Unmatched:
Repairing the U.S. Medical Residency Pipeline

Robert Orr
September 2021
The Niskanen Center is a nonpartisan 501(c)(3) think tank that works to promote an open society.

Niskanen Center
820 1st Street NE, Suite 675
Washington, D.C. 20002
NiskanenCenter.org
Unmatched:
Repairing the U.S. Medical Residency Pipeline

Robert Orr
September 2021
Contents

5 Introduction
7 The three eras of the modern U.S. residency system
10 1945–1980: The era of expansion and abundance
14 1981–2000: The era of modernization, federal retrenchment and downsizing
16 1997–present: Cumulative and persistent imbalances
16 Quantifying the residency bottleneck
20 Our broken, convoluted system of GME financing
25 Additional barriers to primary care residents
27 Policy recommendations
27 Consolidate and standardize federal GME payments
28 Reorganize GME funding
28 Abolish institution-level FTE caps on residency support
30 Vary federal payments to teaching institutions based on program size
30 Increase payments based on the share of PG-1 residents
31 End support for training beyond the 5th year of residency
32 Ensure federal GME payments are geographically uniform
33 Give states flexibility to target Medicaid GME payments
33 Plan for an organized transition to the new system
34 Establish a health care resources development bank
35 Conclusion
36 Appendix 1: Using NRMP data
38 Appendix 2: GME funding sources
39 Appendix 3: Hypothetical funding schedules

About the Author

Robert Orr is a poverty and welfare policy analyst at the Niskanen Center. He earned his BA at Skidmore College and his MA in Economics at George Mason University. His research focuses on welfare, health care, and economic development.

For media inquiries, please contact Louisa Tavlas at ltavlas@niskanencenter.org
Physicians are the cornerstone of any health care delivery system. Nonetheless, the United States employs substantially fewer of these professionals than most other developed countries. Given that the number of physicians per person is a reasonable proxy for the number of service access points within a health system, the United States therefore faces greater challenges in delivering basic, accessible care to its citizens compared to other developed countries.12

At the most abstract level, this lack of physicians is a puzzle. The combined cost of a medical degree and postgraduate residency training is many multiples smaller than the expected lifetime earnings of those who complete residency and become a licensed physician. This raises the question: If the financial incentives are there, why do we have so few doctors?

The key reason for our physician shortfall is that substantial bottlenecks exist in the training and education pipeline, with the most significant of these being the medical residency system. These bottlenecks were initially encouraged by policymakers based upon the misguided belief that limiting physician supply would control rising health care costs.3 Yet these attempts to constrain physician supply growth beginning in the 1980s have been an utter policy failure, leading to no noticeable reduction in health care costs.4 Savings have failed to materialize because when physicians are scarce, the money simply gets spent elsewhere in the health system. All the indications are that squeezing physician supply is much like squeezing a balloon — the costs largely migrate rather than disappear altogether. Because physicians are the U.S. health care system’s primary access point, the health system has effectively been encouraged to pursue high-intensity, low-access care. This shift has been harmful because provision of the most basic medical services is generally recognized to have the greatest marginal impact on population-level health.5 The United States’ failure to allocate

---

2 Health Resources and Services Administration, “MUA Find,” 2021.
health resources effectively helps explain why our health outcomes fall short of what we would expect given our country's exceptional level of health spending. The United States thus has little to lose and much to gain from increasing the supply of physicians across the board. The most effective course of action to achieve this would be to reform and modernize federal support for the physician residency system.

Figure 1: The U.S. has fewer physicians per-capita than most OECD countries

This report unpacks the policy failures currently plaguing the residency system in the United States and proposes how we might solve them. The first section provides an overview of the history of graduate medical education (GME) from the standpoint of both the public and private institutional actors. The second section examines the residency pipeline, confirming the existence of an undesirable bottleneck in the current system. The third section analyzes the shortcomings that result from the current mechanisms of GME financing. The final section focuses on how our GME system might be reformed in order to resolve these issues over the long term.


The three eras of the modern U.S. residency system

Knowing the history of U.S. graduate medical education (GME) is critical to a full understanding of our current predicament. Consequently, this section introduces the key players in GME, and sheds light on their interests and ideas; explains the role of the federal government, in particular; and reveals how our system has been developed in a patchwork fashion. The system supporting GME in our country has ultimately achieved policymakers’ misguided top-line goal of limiting physician supply, but along the way, it has encouraged misallocation of health care resources and entrenched geographic inequities. The current overview will only cover the period following World War II, both because the recognizable outlines of the residency system were still taking shape and because the role of federal policymaking was minimal up to that point.

1945–1980: The era of expansion and abundance

Federal involvement in GME emerged during an era of rapid transformation of the American residency system and the health system more broadly. Following World War II, booming medical school enrollment (financed in part by the Servicemen’s Readjustment Act of 1944, also known as the G.I. Bill, and later by the Health Professions Educational Assistance Act of 1963), on top of increased interest in specialty training among established physicians, resulted in surging demand for postgraduate medical training — that is, additional training beyond the MD level. At the same time, federal matching funds provided through the Hill–Burton Act of 1946 led to a substantial expansion in hospital construction in communities across the country. The resulting confluence of supply- and demand-side support effectively cemented the teaching hospital as the primary venue for postgraduate medical training in the U.S., a distinction it retains to the present day.

The rapid expansion of the U.S. health system during this period created fierce competition among hospitals for labor in the form of medical interns and residents. As a result, hospitals began pressuring medical students to sign commitment contracts earlier and earlier in their medical school tenures, eventually pushing the need for a decision into their second year. The National Intern Matching Program (later renamed the National Resident Matching Program, and colloquially known as “The Match”) was established as a centralized nonprofit clearinghouse to put an end to this rat race. When The Match was first established in 1952, there were only 5,800 medical graduates available to fill roughly 10,500 positions. In other words, the supply of potential residents was exceedingly scarce relative to the demand from hospitals.

With the establishment of Medicare in 1965, the federal government began providing direct payments to cover the hospitals’ teaching costs. Medicare made payments to hospitals to cover its

9 Ibid., Page 170.
10 Ibid., Page 169.
11 Ibid., Page 172.
patients’ share of overall operating costs, submitted retrospectively each fiscal year based upon incurred costs, with teaching costs included in these payments. The intent was to replace similar payments that private insurers had previously been making on behalf of elderly patients who were now covered by Medicare. These payments had grown considerably due to both the increasing complexity of care and the growing demands by interns and residents for salaries and fringe compensation.13 In these early years of the Medicare program, cost control was largely nonexistent, resulting in substantial financial support for medical training and education programs.14 It also resulted in wide disparities in federal support across institutions, mostly stemming from disparities in the cost of training programs as well as from the institutions’ having an incentive to report higher costs in order to maximize their Medicare payments.15 However, despite growing costs, the overriding preoccupation of policymakers was to ensure the sufficiency of health provider capacity in the face of rapidly growing demand for health care services. In addition to the funding delivered through Medicare, some states began contributing federally matched GME funding through the concurrently established Medicaid health

14 Committee on Implementing a National Graduate Medical Education Trust Fund, On Implementing a National Graduate Medical Education Trust Fund, (Washington D.C: National Academy Press, 1997), Pages 54–64.
insurance program for low-income Americans. However, no formal arrangement was ever put in place to earmark GME contributions from private insurers, so the residency system gradually became reliant upon government support, delivered overwhelmingly through Medicare.

Concurrent with the increasing flows of federal dollars into the residency system came calls to regularize and raise training standards. At the time, the disparate residency programs were governed by numerous autonomous Residency Review Committees (RRCs) in the individual specialties and subspecialties. In 1972, the American Medical Association, Association of American Medical Colleges, American Board of Medical Specialties, American Hospital Association, and Council of Medical Specialty Societies came together to create the Liaison Committee for Graduate Medical Education (later reorganized as the Accreditation Council on Graduate Medical Education) to oversee the RRCs. Three years later, the option to pursue a single-year rotating internship as a pathway to general practice was eliminated. From that point onward, physicians looking to pursue a more generalist practice would need to spend at least three years in residency programs, often via internal medicine or the newly created family medicine specialty. The sole point of entry into physician practice was now through specialist residency training (Figure 2).

Beginning in 1974, successive attempts were made to curtail cost growth in Medicare through the placement of limits on per-patient reimbursements to hospitals, initially excluding GME from cost controls, as policymakers also sought to maintain the government’s commitment to training physicians. Thus, even as cost limits continued to tighten, exceptions were eventually established for the financing of GME, two of which continue to inform the basis of Medicare GME financing even today. The first, in 1979, established an exclusion from reimbursement cost limits for all direct costs of teaching at hospitals. The second, established in 1980, increased the general reimbursement cost limits based on a hospital’s ratio of interns and residents to hospital beds, in order to provide compensation for the higher indirect costs of care at teaching hospitals. Even after Medicare phased out cost-based reporting during the 1980s, reimbursement of hospitals based on this conceptual distinction between “Direct Medical Education” (DME) and “Indirect Medical Education” (IME) continued, remaining in roughly the same form up to the present day.

20 Committee on Implementing a National Graduate Medical Education Trust Fund, On Implementing a National Graduate Medical Education Trust Fund.
22 Ibid., Page 224.
23 Continued cost growth in Medicare prompted its transition to the Prospective Payment System (PPS) for acute-care hospitals in 1983, replacing the old system of cost-based reimbursement. The IME and DME funding streams for residency training were formalized and considerably revamped in the process. Indirect Medical Education (IME) was converted into an add-on adjustment to the new PPS system, boosting payments per discharge on the basis of a hospital’s intern-and-resident-to-bed ratio (IRB). For Direct Medical Education (DME), the direct training costs reported in 1984 were carried over in the form of a Per Resident Amount (PRA). As with the initial iteration of DME, Medicare’s support was scaled in line with the hospital’s share of Medicare patients.
1981 to 2000: The era of modernization, federal retrenchment and downsizing

The tide turned decisively in the 1980s and 1990s, as the conventional wisdom on physician supply swung from shortage to surplus. The federal government retrenched its support of GME through Medicare while nongovernmental bodies overseeing the medical training pipeline began to promulgate policies aimed at curbing the growth in the number of U.S. physicians.

In 1981, a federally sponsored report by the Graduate Medical Education National Advisory Committee (GMENAC) concluded that the United States was on the verge of a massive physician surplus. Motivated chiefly by cost concerns, the committee reached this conclusion through estimating the physician workforce needed to meet the country’s projected medical needs. The report recommended immediate action to curtail both the domestic training of physicians and the admittance of those trained outside of the country. Support for both medical schools and medical students was thus sharply scaled-back. These moves coincided with a broader skepticism of government during the “Reagan Revolution” of the 1980s, resulting in a retrenchment of federal support for investment in the health system. Around the same time, federal support for the construction of hospitals, the sites in which residency training takes place, was totally cut off (Figure 3).

Robert Graham, the administrator of the Health Resources and Services Administration, concisely summed up the attitudes of the era toward physician supply when he stated in 1981:

“I believe that the Administration’s position can be interpreted as follows: (1) The general supply of health professionals is adequate or the capacity of the U.S. health professionals schools to produce the needed supply is perceived to be adequate, (2) there will be a minimum federal role for investment in terms of health professions education whether we are talking about direct project grants, institutional assistance, student aid, or other support, and (3) competition will sort out the major issues of distribution, specialty choice, and workforce mix.”

The cause of reducing physician numbers was often taken up with great fervor by members of


25 Physician workforce forecasts models have historically been influential in shaping the conventional wisdom around physician workforce sufficiency, both in the United States as well as abroad. 1) The methods involved in these forecasts have evolved considerably, from rather crude projections based on indicators of workforce size and health care demand to today’s more sophisticated approaches integrating considerations such as productivity and the substitution of tasks between various health professions. 2) Not only has the accuracy of the inputs to the forecasts been demonstrated to be limited in practice, but they also rest upon significant assumptions about the sufficiency of care of the status quo. 3) Important considerations such as the quality of care or provider market power remain outside the scope of these forecasts. 1) Sabine Stordeur and Christian Léonard, “Challenges in physician supply planning: the case of Belgium,” Human Resources for Health Vol. 8 No. 28 (December 2010); 2) Mário Amorim Lopes, Álvaro Santos Almeida, and Bernardo Almada-Lobo, “Handling healthcare workforce planning with care: where do we stand?,” Human Resources for Health Vol. 13 No. 38 (May 2015); 3) Dominique Roberfroid, Christian Leonard, and Sabine Stordeur, “Physician supply forecast: better than peering in a crystal ball?,” Human Resources for Health Vol. 7 No. 10 (February 2009).


28 Barbara Barzansky and Norman Gevitz, Beyond Flexner: Medical Education in the Twentieth Century (United States: Praeger, 1992), Page 119.
the profession, partly out of fear that a failure to reduce numbers would invite greater government control over the profession. As Charles Evarts, the president of the American Orthopaedic Association and a Residency Review Committee member, put it in a 1985 speech:

"Manpower control is mandatory. The size of medical schools must be diminished; there must be a strict limit upon foreign medical school graduates. Certain programs need to reduce their numbers, others must consolidate, and others need to terminate voluntarily or be terminated — not the easiest of actions. Currently there are no official agencies to directly mandate numbers. However, the impending decrease in funding for graduate medical education will directly influence the numbers. Before our destiny is controlled for us, we must step into the breach and exercise considerable leadership. Many pressures will be brought to bear against those who decide to decrease the numbers of orthopedic residents in training. Yet, there is no other choice. We cannot succumb to the temptation of increasing numbers. We must slowly decrease the numbers and continue to increase the quality."

The physician pipeline’s various nongovernmental accrediting and oversight bodies responded by taking swift action to reduce numbers. In 1980, the MD-granting medical schools agreed to a voluntary freeze on new slots as well as to the construction of additional medical schools.

The accrediting boards also took action. The Accreditation Council on Graduate Medical Education — the successor to the Liaison Committee for Graduate Medical Education — and many of its affiliated RRCs raised the standards for teaching hospitals. The intent was to make residency programs harder to establish and less financially attractive to operate. The accreditors’ actions frequently resulted

---


31 A quote from obstetrics–gynecology illustrates these intentions: The Residency Review Committee (RRC) for obstetrics–gynecology has, as do many of the RRCs, new and more stringent special requirements, the rules and regulations under
in a lowering of the caps on the number of residents in approved programs and outright eliminated programs deemed to be of insufficient quality. A side-effect of these retrenchment measures was a doubling down on technical education rather than the vocational aspects of the residency experience, an emphasis that continues to define the U.S. system to the present day.

The growing adoption of health maintenance organizations (HMOs) and other forms of managed care added to the pressure to curb physician supply. HMOs were tightly managed, vertically integrated health-service networks that required central approval for many procedures and doctor visits, with the ultimate goal of curbing health care utilization. Managed care spread rapidly during the 1980s and 1990s, with the majority of Americans being insured under such an arrangement by 1995. In light of this, between 1988 and 2000, the congressionally authorized Council on Graduate Medical Education repeatedly endorsed the need to constrain physician numbers in order to avoid a “surplus” in its annual reports.

which residency programs are approved. Because of stricter scrutiny of residency education programs and because of financial constraints on hospitals, the coming decade will see, at the best, a continued plateau in resident numbers. What is more likely is a decline in the number of approved residency programs and a slow decline in the number of residents per year.


This exuberance toward the prospects of curbing costs through managed care, which tightly manage services offered by
But even as high-level administrators sought to control the supply of physicians, new and distinct RRCs governing the specialties and subspecialties proliferated and created new pipelines for interns and residents. The result was to make limiting overall physician numbers more difficult than anticipated while producing an imbalance in the make-up of the physician workforce: too many specialists, too few primary care doctors.37 While the conventional wisdom over this period remained ardent that the nation was on the verge of an overall “physician surplus,” the rapidly declining share of primary care physicians had begun to raise alarm in some quarters.38

Aiming to preempt government action, the Accreditation Council for Graduate Medical Education (ACGME) enacted a 2-year moratorium on the recognition of new subspecialty RRCs between 1992 and 1993.39 Yet this didn’t stop the Clinton administration’s failed Health Security Act from proposing that the federal government take direct control over the approval of federally funded residency programs.40 Ultimately, though, Congress ended up merely freezing Medicare’s GME inflation adjustments for non-primary care specialties, which lasted for two years.41

Figure 5: Residency bottlenecks incentivized an explosion in subspecialties.

Source: ACGME Program Accreditation Data
This trend of retrenchment of federal support for GME culminated with the 1997 Balanced Budget Act (BBA). The BBA impacted Medicare funding for medical residencies in two important ways. First, the legislation reduced the generosity of the formula through which hospitals that have residents receive increased Medicare reimbursement. Second, and most famously, it capped Medicare’s number of funded residency slots, adjusted for inflation, at 1996 levels. Because GME funding is allocated on a hospital-specific basis, the introduction of the cap largely froze the geographic distribution of these slots. Residency programs could continue to add slots above the cap, but would receive no Medicare funding for doing so.

1997 to present: Cumulative and persistent imbalances

Recent years have seen the culmination of various trends that had been building over the previous two decades, and a gradual shift toward expanding physician supply. Starting in the late 1980s, increasing competition and declining profit margins made it harder for private insurers to pass on the higher costs associated with residency training to consumers through their premiums. The implicit understanding when Medicare’s GME funding stream was established was that it would replace merely the contributions that would have been made otherwise by private insurers had Medicare never existed, yet there was never any mechanism devised to ensure that private reimbursements would be earmarked for training, with the result that medical residency funding became increasingly reliant upon Medicare. Yet since this funding stream was scaled based on a hospital’s Medicare-patient load, it created a bias against the training of pediatric specialists in children’s hospitals. In response, Congress established the Children’s Hospital Graduate Medical Education (CHGME) program in 1999.

By the late 1990s, the constraints on physician numbers had translated into increasing market power, granting doctors leverage over managed care plans. Physicians increasingly felt empowered to stay out of plans that limited utilization, resulting in narrower provider networks for those plans. The tight labor markets at the turn of the new century delivered the finishing blow to what had, up until then, been heralded as the "managed-care revolution." Employers, intent on attracting and retaining employees under conditions of low unemployment, became willing to accept large health insurance premium increases. Limits on access to services and choice among providers largely reverted to how they had been in the pre-managed care days.

42 Legislation enacted in the years following the BBA have since increased DGME’s cap for rural hospitals and redistributed some unused slots, and created a mechanism for reallocating slots that would have otherwise been lost following the closure of resident training programs.


45 The act also established a floor and ceiling on per-resident funding of 70 percent and 140 percent of the national average per resident. Prior to that point, per-resident funding ranged between approximately $10,000 and $240,000. See: Daniel Guss, Ann L. Prestipino, and Harry E. Rubash, "Graduate medical education funding: a Massachusetts General Hospital case study and review," The Journal of Bone and Joint Surgery Vol. 94 No. 4 (February 2012).


The rate of health care consumption growth accelerated, swiftly converging with its pre-managed care trajectory. Thus, ironically, the effort to accommodate the transition toward tightly managed HMO care by restricting the entry of physicians into the profession had instead helped to facilitate its unraveling.

By this time, the balance of power in the residency pipeline had shifted completely from residents to hospitals. Hospitals effectively utilized their power as gatekeepers into the supply-constrained profession, resulting in sluggish resident wage growth despite the long hours. Growing dissatisfaction among residents resulted in organized action against the institutions governing entry into the profession. Opposition among medical students to the National Resident Matching Program’s algorithm led to its design being shifted from program-optimal to applicant-optimal in 1997. In 2002, an antitrust lawsuit was filed alleging anti-competitive collusion between the ACGME, The Match, and various medical associations. Congress responded by passing a law in 2002 exempting matching programs from antitrust action, but still, the message had been received. In 2003, ACGME instituted new duty-hour restrictions on residents. Under the new rule, the hours of clinical and educational activity conducted by medical residents would be capped at 80 hours per week, with specialty boards able to petition for an additional 8 hours.

While Medicare GME cutbacks continued into the middle of the decade, the perceived wisdom on physician supply had begun to evolve. By the early 2000s, awareness began to grow that prior

---


projections of physician surpluses were misguided. In 2005, the Association of American Medical Colleges began to sound the alarm of an impending "physician shortage," calling for a 10 percent expansion in medical school enrollment and raising the recommendation to 30 percent the following year. After two and half decades, the moratorium on allopathic medical enrollment was effectively lifted, resulting in substantial growth in the number of medical school graduates each year.

The various physician associations that had supported and lobbied on behalf of Medicare’s cap on graduate medical education slots, such as the American Medical Association (AMA), began to gradually flip their position on the topic. And by 2009, legislation to repeal the cap on Medicare GME was introduced in Congress; now with the support of the organizations that had originally supported the cap.

Federal policy gradually became more accommodative of GME funding. The Affordable Care Act established the Teaching Health Center Graduate Medical Education program in order to increase funding for primary care training programs operating in community-based ambulatory settings. Like the children’s hospital GME funding program before it, the new program’s impetus responded to biases in Medicare’s GME funding, which had provided an indirect-medical-expense adjustment solely for hospital inpatient care. Yet over the intervening years, the share of care provided in outpatient settings had increased considerably.

The most recent change came with the enactment of the Consolidated Appropriations Act of 2021, which added 1,000 new Medicare GME slots, distributed over the upcoming years, with 40 percent reserved for hospitals meeting certain criteria, such as being in a rural area. The law also made various minor fixes, such as allowing hospitals that had inadvertently established a low reimbursement rate for residency training to have another shot. But as discussed next, these incremental changes fall far short of what is needed to solve our physician shortage.

Quantifying the Residency Bottleneck

The United States’ dearth of physicians per capita relative to peer countries is not because of lower interest, less capable citizens, or less generous remuneration. The United States has fewer physicians because of the remarkable degree to which its physician pipeline filters out entrants into the profession. While medical schools are once again admitting more graduates, the extent to which they can do so is ultimately constrained by the supply of residencies.

Medical residencies are an odd institution from an economic standpoint due to the manner in which they blur the line between education and employment. On the one hand, a residency forms

51 Richard A. Cooper, "There's a Shortage of Specialists: Is Anyone Listening?,” Academic Medicine : Journal of the Association of American Medical Colleges Vol. 77 No. 8 (August 2002); Pages 761–766.
52 Da’Shia Davis, Michael Dill, Projected Shortage of Physicians through 2030, AAMC Workforce Studies Data Snapshot (United States: Association of American Medical Colleges, May 2018).
54 Kevin B. O'Reilly, "Federal funding for Medicaid program should not be capped: AMA," American Medical Association, June 16, 2017.
a crucial step in a physician’s professional education. Residency programs allow medical graduates to gain valuable hands-on experience that isn’t sufficiently imparted during medical school in a supervised fashion. Residents also take part in significant amounts of classroom learning, particularly during the initial year. Much as in university, this includes a specified curriculum and performance evaluations watched over by full-time program administrators, as stipulated by the ACGME and its affiliated RRCs. These requirements place substantial financial demands on teaching institutions, both in terms of the direct expenditures and the opportunity costs imposed on teaching physicians who could otherwise be providing additional services.

Yet unlike other teaching institutions, teaching hospitals pay residents, rather than the other way around. At the most basic level, a resident exchanges their labor for the educational experience and pathway to eventual licensure provided by the hospital, on top of any financial compensation. In return, residents contribute substantial hours of labor to the hospital and are paid relatively modestly compared to their enormous prior investment in education.

If the marginal resident’s direct labor contribution isn’t sufficient to motivate the creation of an additional slot, then that slot won’t be created. As a result, the number of aspiring physicians who are capable of entering and subsequently completing a residency is ultimately determined by the financial costs and benefits of such programs at individual teaching hospitals. And despite the earnings premium on attaining physician licensure being so great that many would take out a student loan in order to obtain a residency placement, this opportunity to internalize the educational costs of residency training in order to eventually reap the benefits does not presently exist – nothing analogous to student loans exists for financing one’s residency training. Instead, if the costs of an additional residency slot exceed the benefits, hospitals will not add the residency unless subsidized by third-party payers to do so, such as through Medicare.

At least a year of residency training is required for physician licensure in all 50 states and in the District of Columbia. If newly minted MDs cannot get that training, they are effectively locked out of the profession they have spent years preparing for. Data released by The Match illustrate this bottleneck and how it has tightened over time.55 The figures below attempt to accurately capture the outlook across time for an American medical school graduate in terms of the graduate entering the profession, adjusting for both the changing composition of medical graduates as well as isolating the outcomes for U.S. citizens. Correcting for this involves grouping allopathic and osteopathic graduates as well as creating a separate measure that also includes Americans who attend medical school abroad. A more detailed description of these choices is available in Appendix 1.

When looked at in terms of raw numbers, we can see that thousands of U.S. medical school graduates remain unmatched each year. Yet even properly categorizing residency applicants to account for the changing composition of medical school degrees still results in a picture of the residency pipeline that is likely too rosy. The Match only includes “active applicants,” a metric that omits certain applicants who either dropped out or didn’t get a single interview offer. Most of the applicants who dropped out likely received a match through minor matching services, such as the independent and highly...

55 Also known as the National Resident Matching Program. See the appendix for an explanation of the author’s calculations.
Figure 7: Thousands of U.S. medical school graduates go unmatched each year.

Source: Author’s calculations using data from the National Resident Matching Program.
The residency bottleneck for entry into the physician profession results in an enormous waste of persons and talent. Though unmatched applicants can reapply, residency applications are expensive and those applying after their senior year of medical school face tough odds to succeed. Those medical graduates who fail to obtain a residency placement generally find they lack a smooth path toward a career that will allow them to apply their clinical skill in a manner commensurate with their training and education. Despite four years of postgraduate education, including two years in clinical settings, unmatched medical graduates cannot simply become a physician assistant (PA) or nurse practitioner (NP) without first going back to school. Yet that's exactly what many unmatched medical graduates do, obtaining three more years of education, including one year of clinical experience, to become an NP or PA. Others pursue certifications in various technician roles or seek employment as liaisons or analysts for the medical device industry or pharmaceutical industry.

The status quo results in tragedy on a personal level as well. Even putting aside the cost of pursuing additional training, failure to ultimately obtain a physician's income after slogging through all the intermediate steps up to residency placement can be financially crippling. The median medical school debt upon graduation is over $200,000, according to the Association of American Medical Colleges, excluding any debt accumulated in pursuit of the preceding bachelor's degree. And that's on top of what generally amounts to tens of thousands of dollars spent on licensing exams, residency application fees, traveling for residency interviews, and other physician pipeline-related expenses. Many unmatched graduates end up with burdensome monthly interest payments and face a high risk of default.

This state of affairs is not normal. Neighboring Canada is instructive here, owing to the similarity of its physician training pipeline. It utilizes both its own matching system for residencies and a dual undergraduate-plus-medical-degree educational track for entry into the residency system. Canada also followed the U.S. lead in pursuing policies aimed at reducing the physician supply, such as freezing medical school enrollment. Yet when Canada reversed course, once again pursuing expansion of its physician workforce, the residency system did not prove to be the same bottleneck that it has in the United States. The match rate for Canadian graduates is 97 percent, surpassing even

56 Bryan Carmody, “What’s the Real Match Rate,” The Sheriff of Sodium, December 5, 2019.
61 Benjamin T.B. Chan, From Perceived Surplus to Perceived Shortage: What Happened to Canada’s Physician Workforce in the 1990s?, Canadian Institute for Health Information (Ottawa: Canadian Institute for Health Information, June 2002).
the 93 percent success rate for the elite subset of U.S. allopathic graduates in The Match.62

Looking beyond North America, the strangeness of the U.S. practice of locking thousands of medical graduates out of the profession each year is even more apparent. Access to some form of postgraduate training for medical school graduates essentially operates as something close to an entitlement throughout most of Europe. So long as an individual is able to pass their respective country’s licensing exam, they can be confident that their investment in medical school won’t be wasted.

Our broken, convoluted system of GME financing

The primary reason why the American residency system is so dysfunctional is that it is tilted against medical graduates. In contrast to when the residency system was in its infancy and hospitals competed for medical graduates, today, medical graduates compete for hospitals. The implementation of increasingly stringent program requirements combined with tightening public funding have made it the case that the training of residents has increasingly become a net cost to hospitals. And, as discussed in the earlier history section, this reversal in bargaining leverage is a product of intentional design, an attempt to dissuade Americans from becoming physicians out of the mistaken fear of an impending “physician surplus.”

The primary goal for reforming the residency system should be to increase the demand for medical graduates among teaching hospitals. The diminished pay experienced by physicians in teaching institutions relative to those in private or hospital settings helps to illustrate the perverse dynamics currently facing the physician training pipeline. A 2017 survey found the average annual salary cut experienced by physicians in academic settings to be worth $123,000 relative to their counterparts in non-academic settings.63 This gap would be even larger if the fact that academic facilities are disproportionately located in higher-cost areas in the American North East were accounted for. After decades of failing to adequately support physician training, their services are now in such high demand that expanding one’s precious time on teaching doesn’t make financial sense for most physicians. The stark imbalance in remuneration between medical services and physician training is impacting not only residency programs but medical schools as well. And a secondary consequence of this remuneration gap that’s worth noting is the financial disincentives it imposes on physicians who would otherwise be interested in pursuing medical research. This reliance of our physician training and education pipeline upon an altruistic “calling” has created a gap that will require additional spending to close.

Still, attributing workforce bottlenecks in the residency system solely to insufficient funding levels fails to take seriously the structural deficiencies in the current financing arrangement. Indeed, despite cutbacks to federal funding enacted in the 1980s and 1990s, many residency programs remain


63 Medical Group Management Association, “MGMA Data Finds Non-Academic Hospital System Physicians Earn As Much As $123,000 More Than Their Academic Counterparts,” CISION PR Newswire, November 29, 2017.
well-funded, with the subsidies for some bordering on lavish. In 2018, the public subsidies for GME totaled approximately $21 billion across all federal programs and Medicaid. Spending for that year worked out at almost $150,000 for each of the approximately 140,000 residents and fellows. The current overall amount of public spending likely isn’t too far short of what would be needed to put the residency system on a substantially better footing — however, the current organization of GME financing is just too poorly structured for doing so effectively.

Over two-thirds of GME subsidies are delivered through Medicare. However, this funding stream has substantial problems. For instance, hospital subsidies are scaled based upon a hospital’s Medicare patient load. This design choice is a product of the false assumption at the time that all insurers would contribute their fair share of funding. As a result, the overwhelming alignment of GME funding with the elderly Medicare population creates funding biases, not only in geographic terms, but also in terms of the specialty mix. And yet a more appropriate alignment of funds around a nationally unrepresentative patient load is merely the tip of the iceberg.

The most straightforward source of federal support for GME is Medicare’s Direct Graduate Medical Education (DME). It’s a total mess though, based largely upon fossilized funding formulas and levels derived from Medicare’s cost-based reimbursement system from roughly 40 years ago. For most facilities, the reimbursement per resident was carried over from Medicare’s cost-based reimbursement system when it was retired. Due to changes in the relative cost of living across

64 See Appendix 2 for a detailed breakdown.
the country over the last 40 years and the fact that the originally established values were often subject to limited oversight, the present subsidies bear only a hazy resemblance to the costs a given facility faces today. Reimbursement based on the reported costs is sort of a silly exercise because the expenditures reported on paper can easily bear no resemblance to the actual economic incidence of GME funding support being described. And this is particularly true when teaching institutions face a minimal threat of having their prospective medical residents (and associated federal funding) migrate elsewhere, as is often the case due to the cap on Medicare-funded slots.

The largest channel through which Medicare funds residencies is actually indirect medical education (IME). IME works by magnifying reimbursements to hospitals made through Medicare’s Prospective Payment System (PPS), which operates on predetermined prices based on a patient’s medical diagnosis. Within IME, there are two separate funding streams: operating and capital. These are based upon the ratio of residents to hospital beds and to the average daily number of patients, respectively — a difference that’s largely an artifact of the staggered transition from cost-based reimbursement to the PPS in the 1980s, with the designers deciding that an adjustment based upon patients was actually superior by the time the capital PPS was rolled out. Indeed, the IME operating adjustment, which uses the ratio between residents and beds, is essentially no longer functioning as originally intended due to a hospital-specific cap on IME payments instituted in 1997 as part of the Balanced Budget Act of that year. Hospitals that reduce beds or add interns no longer gain any additional payment, thus rendering the payment adjustment that hospitals receive today fairly arbitrary.

---

The caps imposed in 1997 largely froze in place the facilities eligible to receive Medicare GME funding. However, the resulting reimbursement disparities are far beyond what can be defended based on inherent or necessary cost differentials between teaching hospitals. Medicare will add 1,000 funded slots, which will be distributed over five years starting in 2023 under the 2021 Consolidated Appropriations Act.67 Even putting aside the direct limitations on funding imposed by the cap, the existing system does a poor job of facilitating the creation of new residency programs for other reasons. Funded slots through Medicare are sometimes made available through reallocation as facilities close down.68 Establishing a residency program involves significant fixed costs and there's

---


68 Though nearby hospitals are prioritized when this occurs.
also a significant risk involved in inadvertently establishing an insufficiently generous per-resident rate of reimbursement. This is because, perversely, Medicare reimbursement rates for residency programs, in almost all cases, are set over an infinite time horizon with no possibility of revision.

Currently, medical residents are trained disproportionately in low-population-growth, high-cost states and metropolitan areas in the Northeast. These residents are, in turn, expected to disperse throughout the rest of the country. Considering investments in the residency system from the standpoint of cost-effectiveness, it makes more sense to do the exact opposite. Owing to differences in the cost of living, a dollar spent to train a medical resident in Georgia goes much further than in New York state, yet New York trains more than 3 times as many medical residents per person as Georgia.

Another issue with the current allocation is simply the mismatch between residents and the demand for additional medical services. This is because medical residents implicitly finance a portion of their

Figure 11: Medical residencies are heavily concentrated in the Northeast.

Source: ACGME.
training through their labor contribution. Yet the opportunities to do this are scarcer in places where residents are not only substantially more prevalent but, as is the case in the Northeast, physicians are as well. The present mismatch is evident in Medicare GME reporting data. Among residency programs with at least 100 FTE residents per year, a meaningful relationship exists between metro-area population growth and institutions opting to train above their assigned Medicare GME caps.69

To the extent private insurers are willing and able to pass along higher premiums to the local population, it’s not at all clear why they should shoulder the burden of financing physician training where such physicians are not needed. Even the “winners” under our broken system of residency financing don’t get a great deal at the end of the day. Almost everyone loses when we train too many doctors in the places where it’s most costly to do so. It is better and more cost-effective to train residents in places where there is more demand for medical services relative to the existing supply. Current policy is not accommodative of this fact.

**Additional barriers to primary care residents**

The United States’ dearth of primary care physicians is also encouraged by the current residency financing system. Because the IME formulas were originally designed to reimburse care at hospitals rather than outpatient settings, the Affordable Care Act created affiliation agreements between the two sorts of facilities so as to reduce IME’s bias against specialties more likely to practice in outpatient settings, notably those in primary care. Though an improvement over the situation it preceded, the need for such agreements is still far from ideal. The current arrangement, in which hospitals effectively loan residents to other facilities, still forces providers to enter into needlessly complicated arrangements in order to pass audits and maximize GME subsidies.70 Furthermore, regardless of the facility type, the primary care residents trained typically produce less revenue for sponsoring institutions as a consequence of the sorts of procedures that they commonly perform, resulting in less generous subsidies for such residents.71 As Bruce Steinwald and co-authors write for the Brookings Institution:72

> The IME add-on for a relatively less expensive case, such as pneumonia, might add about $30 to the diagnosis-related group (DRG) payment, while the add-on for a more complex case, such as bypass surgery, might add ten times as much to the DRG payment.

Making up roughly 30 percent of public funding, Medicaid is the second-largest funding source for postgraduate physician training. Like the rest of the Medicaid program, this financing is provided jointly by federal and state governments. The share of GME funds coming through Medicaid has

---

69 Author’s analysis of Medicare’s HOSP-10 GME Reporting File.
72 Bruce Steinwald, Paul Ginsburg, Caitlin, Brandt, Sobin Lee, and Kavita Patel, Medicare Graduate Medical Education Funding is Not Addressing the Primary Care Shortage: We Need a Radically Different Approach, USC-Brookings Schaeffer Initiative for Health Policy (University of Southern California, Brookings Institution: December 2018).
increased considerably since the 1997 Medicare cap was instituted. As of 2018, 42 states and the District of Columbia operated Medicaid GME funding programs. State-level funding GME funding largely mimics the structure of federal funding streams, distributing funds in a manner that is roughly similar to Medicare’s DME or IME. This design choice is largely a product of states’ need to comply with standards for fiscal matching from the federal government. It also reflects the implicit and false assumption, also embedded in the design of Medicare GME, that private insurers specifically earmark funds for the purpose of training medical residents. The requirement that Medicaid GME funding be delivered as a function of reimbursements for services by the insurer’s enrollees constrains states’ ability to efficiently target workforce gaps in a strategic manner.

The remaining federal GME funding streams attempt to fill gaps created by the Medicare program. For instance, because Veterans Affairs facilities and children’s hospitals receive few or no Medicare patients, specific programs were carved out to provide residencies at facilities serving these populations. Also, the Teaching Health Center Program was established to help train primary care practitioners at smaller-scale clinics that disproportionately service the medically underserved. In addition, the Department of Defense provides funding for a small number of residencies related to its own objectives.

Finally, it’s worth noting that while private insurers do not contribute to the financing of GME in any transparent or consistent manner, that doesn’t mean that some residency programs do not benefit indirectly from such funding. In particular, large teaching hospitals are more capable of extracting higher reimbursement from private insurers. This pattern of elevated reimbursement is attributable to a hospital’s reputation or the extent to which it has cornered the local hospital market, thus increasing their capacity to raise premiums on local residents since cutting the facility out of the insurer’s network isn’t a realistic option. This is a horrible way to finance GME because it does little to ensure residents are being produced in the communities most capable of supporting them or most in need of them. Furthermore, funds resulting from elevated reimbursement attributable to market power are treated as general funds rather than earmarked for the specific purpose of training residents. Consequently, the status quo here simply doesn’t work.

74 Department of Health and Human Services Centers for Medicare & Medicaid Services, “Medicaid Program; Graduate Medical Education Proposed Rule”.
75 Another non-public funding source for residencies worth noting is corporate donations. Many teaching hospitals, particularly larger ones, also receive some funding from medical equipment companies. Such companies donated $832 million in 2018, accounting for 91% of all non-research gifts to teaching hospitals. These companies obviously have a vested interest in influencing the treatment decisions of new physicians. Timothy Anderson, Walid Gellad, and Chester Good, “Characteristics of Medical Industry Payments To Teaching Hospitals,” Health Affairs Vol. 39 No. 9 (September 2020): Page 1583-1591.
Policy recommendations

The last half-century of policy aimed at postgraduate residency training, which facilitates physicians’ entry into professional practice, has been hamstrung by its fractious design and needless complexity. A better, more well-designed GME financing system would encourage hospitals to demand more residents or new residency programs. It would also encourage the training of physicians in places where they are most needed while generating incentives for cost-effectiveness and better programs.

Failure to rationally coordinate GME funding sooner rather than later will only compound the difficulties. A solution should involve:

1. Consolidating and modernizing existing GME financing;
2. Augmenting the federal coordination of capital investments in the health sector.

The following policy recommendations aim not only to clean up and modernize the existing system but to create a stable yet flexible framework that will continue to ensure the physician workforce we need long into the future.

Consolidate and standardize federal GME payments

The current system of federal GME financing is an arbitrary patchwork that needs to be phased out and rebuilt from the ground up. The training of medical residents is a public good with benefits that extend far beyond the training institution. Yet funding is currently structured under the fictitious pretense that the various public and private actors in a given medical market are each contributing their fair share to funding this good. And while some have advocated an all-payer financing pool wherein private and public insurers would be forced to contribute their fair share, the United States’ health financing system is likely too decentralized and uncoordinated for arrangements of this sort to function smoothly.77

A better approach would be for the federal government to distribute residency funding on a relatively uniform per-resident basis, severing the link between residency funding and service reimbursement entirely. Severing this link would constitute a major administrative simplification both for hospitals and the federal government, requiring substantially less effort in terms of paperwork and auditing. Reported-cost reimbursement is ultimately a poor means of ensuring cost control and produces a lot of arbitrary funding disparities – therefore, we should stop attempting to figure out the costs at specific institutions. Instead, the federal government should provide an amount that should be sufficient at most teaching institutions and then rely upon competition among teaching facilities for residents, combined with the quality standards imposed by accreditation, to act as the primary disciplining mechanism for funding within the residency system.

Unlike the current system, this payment should be blind with respect to the type of insurer or the diagnoses of patients at the training facility. This more stable and predictable approach to funding

residencies has many virtues. For instance, there’s no good reason why events such as a natural disaster should threaten to upend residency funding. This scenario is particularly relevant in the case of funding through Medicare’s IME operating adjustment, which becomes less generous as a teaching institution increases the number of hospital beds.

This more uniform approach has the added benefit of being much more accommodative of different training models, since they’d have no intrinsic connection with the inpatient facilities. While the vast majority of residents are trained in inpatient hospitals, the vast majority of care these days takes place in outpatient settings. Making funding blind with regard to the care setting would promote greater experimentation in settings that are more likely to be relevant to a physician’s future medical practice.

**Reorganize GME funding**

The consolidation of GME financing should involve the merger of the present three Medicare GME funding streams (DME, IME-Capital, and IME-Operating) along with the Teaching Health Center GME program, the Children’s Hospital GME program, and the Veterans Affairs GME program. Organizing this consolidated GME program within Medicare, as recommended by a 2014 report by the National Academies of Medicine, makes sense. As a non-discretionary funding stream, Medicare provides the degree of certainty necessary for institutions to plan over the medium term without having to worry about year-to-year fluctuations due to budgetary wrangling. Certainty of this sort would be highly valuable given the extended duration of many U.S. residencies. The exception to federal consolidation should be Department of Defense GME funding, which is comparatively tiny as well as intended for the distinct goal of fostering training in areas of strategic interest to the military.

The reorganization of direct federal GME financing (Figure 10) should therefore be paired with the liberalization of reimbursements for GME through Medicaid, allowing contributions to be similarly de-linked in order to foster a more strategic allocation of residency positions, which is discussed in further detail later.

**Abolish institution-level FTE caps on residency support**

The prevailing caps on Medicare GME financing represent an ill-conceived policy on many levels. The primary adverse impact of the caps is to freeze the geographic distribution of physicians, with the resulting distribution of residency training being highly discordant with current population patterns.

---

78 The National Academies of Sciences, Engineering, and Medicine, “Governance and Financing of Graduate Medical Education,” 2014.

79 Federal GME spending is currently adjusted upward based on the broader consumer price index, but this is a mistake. Since it is physicians that instruct and oversee residency training, the relevant bundle of prices to track from the perspective of residency financing is the opportunity of cost time spent teaching that could otherwise be spent performing medical services. Annual inflation adjustments should therefore be tethered to growth in the cost of broader hospital and related services. U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: Hospital and Related Services in U.S. City Average, Federal Reserve Economic Data (FRED), CUUR0000SEMD, (St. Louis: Federal Reserve Bank of St. Louis, August 25, 2021)
New programs are made more difficult to start up not just by the funding limits imposed by the caps themselves, but also by the administrative burdens imposed in order to maximize funding while maintaining the caps. For instance, outpatient facilities must enter into agreements when receiving resident rotations from inpatient facilities in order to not forfeit any funding while remaining under the total resident cap.80

Yet there are also less obvious reasons why the cap is harmful. Without meaningful competition among teaching institutions for medical graduates, teaching institutions are encouraged to treat federal subsidies for residencies as general funds — free to be spent on anything — rather than resources that should first and foremost be spent on the residency program itself. So long as residency opportunities are scarce from the perspective of medical graduates, federal assistance runs the risk of being effectively diverted to supporting more lucrative medical activities at the expense of the program. This tendency is a result of the fact that the benefits of residency programs accrue largely to the residents themselves in the form of higher future wages. Eliminating the cap helps address this imbalance — allowing institutions that more adequately invest in their programs to grow.

A key goal in reforming federal GME support should be to make access to funding for residency training something close to an entitlement for medical school graduates that have passed relevant licensing examinations. Federal funding should be available to all residents who are U.S. citizens and lawful permanent residents.81 States would, of course, still be free to provide GME funding for non-LPRs through their Medicaid programs.82

---

80 Department of Health and Human Services Centers for Medicare & Medicaid Services, "Direct Graduate Medical Education (DGME)," December 22, 2020.

81 It would be prudent to place restrictions on medical graduates that aren’t either U.S. citizens or lawful permanent residents on the grounds that it places reasonable bounds on anticipated cost-growth and preserves the political viability of an entitlement-style approach to financing residency slots.

82 Federal GME funding for non-residents should be made contingent upon alignment with the Conrad 30 program, which
Vary federal payments to teaching institutions based on program size

Previous analyses have identified program scale as an important factor in determining the financial burden imposed on teaching hospitals by residency programs. Teaching hospitals with more residents are better able to spread out the fixed costs across a larger number of per-resident reimbursement as well as to economize on shared facilities.

Yet the current advantage experienced by institutions with larger quantities of resident runs has likely contributed to the geographic over-centralization of resident training at legacy institutions. For this reason, the marginal per-resident payments to teaching facilities should be made to vary accordingly, declining as the number of total medical residents increases. Other countries that finance GME on a per-resident basis, such as the Netherlands, perform a scale adjustment to their system’s GME payments similar to what is being described here.

Three sorts of facilities may be worth excluding from the program-scale adjustment:

1. Academic medical centers that are organizationally and administratively integrated with medical schools. These institutions are typically larger and maintain a substantial non-market research focus in addition to providing services to communities — these might need a bit more funding relative to their economies of scale to reflect these institutions’ broader, and often less lucrative, mission;

2. Teaching health centers. These are smaller non-hospital teaching settings in areas designated as underserved — these might warrant an additional boost as well.

3. Veterans Affairs (VA) facilities. The danger of the VA overinvesting in its residency program is minimal. Yet VA programs should still retain access to non-discretionary financing in amounts comparable to and therefore competitive with non-VA facilities.

Increase payments based on the share of PG-1 residents

The first year of postgraduate residency training (PG-1), sometimes referred to as the internship year, should receive a boosted payment rate (see Appendix 3 for hypothetical funding schedules).
The value that first-year residents generate for teaching institutions is lower, and then gradually increases as they advance through the program. A boosted year-1 payment can help new institutions looking to start up a new residency program, in which a higher per-resident payment can help defray the sizeable fixed costs as they ramp up to scale.

Furthermore, boosting the payment rate for first-year residents implicitly subsidizes the generally shorter residencies undertaken by physicians in specialties most likely to enter into primary care, such as family medicine. This makes sense as the broader social benefits of practicing primary care are likely in excess of the market compensation relative to other sorts of physicians. A family doctor is compensated for the care avoided due to catching symptoms when they are manageable, encouraging vaccination, or building trust between persons and the broader medical system in the same way that a heart surgeon is compensated for his discreet procedures. Also, boosting payments during the first year would act as an implicit incentive to improve the volume of physicians trained, thereby increasing the productivity of the residency system.

There are two ways to achieve such a boost. First, a flat increase in the payment based on the number of first-year residents; for instance, $10,000 per PG-1 resident. Another way to go about doing this would be to adjust the total teaching institution payments upward at a set rate based upon the institution’s PG-1 share of total residents. For instance, a hospital receiving $2 million in annual federal support for residency training might see a 20% upward funding adjustment based upon having 25% of residents in their first year, thus raising their total federal payment to $2,100,000 for that year, or a 10% increase. The differences between these two approaches are actually rather small, with the latter approach being slightly more favorable to institutions utilizing the shorter residencies that are disproportionately found among specialties geared toward primary care.

End support for training beyond the 5th year of residency

Currently, residency programs beyond their fifth year are subsidized by Medicare at one-half the standard rate, yet the case for any public funding is fairly dubious. Almost all residency programs can be completed by the physician’s fifth year, with the exception of a few specialties, such as neurosurgery and plastic surgery. As a result, the current target of this extended funding is overwhelmingly subspecialty fellowship programs.

Yet by a physician’s fifth year of residency training, they are more than capable of performing highly remunerative tasks from the standpoint of teaching facilities, thereby indirectly self-financing this additional training. And, indeed, it’s been estimated that the majority of subspecialty fellowships in fields such as radiology lie outside ACGME accreditation, meaning they receive no federal GME support whatsoever.86

When considering the value of subsidizing further subspecialization, the public should be wary of the extent to which doing so encourages “ball-hogging” by the recipients of more prestigious

credentials without producing value for the public commensurate with its investment. While the pitfalls associated with overspecialization, namely market power and workforce rigidity, have been remarked upon and accounted for by equivalent regulatory bodies in Europe, scant attention has been given to these potential issues in the United States despite its vastly more specialized physician workforce, both in terms of the prevalence and the much higher number of distinct specialties and subspecialties.87

The legality and pervasive use of “exclusive practice contracts” in the United States, wherein specific services or whole departments are handed over to certain physicians at the exclusion of other non-contracting physicians, provides a clear instance of potential ball-hogging exercised through the accumulation of market power.

“Turf wars” over medical funding are one of the defining features of America’s health care political economy, waged not just between different physician specialties but also between the various health care professions. It’s also arguably one of the most perverse features, prompting practitioners to pursue increasing amounts of credentialing in order to bargain against their colleagues.88 America’s outlier status in terms of training volumes and specialization has produced dubious benefits, and public policy has no business subsidizing zero-sum arms races. For this reason, federal funding for postresidency fellowships should be zeroed out, or at least remain capped in terms of funding if the complete withdrawal of federal support for these proves politically infeasible. Given our scarce supply of physicians, it just doesn’t make sense for the public to subsidize ever-greater levels of specialization when the benefits of doing so are overwhelmingly captured by the physicians themselves through higher remuneration.

Ensure federal GME payments are geographically uniform

A reasonably uniform payment rate would ensure that new residency programs have greater upfront certainty with regard to revenue, rather than praying that they can establish a decent enough per-resident rate in the initial years through Medicare DME to keep them operating in an adequate financial position over the long term. Payments through Medicare IME are even more difficult to estimate due to their dependence upon the diagnoses of the patient population in a given year.

Furthermore, the current allocation of GME spending across institutions is not only wildly unfair, with some facilities receiving many multiples of others, but also inefficient. A uniform payment would encourage teaching institutions to compete based upon their ability to direct federal support toward the program itself in a cost-effective manner. A facility in New York state that requires three times as much funding in order to offer a residency experience comparable to one offered in Texas shouldn’t have federal funding simply adjusted upward in order to make it so. For comparison, Pell Grants, used to subsidize college education, are not adjusted based on local costs because the students are not

87 General Medical Council Intelligence Unit Research; Specialties, sub-specialties and progression through training: the international perspective; (United Kingdom: General Medical Council, August 2011).
tied to the location that they are educated in after they graduate.

This same logic should apply to postgraduate physician training. We shouldn’t expect greater geographic parity to result in anything more than a modest rebalancing away from higher-cost training areas. Prestigious institutions in higher-cost areas will be more than capable of leveraging their name brand to attract sufficient trainees. More significantly though, these higher-cost training areas largely overlap with richer, high-fiscal-capacity states, where leveraging additional federally matched funds through Medicaid shouldn’t be an issue, and may not even be necessary.

**Give states flexibility to target Medicaid GME payments**

Current Medicaid policy allows states to receive federal reimbursement for GME payments made as an add-on to the program’s inpatient and outpatient reimbursements. This provides states with little flexibility to target funds strategically, effectively forcing them to finance GME through an IME/DME setup. To illustrate the current difficulties here, Massachusetts dropped its Medicaid financing of GME upon finding that the current rules on federal reimbursement rendered target investments in primary care and psychiatric training too restrictive to be allocated effectively. A better way to go about this would be to allow states to be directly reimbursed for their GME expenditures at the Medicaid rate.

This would involve replacing the current Medicaid guidance limiting federally matched GME financing to simply a system with Medicaid service add-on payments.\(^{89}\) However, these payments would be actually less transparent than direct appropriation since a teaching facility’s actual payment depends upon their specific Medicaid service load. In this case, Congress could merely make receipt of Medicaid GME conditional upon the performance of any Medicaid-reimbursed services. This would allow states to directly address physician workforce gaps, both in terms of geographical location as well as specialty mix. States would be able to dip into these federally matched funds should the uniform federal payment prove insufficient in higher-cost areas.

In order to ensure accountability, a fix to GME funding through Medicaid should also incorporate provisions to simply keep track of how much states are spending on GME, how much each hospital is receiving, and how many medical residents are being supported.

**Plan for an organized transition to the new system**

Transitioning here should be done with care, but it is imminently feasible. Particularly during the earlier years, policymakers shouldn’t be averse to spending more money. The long-term gains from a better functioning system would be significant and meaningful, both for the productivity of the residency system itself as well as for the functioning of the health system more broadly.

Paired with reforms laid out in this report, the downsides to over-investing in the residency system are relatively minimal. Since federal payments would be distributed uniformly based upon facilities’

---

89 Department of Health and Human Services Centers for Medicare & Medicaid Services, “Medicaid Program; Graduate Medical Education Proposed Rule.”
ability to compete for and attract potential medical residents, the worst that could happen is that residency programs might be made more attractive, such as through higher pay for medical residents, more competitive remuneration for academic physicians, or the augmenting of the facilities and relevant research functions at the teaching institutions. Indeed, one could even argue that squeezing individuals’ wages until they are in their mid-to-late-thirties is a toxic dynamic, particularly for those looking to start families. And we should want teaching physicians to have reasonably competitive wages with those of non-teaching ones.

Furthermore, states are more than capable of scaling their Medicaid contributions should funding prove excessive in some portions as the payment rates are being ironed out. The far greater danger is that reformers instead continue to tinker around the edges of our currently broken financing system or instead propose reforms that are too stingy to satisfy existing stakeholders. And if the overall base rate of payment were to somehow still end up as excessive, Congress would be more than capable of slowing or temporarily freezing the rate of inflation adjustments here.

On the most basic level though, the United States has nothing to fear from even substantially more physicians. Most other developed countries utilize far more physicians than we do, resulting in better and more accessible health care services. Even with the reforms laid out above, we’d have a tough time reaching physician numbers comparable to what is common throughout Europe. For that reason, it’s worth exploring additional policy measures, outside the strict confines of the residency system itself, aimed at better allocating and augmenting the provision of medical care.

Establish a health care resources development bank

A development bank for health care – capable of coordinating strategic grants and guaranteed loans – is sorely needed. Greater facilitation of capital investment in the health care sector would improve the performance of the U.S. healthcare system. Other developed countries play a much more active role than the U.S. in facilitating such investments.90

U.S. health care faces unusual difficulties in terms of accessing and allocating capital in a manner that is consistent with maximizing public health. The return on hospital capital investment has been found to be low relative to the rest of the health sector.91 Restrictions on the corporate practice of medicine make private capital investments even more difficult.92 And public and nonprofit hospitals face the additional difficulty of being unable to tap equity markets.93 The deeper issue, however, is that the health care investments generating the highest returns are often not the ones that produce

the greatest gains in terms of health outcomes. A federal health system development bank capable of providing financial support to qualified health infrastructure projects would be of great use here, investing in sustainable projects with the greatest potential to improve access and health outcomes. This support would take the form of grants, direct loans, and loan guarantees to state and local governments that issue debt for qualified projects.

This financial support would be of great use to the physician training pipeline as well. For instance, federally coordinated investment would be of great use in the residency pipeline, which is reliant upon a highly uneven distribution of teaching hospitals throughout the country. Capital secured through this means could be used to construct new hospitals or outpatient facilities and could assist with capital expenditures necessary to prepare new facilities to host a residency program.

This bank’s mission should potentially also include providing support for the establishment of additional medical schools. Similar to residency programs, medical schools are currently over-concentrated in the Northeast. The amelioration of bottlenecks within the residency system will hopefully lead to an increase in demand among Americans to enroll in medical school. Here it may be worth investing in, for instance, programs that offer 6-year medical degrees without needing to first obtain a bachelor’s degree, similar to the medical education format practiced basically everywhere outside of North America.\(^\text{94}\) This approach is much cheaper for graduates, saddling them with less debt and allowing them to start their careers at an earlier age. And the more integrated format enables institutions to offer a superior educational experience.

In terms of structuring, the Health Resources and Services Administration (HRSA) already does a small amount of grant-making related to residency startup costs, so simply augmenting HRSA may make sense; though the mission may be sufficiently expansive to necessitate the creation of a totally separate dedicated body.

**Conclusion**

Physicians are the cornerstone of the U.S. health care system. Yet we have too few of them. This is a product of decades of misguided assumptions about the threat of a “physician surplus,” combined with overly complex and antiquated institutional arrangements. And while the misguided assumptions have largely been shed, those Byzantine institutions have failed to deliver on the new consensus that we need far more doctors. One key reason for this failure is that they produce a shortage of residency slots for holders of medical degrees.

The residency bottleneck in the U.S. excludes thousands of potential doctors from working in the profession each year, but also dissuades capable individuals from pursuing the profession as well as medical schools from increasing enrollment. This problem is eminently within the powers of the federal government to fix, as the primary sources of funding for medical residencies are the Medicare and Medicaid programs. The reforms recommended here can help put the system on a stable path going forward.

Appendix 1: Using NRMP data

In order to properly analyze the National Resident Matching Program (NRMP)'s data, it is important to keep a number of things in mind. First, only the outcomes of U.S. citizens should be considered when evaluating the extent to which a bottleneck exists. This is because the supply of international medical school graduates is functionally unbounded and has ebbed and flowed considerably over time. Nonetheless, a U.S. residency system capable of training doctors at a scale commensurate with demand from foreign medical graduates would certainly amount to an enormous benefit to the American people.

Second, the composition of medical graduates among U.S. citizens isn’t what it was in the 1980s. The numbers of osteopathic physicians (DO) and graduates of international medical schools who are U.S. citizens (US-IMGs) have increased precipitously since the 1980s when MD-granting allopathic schools instituted a 25-year moratorium. Osteopathic medical schools didn’t take part in the moratorium, instead growing from a relatively small sum to now constituting a relatively sizable share of overall U.S. physicians. More limited medical slots also drew many aspiring U.S. physicians to overseas medical schools, particularly in the Caribbean. These days, it’s helpful to think of the U.S. physician workforce as operating under a 3-tier prestige hierarchy, with allopathic MD-degree physicians on top, followed by the osteopathic DO-degree physicians, and international medical school graduates at the bottom. As a result, looking at The Match outcomes for MDs in isolation effectively cherry-picks an increasingly elite cohort of medical graduates.

Third, the extent to which The Match represents a comprehensive portrait of residency outcomes has improved owing to the institution of the NRMP’s “All-In” policy in 2012 and the phasing out of the parallel matching process for DOs. Under the NRMP’s “All-In” policy, residency providers lost the
ability to fill some positions within The Match, while pre-matching other slots. Starting in 2014, the parallel DO-match organized by the American Osteopathic Association began to be phased out, with the matching process for all American physicians now done through the NRMP’s unified system. Both of these events led to outsized short-term increases in the number of residency slots offered through the program, while also increasing the number of applicants somewhat by some amount by including those persons who would have previously landed a residency outside the NRMP. Failing to account for the compositional change in the degrees held by medical graduates produces a somewhat rosier picture than in reality.

The official Match figures contain two sources of arguable bias which this paper does not attempt to correct for. First, official NRMP statistics only include “active applicants”, or applicants that submitted a rank-order list. This leads to an under-counting of less competitive applicants, since applicants that do not receive an interview have no programs to rank. Precisely adjusting for this sort of bias is interpretively difficult, however, because some of these applicants likely dropped out voluntary, opting to instead pursue one of the much smaller alternative matching such as the independent Urology Match. A second source of arguable bias is the Supplemental Offer and Acceptance Program (SOAP), wherein unfilled residency slots are made available to applicants that did not match in the primary matching process. These sources of bias point in opposite directions. Yet because the number of applicants that eventually obtain residencies through SOAP is credibly much smaller than the number of applicants that failed to land an interview, it was deemed that adjusting for SOAP without attempting to account for the other form of bias would illustrate a misleadingly rosy picture of match probability.

**Figure A2: Residency match probability by applicant type**

![Figure A2: Residency match probability by applicant type](image)
## Appendix 2: GME funding sources

**Figure A3: Major sources of GME Funding: 1998 and 2018**

<table>
<thead>
<tr>
<th>Program</th>
<th>2018</th>
<th>1998 (Inflation Adjusted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicare</td>
<td>$13 Billion</td>
<td>$10.9 Billion</td>
</tr>
<tr>
<td>DME</td>
<td>$3.7 Billion</td>
<td>-</td>
</tr>
<tr>
<td>IME - Inpatient Operating</td>
<td>$8.9 Billion</td>
<td>-</td>
</tr>
<tr>
<td>IME - Inpatient Capital</td>
<td>$0.4 Billion</td>
<td>-</td>
</tr>
<tr>
<td>Medicaid</td>
<td>$5.6 Billion</td>
<td>$3.7 Billion</td>
</tr>
<tr>
<td>State</td>
<td>$2.3 Billion</td>
<td>-</td>
</tr>
<tr>
<td>Federal</td>
<td>$3.3 Billion</td>
<td>-</td>
</tr>
<tr>
<td>Veterans' Affairs:</td>
<td>$1.92 Billion</td>
<td>$0.4 Billion</td>
</tr>
<tr>
<td>Trainees (Direct)</td>
<td>$0.986 Billion</td>
<td>-</td>
</tr>
<tr>
<td>Educational Support (Indirect)</td>
<td>$0.935 Billion</td>
<td>-</td>
</tr>
<tr>
<td>Children's Hospital GME (CHGME)</td>
<td>$0.325 Billion</td>
<td>-</td>
</tr>
<tr>
<td>Teaching Health Centers GME</td>
<td>$0.126 Billion</td>
<td>-</td>
</tr>
<tr>
<td>Other</td>
<td>$0.02 Billion</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>$21 Billion</td>
<td>$15 Billion</td>
</tr>
</tbody>
</table>

---


96 Tim M. Henderson, "Medicaid's Role In Financing Graduate Medical Education."

97 Approximate, based on typical Medicaid state-federal cost-sharing.

98 Approximate, based on typical Medicaid state-federal cost-sharing.

99 Figures collected through email correspondence with Veterans Health Administration.


102 The U.S Department of Defense also contributes a small amount of funding as well, totaling 16.5 million as of the most recent figures released in FY2012. See page 2: [https://www.ama-assn.org/system/files/2020-08/2020-gme-compendium-report.pdf](https://www.ama-assn.org/system/files/2020-08/2020-gme-compendium-report.pdf)
## Appendix 3: Hypothetical funding schedules

**Figure A4:** Hypothetical funding schedules for residency and GME programs

<table>
<thead>
<tr>
<th>Stage of Training</th>
<th>Academic Medical Centers</th>
<th>Federally Qualified Health Centers</th>
<th>Teaching Hospital</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-49 FTE</td>
<td>50-149 FTE</td>
<td>150-250 FTE</td>
</tr>
<tr>
<td>PG1 Resident</td>
<td>$155,000</td>
<td>$165,000</td>
<td>$165,000</td>
</tr>
<tr>
<td>PG2-3 Resident</td>
<td>$125,000</td>
<td>$130,000</td>
<td>$130,000</td>
</tr>
<tr>
<td>PG4-5 Residency</td>
<td>$100,000</td>
<td>$100,000</td>
<td>$110,000</td>
</tr>
</tbody>
</table>
The Niskanen Center is a nonpartisan 501(c)(3) think tank that works to promote an open society.

Niskanen Center
820 1st Street NE, Suite 675
Washington, D.C. 20002
NiskanenCenter.org