rouphael NG, et al. An mRNA vaccine against SARS-CoV-2 — preliminary report. N Engl J Med 2020;383:1920-31.
2. Anderson EJ, Rouphael NG, Widge AT, et al. Safety and immunogenicity of SARS-CoV-2 mRNA-1273 vaccine in older adults. N Engl J Med 2020;383:2427-38.
3. Baden LR, El Sahly HM, Essink B, et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. N Engl J Med 2021;384:403-16.

DOI: 10.1056/NEJMc2102179

Cardiovascular Outcomes with Sotagliflozin

TO THE EDITOR: In the SOLOIST-WHF trial (Effect of Sotagliflozin on Cardiovascular Events in Patients with Type 2 Diabetes Post Worsening Heart Failure), Bhatt et al. (Jan. 14 issue)[1] report that death from cardiovascular causes and hospitalizations and urgent visits for heart failure (the composite primary outcome) occurred in fewer patients in the sotagliflozin group than in the placebo group. This benefit was also seen in the SCORED trial (Effect of Sotagliflozin on Cardio-

vascular and Renal Events in Patients with Type 2 Diabetes and Moderate Renal Impairment Who Are at Cardiovascular Risk) reported by Bhatt et al. (also in Jan. 14 issue)[2] involving patients with chronic kidney disease, with or without albuminuria. Although formally correct, the reporting of the overall composite outcome distracts from the statistical significance of its individual components.[3] Chronic heart failure is associated with polypharmacy[4] and excess mortality.[5] The question in clinical practice is whether the inconvenience and cost of adding to polypharmacy are justified by a reduction in cardiovascular mortality. As Table 2 in the article on the SOLOIST-WHF trial and Table 2 in the article on the SCORED trial show, when this outcome is assessed individually, sotagliflozin does not reach statistical significance in reducing death from cardiovascular causes (hazard ratio, 0.84; 95% confidence interval [CI], 0.58 to 1.22, P=0.36; and hazard ratio, 0.90; 95% CI, 0.73 to 1.12, P=0.35, in the two articles, respectively). Since both trials were adequately powered, this finding is unlikely to be due to a type 2 error.

Piero Baglioni, M.D.

Prince Charles Hospital
Merthyr Tydfil, United Kingdom
Piero.Baglioni@wales.nhs.uk

No potential conflict of interest relevant to this letter was reported.

1. Bhatt DL, Szarek M, Steg PG, et al. Sotagliflozin in patients with diabetes and recent worsening heart failure. N Engl J Med 2021;384:117-28.
2. Bhatt DL, Szarek M, Pitt B, et al. Sotagliflozin in patients with diabetes and chronic kidney disease. N Engl J Med 2021;384:129-39.
3. Cordoba G, Schwartz L, Woloshin S, Bae H, Gøtzsche PC. Definition, reporting, and interpretation of composite outcomes in clinical trials: systematic review. BMJ 2010;341:c3920.
4. Baron-Franco B, McLean G, Mair FS, Roger VL, Guthrie B, Mercer SW. Comorbidity and polypharmacy in chronic heart failure: a large cross-sectional study in primary care. Br J Gen Pract 2017;67(658):e314-e320.
5. Butler J, Yang M, Manzi MA, et al. Clinical course of patients with worsening heart failure with reduced ejection fraction. J Am Coll Cardiol 2019;73:935-44.

DOI: 10.1056/NEJMc2102961

TO THE EDITOR: As reported by Bhatt et al., in the SOLOIST-WHF trial, severe hypoglycemia was more frequent with sotagliflozin than with placebo (1.5% vs. 0.3%). At baseline, type 2 diabetes was quite well controlled in the sotagliflozin and placebo groups (median glycated hemoglobin level, 7.1% [interquartile range, 6.4 to 8.3] and 7.2% [interquartile range, 6.4 to 8.2], respectively). Whether patients who presented with severe hypoglycemia were treated concomitantly with a sulfonylurea or insulin was not specified, as sodium–glucose cotransporter 2 (SGLT2) inhibitors or sotagliflozin alone do not increase this risk.[1]

There has been an accumulation of studies showing significant benefits with respect to heart failure and hard renal end points with this new drug class beyond its glucose-lowering effect.[1] Therefore, without doubt, SGLT2 inhibitors will soon be frequently prescribed by cardiologists and nephrologists. There will be a crucial need for more collaboration between these specialists and diabetologists, to adapt antidiabetic treatments at the time of introduction of SGLT2 inhibitors (especially for patients whose diabetes is well controlled, and for those treated with a sulfonylurea or insulin) to avoid the burden of severe hypoglycemia. We have an opportunity to think about how to enhance our collaboration and organization in order to improve patient care.

Lucien Marchand, M.D.
Emmanuel Villar, M.D., Ph.D.
Lisa Green, M.D.

Centre Hospitalier Saint Joseph Saint Luc
Lyon, France
lmarchand@ch-stjoseph-stluc-lyon.fr

Dr. Marchand reports receiving honoraria for speaking from Abbott, Lilly, Novo Nordisk, and Sanofi. No other potential conflict of interest relevant to this letter was reported.

1. Scheen AJ. Sodium-glucose cotransporter type 2 inhibitors for the treatment of type 2 diabetes mellitus. Nat Rev Endocrinol 2020;16:556-77.

DOI: 10.1056/NEJMc2102961

THE AUTHORS REPLY: The SOLOIST-WHF and SCORED trials both showed significant reductions in their primary composite end points, with consistency across subgroups and in the individual components of the primary end point, including death from cardiovascular causes. Neither trial was sufficiently powered for cardiovascular death since both trials ended prematurely owing to loss of funding at the onset of the Covid-19 pandemic.[1] However, the point estimate for cardiovascular death is entirely in line with a recent meta-analysis of SGLT2 inhibitors, and thus, as a class, SGLT2 inhibitors in high-risk populations

Table 1. Use of Sulfonylureas and Insulin at Baseline among Patients in the SOLOIST-WHF and SCORED Trials.*

Cohort and Variable	SOLOIST-WHF		SCORED	
	Sotagliflozin (N=608)	Placebo (N=614)	Sotagliflozin (N=5292)	Placebo (N=5292)
	number (percent)			
All enrolled patients				
Either sulfonylurea or insulin	315 (51.8)	314 (51.1)	4351 (82.2)	4385 (82.9)
Sulfonylurea	114 (18.8)	114 (18.6)	1400 (26.5)	1486 (28.1)
Insulin	217 (35.7)	217 (35.3)	3389 (64.0)	3333 (63.0)
Both sulfonylurea and insulin	16 (2.6)	17 (2.8)	438 (8.3)	434 (8.2)
Enrolled patients with severe hypoglycemia	9 (1.5)	2 (0.3)	53 (1.0)	55 (1.0)
Either sulfonylurea or insulin	7 (1.2)	1 (0.2)	51 (1.0)	53 (1.0)
Sulfonylurea	1 (0.2)	0	7 (0.1)	12 (0.2)
Insulin	6 (1.0)	1 (0.2)	49 (0.9)	49 (0.9)
Both sulfonylurea and insulin	0	0	5 (0.1)	8 (0.2)

* SCORED denotes Effect of Sotagliflozin on Cardiovascular and Renal Events in Patients with Type 2 Diabetes and Moderate Renal Impairment Who Are at Cardiovascular Risk, and SOLOIST-WHF Effect of Sotagliflozin on Cardiovascular Events in Patients with Type 2 Diabetes Post Worsening Heart Failure.

with sufficient duration of follow-up clearly reduce cardiovascular mortality.[2]

In the SOLOIST-WHF trial, in patients with diabetes who presented with acute decompensated heart failure, there was a small increase in severe hypoglycemia with sotagliflozin as compared with placebo (9 events vs. 2 events; P=0.04), which was carefully captured as an adverse event of special interest. A total of 18.8% of patients in the sotagliflozin group and 18.6% in the placebo group were receiving sulfonylureas at baseline; 35.7% and 35.3%, respectively, were receiving insulin (Table 1). During the double-blind treatment period, the management of glycemia and diabetes complications was left to the physician's judgment, informed by clinical guidelines. Investigators were instructed to record as serious adverse events episodes of hypoglycemia that required the assistance of another person to administer carbohydrate, glucagon, intravenous glucose, or other resuscitative measures. However, in the larger SCORED trial involving patients with stable diabetes and chronic kidney disease, there was no difference with respect to severe hypoglycemia between the sotagliflozin group and the placebo group (53 events and 55 events, respectively; P=0.84); 26.5% of patients in the sotagliflozin group and 28.1% in the placebo group were receiving sulfonylureas at baseline, and 64% and 63%, respectively, were receiving insulin. The majority of instances of severe hypoglycemia in both trials occurred in patients receiving insulin. Nevertheless, caution regarding concomitant diabetes medications is warranted at the time of initiation of SGLT2 inhibitors, especially in a patient recovering from a recent exacerbation of heart failure, where multiple medication additions and adjustments may be occurring simultaneously.

We agree that on the basis of the data, including the most recent additions from the SOLOIST-WHF and SCORED trials, SGLT2 inhibitors should soon be in common use by cardiologists and nephrologists,[3] and as a medical community we must ensure that there is close collaboration between these specialists and the primary care physicians and endocrinologists who are managing the overall glycemic control of the patient and, as for any drug, that there is awareness of the safety profile of this class among physicians in all medical fields who will be involved with prescribing these drugs.

Deepak L. Bhatt, M.D., M.P.H.

Harvard Medical School
Boston, MA
dlbhattmd@post.harvard.edu

P. Gabriel Steg, M.D.

Université de Paris
Paris, France

Bertram Pitt, M.D.

University of Michigan
Ann Arbor, MI

Since publication of their articles, the authors report no further potential conflict of interest.

1. Bagiella E, Bhatt DL, Gaudino M. The consequences of the COVID-19 pandemic on non-COVID-19 clinical trials. J Am Coll Cardiol 2020;76:342-5.
2. Zelniker TA, Wiviott SD, Raz I, et al. SGLT2 inhibitors for primary and secondary prevention of cardiovascular and renal outcomes in type 2 diabetes: a systematic review and meta-analysis of cardiovascular outcome trials. Lancet 2019;393:31-9.
3. Vaduganathan M, Sathiyakumar V, Singh A, et al. Prescriber patterns of SGLT2i after expansions of U.S. Food and Drug Administration labeling. J Am Coll Cardiol 2018;72:3370-2.

DOI: 10.1056/NEJMc2102961

Tocilizumab in Patients Hospitalized with Covid-19 Pneumonia

TO THE EDITOR: The *Journal* recently published two trials of tocilizumab for coronavirus disease 2019 (Covid-19) (Dec. 10 and Jan. 7 issues).[1,2] According to the articles, the primary outcome in both trials was mechanical ventilation or death. However, there are important discrepancies between those articles and the trial protocols and registry information.

In both trials, the original protocols and ClinicalTrials.gov records specify mechanical ventilation as the primary outcome. In the trial by Stone et al., the primary outcome was changed to include death in the final protocol and, almost 2 months after enrollment started, in the ClinicalTrials.gov record. In the trial by Salama et al., the primary outcome was changed in an amended protocol after enrollment had started, but the ClinicalTrials.gov record still specifies mechanical ventilation as the primary outcome. This protocol change is of particular importance, since the results for the "new" primary outcome were significant but the results for the original primary outcome were not reported in the article.

Outcome switching in trials may lead to bias,[3,4] and it is not clear why these important protocol changes were not described in the articles.

Andreas Lundh, M.D., Ph.D.

Hvidovre University Hospital
Hvidovre, Denmark
andreas.lundh@gmail.com

No potential conflict of interest relevant to this letter was reported.

This letter was published on March 3, 2021, at NEJM.org.

1. Stone JH, Frigault MJ, Serling-Boyd NJ, et al. Efficacy of tocilizumab in patients hospitalized with Covid-19. N Engl J Med 2020;383:2333-44.
2. Salama C, Han J, Yau L, et al. Tocilizumab in patients hospitalized with Covid-19 pneumonia. N Engl J Med 2021;384:20-30.
3. Altman DG, Moher D, Schulz KF. Harms of outcome switching in reports of randomised trials: CONSORT perspective. BMJ 2017;356:j396.
4. Chan A-W, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 2004;291:2457-65.

DOI: 10.1056/NEJMc2100217

DRS. SALAMA AND MOHAN REPLY: Lundh references potential bias introduced by postenrollment modification of the primary outcome with knowledge of the data. Such bias was not introduced in our trial, because the original planned analysis was unchanged and interim efficacy analyses were not performed. Death represents a competing risk for mechanical ventilation; therefore, a composite outcome is needed. Language to clarify the primary outcome was added in version 2 of the protocol to technically align with the original planned analysis, in which death before mechanical ventilation was considered a qualifying event in the evaluation of the primary outcome. Furthermore, version 2 of the protocol was finalized when less than 5% of the patients had been enrolled, while the trial was fully blinded, and well in advance of the internal review of unblinded safety data by the internal monitoring committee. For these reasons, we refute the suggestion that bias may have resulted from updating of the wording describing the primary outcome in the protocol.

P. Gabriel Steg, M.D.

Université de Paris
Paris, France

Bertram Pitt, M.D.

University of Michigan
Ann Arbor, MI

Since publication of their articles, the authors report no further potential conflict of interest.

1. Bagiella E, Bhatt DL, Gaudino M. The consequences of the COVID-19 pandemic on non-COVID-19 clinical trials. J Am Coll Cardiol 2020;76:342-5.
2. Zelniker TA, Wiviott SD, Raz I, et al. SGLT2 inhibitors for primary and secondary prevention of cardiovascular and renal outcomes in type 2 diabetes: a systematic review and meta-analysis of cardiovascular outcome trials. Lancet 2019;393:31-9.
3. Vaduganathan M, Sathiyakumar V, Singh A, et al. Prescriber patterns of SGLT2i after expansions of U.S. Food and Drug Administration labeling. J Am Coll Cardiol 2018;72:3370-2.

DOI: 10.1056/NEJMc2102961

Tocilizumab in Patients Hospitalized with Covid-19 Pneumonia

TO THE EDITOR: The *Journal* recently published two trials of tocilizumab for coronavirus disease 2019 (Covid-19) (Dec. 10 and Jan. 7 issues).[1,2] According to the articles, the primary outcome in both trials was mechanical ventilation or death. However, there are important discrepancies between those articles and the trial protocols and registry information.

In both trials, the original protocols and ClinicalTrials.gov records specify mechanical ventilation as the primary outcome. In the trial by Stone et al., the primary outcome was changed to include death in the final protocol and, almost 2 months after enrollment started, in the ClinicalTrials.gov record. In the trial by Salama et al., the primary outcome was changed in an amended protocol after enrollment had started, but the ClinicalTrials.gov record still specifies mechanical ventilation as the primary outcome. This protocol change is of particular importance, since the results for the "new" primary outcome were significant but the results for the original primary outcome were not reported in the article.

Outcome switching in trials may lead to bias,[3,4] and it is not clear why these important protocol changes were not described in the articles.

Andreas Lundh, M.D., Ph.D.

Hvidovre University Hospital
Hvidovre, Denmark
andreas.lundh@gmail.com

No potential conflict of interest relevant to this letter was reported.

This letter was published on March 3, 2021, at NEJM.org.

1. Stone JH, Frigault MJ, Serling-Boyd NJ, et al. Efficacy of tocilizumab in patients hospitalized with Covid-19. N Engl J Med 2020;383:2333-44.
2. Salama C, Han J, Yau L, et al. Tocilizumab in patients hospitalized with Covid-19 pneumonia. N Engl J Med 2021;384:20-30.
3. Altman DG, Moher D, Schulz KF. Harms of outcome switching in reports of randomised trials: CONSORT perspective. BMJ 2017;356:j396.
4. Chan A-W, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 2004;291:2457-65.

DOI: 10.1056/NEJMc2100217

DRS. SALAMA AND MOHAN REPLY: Lundh references potential bias introduced by postenrollment modification of the primary outcome with knowledge of the data. Such bias was not introduced in our trial, because the original planned analysis was unchanged and interim efficacy analyses were not performed. Death represents a competing risk for mechanical ventilation; therefore, a composite outcome is needed. Language to clarify the primary outcome was added in version 2 of the protocol to technically align with the original planned analysis, in which death before mechanical ventilation was considered a qualifying event in the evaluation of the primary outcome. Furthermore, version 2 of the protocol was finalized when less than 5% of the patients had been enrolled, while the trial was fully blinded, and well in advance of the internal review of unblinded safety data by the internal monitoring committee. For these reasons, we refute the suggestion that bias may have resulted from updating of the wording describing the primary outcome in the protocol.

In addition, ClinicalTrials.gov updates may lag behind requests for updates. We thank Lundh for bringing this administrative issue to our attention and have requested that the posting be corrected.

Carlos Salama, M.D.

Elmhurst Hospital Center–Icahn School of Medicine at Mount Sinai Hospital
New York, NY

Shalini V. Mohan, M.D.

Genentech
South San Francisco, CA
mohan.shalini@gene.com

Since publication of their article, the authors report no further potential conflict of interest.

This letter was published on March 3, 2021, at NEJM.org.

DOI: 10.1056/NEJMc2100217

DR. STONE REPLIES: I appreciate Lundh's point and agree that ideally a clinical trial is planned in elaborate detail, relatively free of time pressure, and none of the well-conceived plans are altered during its conduct. The reality of Boston in April 2020 dictated different circumstances for the start of our trial. Boston was in the throes of the first pandemic surge. At that time, not only was there no standard of care for Covid-19, but the appropriate outcome measures were also uncertain. After enrollment had begun and we were able to think further about analyzing the results to come, it made sense to expand the primary outcome to include not only the prevention of mechanical ventilation but also the prevention of death. We formally adjusted the primary end point in the protocol 6 weeks into the trial, notifying our institutional review board immediately and updating ClinicalTrials.gov at a later date (before the conclusion of the trial and before the locking and analysis of the data set). Ultimately, this modification of the primary outcome did not alter the primary or secondary conclusions of the trial.

John H. Stone, M.D., M.P.H.

Massachusetts General Hospital
Boston, MA
jhstone@mgh.harvard.edu

for the BACC Bay Tocilizumab Trial Investigators

Since publication of his article, the author reports no further potential conflict of interest.

This letter was published on March 3, 2021, at NEJM.org.

DOI: 10.1056/NEJMc2100217

Phase 3 Trial of Interleukin-1 Trap Rilonacept in Recurrent Pericarditis

TO THE EDITOR: Klein et al. (Jan. 7 issue)[1] report that recurrences of pericarditis after treatment with rilonacept occurred in 2 of 30 patients (7%), as compared with 23 of 31 patients (74%) who received placebo. This incidence of recurrence is far higher than the 15 to 30% reported in other studies, as cited by the authors. Although the number of patients in the trial is relatively small, this high incidence suggests that there might be something unusual about the population assessed. Do the authors have any explanation for this extraordinarily high incidence?

Edward Tabor, M.D.

Bethesda, MD
thousandyears@comcast.net

No potential conflict of interest relevant to this letter was reported.

1. Klein AL, Imazio M, Cremer P, et al. Phase 3 trial of interleukin-1 trap rilonacept in recurrent pericarditis. N Engl J Med 2021;384:31-41.

DOI: 10.1056/NEJMc2101978

TO THE EDITOR: The phase 3 trial RHAPSODY (Rilonacept Inhibition of Interleukin-1 Alpha and Beta for Recurrent Pericarditis: A Pivotal Symptomatology and Outcomes Study) assessed the interleukin-1 trap rilonacept in recurrent pericarditis. The results of this trial provide further confirmation of the efficacy of interleukin-1 inhibition in the treatment of pericarditis that has recurred despite treatment with glucocorticoids or colchicine. The striking effects of interleukin-1 inhibition in adults with recurrent pericarditis have been previously observed with the

In addition, ClinicalTrials.gov updates may lag behind requests for updates. We thank Lundh for bringing this administrative issue to our attention and have requested that the posting be corrected.

Carlos Salama, M.D.

Elmhurst Hospital Center–Icahn School of Medicine at Mount Sinai Hospital
New York, NY

Shalini V. Mohan, M.D.

Genentech
South San Francisco, CA
mohan.shalini@gene.com

Since publication of their article, the authors report no further potential conflict of interest.

This letter was published on March 3, 2021, at NEJM.org.

DOI: 10.1056/NEJMc2100217

DR. STONE REPLIES: I appreciate Lundh's point and agree that ideally a clinical trial is planned in elaborate detail, relatively free of time pressure, and none of the well-conceived plans are altered during its conduct. The reality of Boston in April 2020 dictated different circumstances for the start of our trial. Boston was in the throes of the first pandemic surge. At that time, not only was there no standard of care for Covid-19, but the appropriate outcome measures were also uncertain. After enrollment had begun and we were able to think further about analyzing the results to come, it made sense to expand the primary outcome to include not only the prevention of mechanical ventilation but also the prevention of death. We formally adjusted the primary end point in the protocol 6 weeks into the trial, notifying our institutional review board immediately and updating ClinicalTrials.gov at a later date (before the conclusion of the trial and before the locking and analysis of the data set). Ultimately, this modification of the primary outcome did not alter the primary or secondary conclusions of the trial.

John H. Stone, M.D., M.P.H.

Massachusetts General Hospital
Boston, MA
jhstone@mgh.harvard.edu

for the BACC Bay Tocilizumab Trial Investigators

Since publication of his article, the author reports no further potential conflict of interest.

This letter was published on March 3, 2021, at NEJM.org.

DOI: 10.1056/NEJMc2100217

Phase 3 Trial of Interleukin-1 Trap Rilonacept in Recurrent Pericarditis

TO THE EDITOR: Klein et al. (Jan. 7 issue)[1] report that recurrences of pericarditis after treatment with rilonacept occurred in 2 of 30 patients (7%), as compared with 23 of 31 patients (74%) who received placebo. This incidence of recurrence is far higher than the 15 to 30% reported in other studies, as cited by the authors. Although the number of patients in the trial is relatively small, this high incidence suggests that there might be something unusual about the population assessed. Do the authors have any explanation for this extraordinarily high incidence?

Edward Tabor, M.D.

Bethesda, MD
thousandyears@comcast.net

No potential conflict of interest relevant to this letter was reported.

1. Klein AL, Imazio M, Cremer P, et al. Phase 3 trial of interleukin-1 trap rilonacept in recurrent pericarditis. N Engl J Med 2021;384:31-41.

DOI: 10.1056/NEJMc2101978

TO THE EDITOR: The phase 3 trial RHAPSODY (Rilonacept Inhibition of Interleukin-1 Alpha and Beta for Recurrent Pericarditis: A Pivotal Symptomatology and Outcomes Study) assessed the interleukin-1 trap rilonacept in recurrent pericarditis. The results of this trial provide further confirmation of the efficacy of interleukin-1 inhibition in the treatment of pericarditis that has recurred despite treatment with glucocorticoids or colchicine. The striking effects of interleukin-1 inhibition in adults with recurrent pericarditis have been previously observed with the

interleukin-1 receptor antagonist anakinra in a randomized trial and in the International Registry of Anakinra for Pericarditis (IRAP) study.[1,2] Despite these impressive results observed while patients were receiving anti–interleukin-1 treatment, our previously published studies, as well as the results of the IRAP study, have shown that abrupt discontinuation of anakinra is associated with a high incidence of relapse (50 to 70%).[1,3,4] In light of these findings, the results from the extension period of the RHAPSODY trial are eagerly awaited. Until then, it should be emphasized that the excellent outcomes observed with rilonacept apply only during the time when patients are receiving treatment. Maintenance of long-term disease remission with well-characterized tapering protocols is the main challenge for the near future.

George Lazaros, M.D.
Konstantinos Tsioufis, M.D.
Dimitrios Vassilopoulos, M.D.

National and Kapodistrian University of Athens
Athens, Greece
glaz35@hotmail.com

Dr. Lazaros reports receiving advisory board compensation from Kiniksa Pharmaceuticals. No other potential conflict of interest relevant to this letter was reported.

1. Imazio M, Andreis A, De Ferrari GM, et al. Anakinra for corticosteroid-dependent and colchicine-resistant pericarditis: the IRAP (International Registry of Anakinra for Pericarditis) study. Eur J Prev Cardiol 2020;27:956-64.
2. Brucato A, Imazio M, Gattorno M, et al. Effect of anakinra on recurrent pericarditis among patients with colchicine resistance and corticosteroid dependence: the AIRTRIP randomized clinical trial. JAMA 2016;316:1906-12.
3. Lazaros G, Vasileiou P, Koutsianas C, et al. Anakinra for the management of resistant idiopathic recurrent pericarditis: initial experience in 10 adult cases. Ann Rheum Dis 2014;73:2215-7.
4. Lazaros G, Antonatou K, Vassilopoulos D. The therapeutic role of interleukin-1 inhibition in idiopathic recurrent pericarditis: current evidence and future challenges. Front Med (Lausanne) 2017;4:78.

DOI: 10.1056/NEJMc2101978

THE AUTHORS REPLY: We thank Tabor for his comments. Patients in the RHAPSODY trial were representative of the broad population of patients who have multiple recurrences of pericarditis and persistent disease. Although 15 to 30% of patients with incident pericarditis have a recurrence,[1] both risk and tempo of recurrent pericarditis events increase with each subsequent recurrence; patients with a second recurrence have a 54% chance of having a third recurrence.[2] In the RHAPSODY trial, 74% of patients in the placebo group had a recurrence of pericarditis, a finding consistent with these observational data. These data illustrate the tenacity of recurrent pericarditis in patients with persistent autoinflammation.

We concur with Lazaros et al. that the treatment objective in recurrent pericarditis is durable remission. Epidemiologic data suggest a 2-year mean disease duration.[3] Although the causes of the frequent recurrences in patients in the IRAP study who received treatment for only 6 as opposed to 9 months are unclear,[4] premature termination of interleukin-1 blockade in the context of persistent underlying autoinflammation may have been responsible for the observed outcomes. The median duration of rilonacept treatment in the RHAPSODY trial was 9 months (maximum, 14). The annualized incidence of pericarditis recurrence with treatment was 0.15 episodes per year, as compared with 4.42 events per year among all patients who received standard therapy before entering the trial. The two recurrences in the rilonacept group occurred during temporary treatment interruptions, thereby supporting the concept that continued rilonacept therapy resulted in continued treatment response; 74 of 75 eligible patients continued into the long-term extension period. The distinct mechanism of action and the gradual washout pharmacokinetics of rilonacept over approximately 5 to 8 weeks yielded a more measured magnitude and onset of recurrence events in the placebo group, indicating that tapering of rilonacept doses was unnecessary in the RHAPSODY trial. Recurrences resolved after reinitiation of rilonacept.

Decisions regarding treatment continuation or cessation involve many factors, including consideration of baseline characteristics as well as clinical status, biomarkers, and imaging at periodic intervals. Patients in the long-term extension phase will be assessed similarly, and the planned treatment duration of up to 2 years may be sufficient to allow underlying autoinflammation to resolve.[5] The planned Registry of the Natural History of Recurrent Pericarditis in Pediatric and Adult Patients (RESONANCE; ClinicalTrials.gov number, NCT04687358) trial will follow a broad patient population in which real-world treatment strategies will be used.

Allan L. Klein, M.D.

Cleveland Clinic
Cleveland, OH
kleina@ccf.org

Massimo Imazio, M.D.

Azienda Ospedaliero–Universitaria Città della Salute e della Scienza di Torino
Turin, Italy

John F. Paolini, M.D., Ph.D.

Kiniksa Pharmaceuticals
Lexington, MA

Since publication of their article, the authors report no further potential conflict of interest.

1. Adler Y, Charron P, Imazio M, et al. 2015 ESC Guidelines for the diagnosis and management of pericardial diseases: the Task Force for the Diagnosis and Management of Pericardial Diseases of the European Society of Cardiology (ESC) Endorsed by: the European Association for Cardio-Thoracic Surgery (EACTS). Eur Heart J 2015;36:2921-64.
2. Klein A, Cremer P, Kontzias A, et al. Recurrent pericarditis epidemiology in the United States (2013-2018). American College of Epidemiology 2019 (ACE), Pasadena, CA, September 7–10, 2019. poster.
3. Lin D, Laliberté F, Magestro M, et al. Recurrence burden in recurrent pericarditis: a US-based retrospective study of administrative healthcare claims. Circ Cardiovasc Qual Outcomes 2020;13:Suppl 1:A248. abstract.
4. Imazio M, Andreis A, De Ferrari GM, et al. Anakinra for corticosteroid-dependent and colchicine-resistant pericarditis: the IRAP (International Registry of Anakinra for Pericarditis) study. Eur J Prev Cardiol 2020;27:956-64.
5. Klein AL, Imazio M, Brucato A, et al. RHAPSODY: rationale for and design of a pivotal phase 3 trial to assess efficacy and safety of rilonacept, an interleukin-1α and interleukin-1β trap, in patients with recurrent pericarditis. Am Heart J 2020;228:81-90.

DOI: 10.1056/NEJMc2101978

Ewing's Sarcoma

TO THE EDITOR: In their review article on Ewing's sarcoma, Riggi et al. (Jan. 14 issue)[1] summarize the clinical, genetic, and epigenetic features of this aggressive bone sarcoma. However, their discussion of the experimental approaches is outdated and overly pessimistic, and they do not acknowledge the strong signals of activity that have been observed in phase 2 trials. In these trials, though most Ewing's sarcomas treated with single-agent insulin-like growth factor I receptor (IGF-IR)–directed therapies progressed, 10 to 14% of the patients had striking, albeit short-lived, tumor regression.[2,3]

Also, although the authors correctly highlight impediments to the clinical development of YK-4-279 (a preclinical lead compound),[4,5] it is surprising that they do not reference the exciting preliminary results of the ongoing multicenter, United States–based phase 1–2 study of the reformulated analog of YK-4-279 (TK-216).[6] The results of that study, which were presented at the European Society for Medical Oncology Congress in 2020, show that a subgroup of patients who received TK-216 had a complete and durable radiographic response. Given the heterogeneity described by Riggi and colleagues with respect to the cell of origin and fusion protein expression, the likely path forward will involve a nuanced treatment approach that matches each patient to patient-specific biomarkers.

Joseph A. Ludwig, M.D.

M.D. Anderson Cancer Center
Houston, TX

Paul A. Meyers, M.D.

Memorial Sloan Kettering Cancer Center
New York, NY

Uta Dirksen, M.D.

University Hospital Essen
Essen, Germany
uta.dirksen@uk-essen.de

No potential conflict of interest relevant to this letter was reported.

1. Riggi N, Suvà ML, Stamenkovic I. Ewing's sarcoma. New Engl J Med 2021;384:154-64.
2. Juergens H, Daw NC, Geoerger B, et al. Preliminary efficacy of the anti-insulin-like growth factor type 1 receptor antibody figitumumab in patients with refractory Ewing sarcoma. J Clin Oncol 2011;29:4534-40.
3. Pappo AS, Patel SR, Crowley J, et al. R1507, a monoclonal antibody to the insulin-like growth factor 1 receptor, in patients with recurrent or refractory Ewing sarcoma family of tumors: results of a phase II Sarcoma Alliance for Research through Collaboration study. J Clin Oncol 2011;29:4541-7.
4. Zöllner SK, Selvanathan SP, Graham GT, et al. Inhibition of the oncogenic fusion protein EWS-FLI1 causes G_2-M cell cycle arrest and enhanced vincristine sensitivity in Ewing's sarcoma. Sci Signal 2017;10(499):eaam8429.
5. Conn E, Hour S, Allegakoen D, et al. Development of an Ewing sarcoma cell line with resistance to EWS-FLI1 inhibitor YK-4-279. Mol Med Rep 2020;21:1667-75.
6. Ludwig JA, Federman N, Anderson P, et al. Phase I study of TK216, a novel anti-ETS agent for Ewing sarcoma. Ann Oncol 2020;31:Suppl 4:S972. abstract.

DOI: 10.1056/NEJMc2102423

Allan L. Klein, M.D.

Cleveland Clinic
Cleveland, OH
kleina@ccf.org

Massimo Imazio, M.D.

Azienda Ospedaliero–Universitaria Città della Salute e della Scienza di Torino
Turin, Italy

John F. Paolini, M.D., Ph.D.

Kiniksa Pharmaceuticals
Lexington, MA

Since publication of their article, the authors report no further potential conflict of interest.

1. Adler Y, Charron P, Imazio M, et al. 2015 ESC Guidelines for the diagnosis and management of pericardial diseases: the Task Force for the Diagnosis and Management of Pericardial Diseases of the European Society of Cardiology (ESC) Endorsed by: the European Association for Cardio-Thoracic Surgery (EACTS). Eur Heart J 2015;36:2921-64.
2. Klein A, Cremer P, Kontzias A, et al. Recurrent pericarditis epidemiology in the United States (2013-2018). American College of Epidemiology 2019 (ACE), Pasadena, CA, September 7–10, 2019. poster.
3. Lin D, Laliberté F, Magestro M, et al. Recurrence burden in recurrent pericarditis: a US-based retrospective study of administrative healthcare claims. Circ Cardiovasc Qual Outcomes 2020;13:Suppl 1:A248. abstract.
4. Imazio M, Andreis A, De Ferrari GM, et al. Anakinra for corticosteroid-dependent and colchicine-resistant pericarditis: the IRAP (International Registry of Anakinra for Pericarditis) study. Eur J Prev Cardiol 2020;27:956-64.
5. Klein AL, Imazio M, Brucato A, et al. RHAPSODY: rationale for and design of a pivotal phase 3 trial to assess efficacy and safety of rilonacept, an interleukin-1α and interleukin-1β trap, in patients with recurrent pericarditis. Am Heart J 2020;228:81-90.

DOI: 10.1056/NEJMc2101978

Ewing's Sarcoma

TO THE EDITOR: In their review article on Ewing's sarcoma, Riggi et al. (Jan. 14 issue)[1] summarize the clinical, genetic, and epigenetic features of this aggressive bone sarcoma. However, their discussion of the experimental approaches is outdated and overly pessimistic, and they do not acknowledge the strong signals of activity that have been observed in phase 2 trials. In these trials, though most Ewing's sarcomas treated with single-agent insulin-like growth factor I receptor (IGF-IR)–directed therapies progressed, 10 to 14% of the patients had striking, albeit short-lived, tumor regression.[2,3]

Also, although the authors correctly highlight impediments to the clinical development of YK-4-279 (a preclinical lead compound),[4,5] it is surprising that they do not reference the exciting preliminary results of the ongoing multicenter, United States–based phase 1–2 study of the reformulated analog of YK-4-279 (TK-216).[6] The results of that study, which were presented at the European Society for Medical Oncology Congress in 2020, show that a subgroup of patients who received TK-216 had a complete and durable radiographic response. Given the heterogeneity described by Riggi and colleagues with respect to the cell of origin and fusion protein expression, the likely path forward will involve a nuanced treatment approach that matches each patient to patient-specific biomarkers.

Joseph A. Ludwig, M.D.

M.D. Anderson Cancer Center
Houston, TX

Paul A. Meyers, M.D.

Memorial Sloan Kettering Cancer Center
New York, NY

Uta Dirksen, M.D.

University Hospital Essen
Essen, Germany
uta.dirksen@uk-essen.de

No potential conflict of interest relevant to this letter was reported.

1. Riggi N, Suvà ML, Stamenkovic I. Ewing's sarcoma. New Engl J Med 2021;384:154-64.
2. Juergens H, Daw NC, Geoerger B, et al. Preliminary efficacy of the anti-insulin-like growth factor type 1 receptor antibody figitumumab in patients with refractory Ewing sarcoma. J Clin Oncol 2011;29:4534-40.
3. Pappo AS, Patel SR, Crowley J, et al. R1507, a monoclonal antibody to the insulin-like growth factor 1 receptor, in patients with recurrent or refractory Ewing sarcoma family of tumors: results of a phase II Sarcoma Alliance for Research through Collaboration study. J Clin Oncol 2011;29:4541-7.
4. Zöllner SK, Selvanathan SP, Graham GT, et al. Inhibition of the oncogenic fusion protein EWS-FLI1 causes G_2-M cell cycle arrest and enhanced vincristine sensitivity in Ewing's sarcoma. Sci Signal 2017;10(499):eaam8429.
5. Conn E, Hour S, Allegakoen D, et al. Development of an Ewing sarcoma cell line with resistance to EWS-FLI1 inhibitor YK-4-279. Mol Med Rep 2020;21:1667-75.
6. Ludwig JA, Federman N, Anderson P, et al. Phase I study of TK216, a novel anti-ETS agent for Ewing sarcoma. Ann Oncol 2020;31:Suppl 4:S972. abstract.

DOI: 10.1056/NEJMc2102423

TO THE EDITOR: In response to the review article on Ewing's sarcoma by Riggi et al., we would like to add information from the French nationwide NetSarc (www.netsarc.org) database.[1-3] This database has been active in France since 2010 and includes expert pathological data and molecular review.

From 2013 through 2016 (the first period of the exhaustive data collection), 614 cases of Ewing's sarcoma were diagnosed among 66 million inhabitants of France (an incidence of 2.33 cases per 1 million inhabitants per year). The incidence of Ewing's sarcoma during this period was therefore 3 times as high as the generally reported incidence.[1] In 385 patients with Ewing's sarcoma (62.7%), the primary metastatic sites were the bone and bone marrow. Most of the patients were adults; the median age was 20 years (range, 1 to 93), 243 patients (39.6%) were younger than 18 years of age, and 188 (30.6%) were older than 30 years.

The management of soft-tissue sarcoma in national reference centers is associated with increased survival.[2,3] In the nationwide series in France from 2010 through 2016, survival among all adult patients with Ewing's sarcoma was significantly longer among patients treated initially in NetSarc reference centers than among those who were not treated in these centers. The risk of death was 46% lower (hazard ratio, 0.54) among the patients treated initially in the reference centers (Fig. 1). The treatment center was an independent prognostic factor in a multivariate analysis (data not shown). The initial therapeutic management of Ewing's sarcoma in a sarcoma reference center is a simple measure that is associated with a major reduction of the risk of death.

Jean-Yves Blay, M.D., Ph.D.

Centre Léon Bérard
Lyon, France
jean-yves.blay@lyon.unicancer.fr

Gonzague De Pinieux, M.D., Ph.D.

Université de Tours
Tours, France

Francois Gouin, M.D., Ph.D.

Centre Léon Bérard
Lyon, France

No potential conflict of interest relevant to this letter was reported.

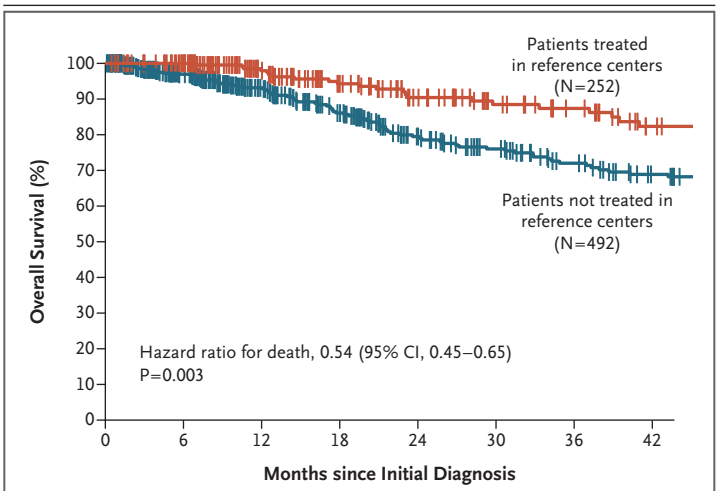


Figure 1. Overall Survival among Adults with Ewing's Sarcoma in France, 2010–2016.

Data on patients who received treatment at NetSarc reference centers and those who did not receive treatment at these centers are shown. CI denotes confidence interval.

1. Biermann JS, Chow W, Reed DR, et al. NCCN Guidelines insights: bone cancer, version 2.2017. J Natl Compr Canc Netw 2017;15:155-67.
2. Blay J-Y, Honoré C, Stoeckle E, et al. Surgery in reference centers improves survival of sarcoma patients: a nationwide study. Ann Oncol 2019;30:1143-53.
3. Bonvalot S, Gaignard E, Stoeckle E, et al. Survival benefit of the surgical management of retroperitoneal sarcoma in a reference center: a nationwide study of the French Sarcoma Group from the NetSarc database. Ann Surg Oncol 2019;26:2286-93.

DOI: 10.1056/NEJMc2102423

THE AUTHORS REPLY: In response to Ludwig et al.: it was our intent to focus the discussion of current and experimental therapies primarily on patients who had undergone sufficiently extensive evaluation in clinical trials to determine the usefulness of these therapies. We adopted the premise that successful therapy for Ewing's sarcoma, or any cancer for that matter, should imply long-term benefit. That is why we did not consider the short-term response to IGF-IR inhibitors in 10 to 14% of the patients in the trials mentioned by Ludwig et al. to be a success and why we did not include it in our discussion. Several studies have highlighted the plethora of experimental, targeted, anticancer therapies that may elicit transient responses but provide little if any long-term benefit to patients, often at a disproportionate cost to both society and the patients themselves.[1-3] Ac-

cordingly, we chose to limit our discussion to currently accepted therapy for Ewing's sarcoma and to describe the directions in which the field is moving by covering some of the mechanism-based experimental approaches that have been tried. However, the limited effectiveness of experimental therapies in Ewing's sarcoma thus far in no way casts a pessimistic view on the potential success of future endeavors. The phase 1–2 trial results with TK-216 that Ludwig et al. describe are of interest but were published when our review article was already in press. Furthermore, although these results may be encouraging, they are preliminary, and a longer perspective is needed to determine the long-term effectiveness of TK-216 and its potential analogues.

Nicolò Riggi, M.D., Ph.D.

Centre Hospitalier Universitaire Vaudois
Lausanne, Switzerland

Mario L. Suvà, M.D., Ph.D.

Massachusetts General Hospital
Boston, MA

Ivan Stamenkovic, M.D.

Centre Hospitalier Universitaire Vaudois
Lausanne, Switzerland
ivan.stamenkovic@chuv.ch

Since publication of their article, the authors report no further potential conflict of interest.

1. Fojo T, Mailankody S, Lo A. Unintended consequences of expensive cancer therapeutics — the pursuit of marginal indications and a me-too mentality that stifles innovation and creativity. JAMA Otolaryngol Head Neck Surg 2014;140:1225-36.
2. Davis C, Naci H, Gurpinar E, Poplavska E, Pinto A, Aggarwal A. Availability of evidence of benefits on overall survival and quality of life of cancer drugs approved by European Medicines Agency: retrospective cohort study of drug approvals 2009-13. BMJ 2017;359:j4530.
3. Seruga B, Ocana A, Amir E, Tannock IF. Failures in phase III: causes and consequences. Clin Cancer Res 2015;21:4552-60.

DOI: 10.1056/NEJMc2102423

INSTRUCTIONS FOR LETTERS TO THE EDITOR

Letters to the Editor are considered for publication, subject to editing and abridgment, provided they do not contain material that has been submitted or published elsewhere.

Letters accepted for publication will appear in print, on our website at NEJM.org, or both.

Please note the following:

- Letters in reference to a *Journal* article must not exceed 175 words (excluding references) and must be received within 3 weeks after publication of the article.
- Letters not related to a *Journal* article must not exceed 400 words.
- A letter can have no more than five references and one figure or table.
- A letter can be signed by no more than three authors.
- Financial associations or other possible conflicts of interest must be disclosed. Disclosures will be published with the letters. (For authors of *Journal* articles who are responding to letters, we will only publish new relevant relationships that have developed since publication of the article.)
- Include your full mailing address, telephone number, fax number, and email address with your letter.
- All letters must be submitted through our online submission system at NEJM.org.

Letters that do not adhere to these instructions will not be considered. We will notify you when we have made a decision about possible publication. Letters regarding a recent *Journal* article may be shared with the authors of that article. We are unable to provide prepublication proofs. Submission of a letter constitutes permission for the Massachusetts Medical Society, its licensees, and its assignees to use it in the *Journal*'s various print and electronic publications and in collections, revisions, and any other form or medium.

THE *JOURNAL*'S WEB AND EMAIL ADDRESSES

To submit a letter to the Editor: authors.NEJM.org

For information about the status of a submitted manuscript: authors.NEJM.org

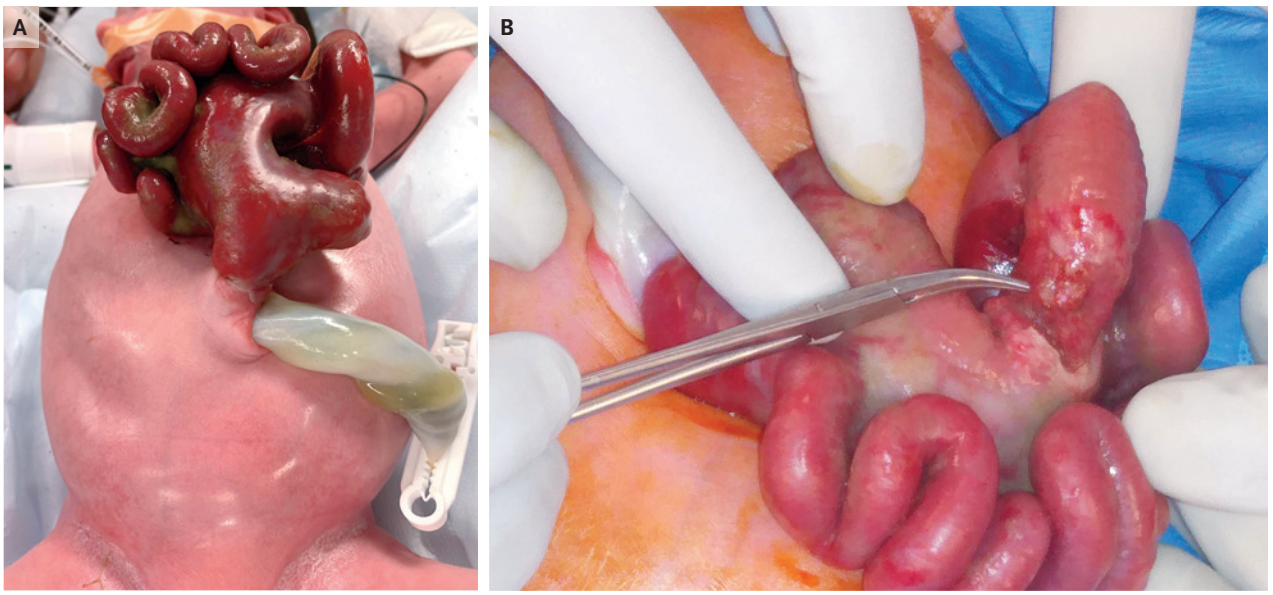
To submit a meeting notice: meetingnotices@NEJM.org

The *Journal*'s web pages: NEJM.org

IMAGES IN CLINICAL MEDICINE

Chana A. Sacks, M.D., *Editor*

Closed Gastroschisis



Tyler D. Robinson, M.D., M.P.H.
Christine Whyte, M.D.

Albany Medical Center
Albany, NY
robinst8@amc.edu

This article was published on April 10, 2021, at NEJM.org.

In a male neonate born at 37 weeks of gestation, a narrow stalk of small bowel and proximal colon protruded through a 1.5-cm defect in the abdominal wall. A diagnosis of gastroschisis had been made antenatally on the basis of findings from prenatal ultrasonography. In the area where the bowel had traversed the defect, the jejunum was atretic and the colon was stenosed (Panel A). The term "gastroschisis" refers to the evisceration of the intestines through a defect in the abdominal wall; the condition is referred to as closed gastroschisis when the defect closes, causing incarceration of the eviscerated bowel. In this infant, laparotomy was performed on the first day of life. The cecum was found to be perforated (Panel B), and the proximal jejunum ended blindly within the abdomen. The colonic perforation was sutured, the bowels were reduced, and the abdomen was closed. An additional laparotomy was performed 19 days later to repair the jejunal atresia and colonic stenosis. After transitioning from parenteral to enteral nutrition at 27 days, the boy was discharged home, at 1 month of age.

DOI: 10.1056/NEJMicm2029281

EDITORIAL

Audio Interview: Vaccine Successes and Vaccine Adverse Events

Eric J. Rubin, M.D., Ph.D., Lindsey R. Baden, M.D., and Stephen Morrissey, Ph.D.

The continuing spread of SARS-CoV-2 remains a Public Health Emergency of International Concern. What physicians need to know about transmission, diagnosis, and treatment of Covid-19 is the subject of ongoing updates from infectious disease experts at the *Journal*.

In this audio interview conducted on April 14, 2021, the editors discuss the current state of Covid-19 vaccination, including the rare occurrence of thrombotic thrombocytopenia in recipients of the ChadOx1 nCoV-19 and Ad26.COV2.S vaccines.

Disclosure forms provided by the authors are available at NEJM.org.

DOI: 10.1056/NEJMe2106379

An audio interview with Eric Rubin and Lindsey Baden is available at NEJM.org