Measures and Methodology for International Comparisons of Health Care System Performance

Final Report

Laurie T. Martin, Carl Berdahl, Rachel M. Burns, Emily Hoch,
Evan D. Peet, Peter S. Hussey
Preface

Compared with other Organisation for Economic Co-operation and Development (OECD) member countries, the United States ranks lower on several notable health-related outcomes, such as infant mortality, that tend to be overemphasized by stakeholders and used as a marker of health care system performance. However, the driving factors contributing to many of these rankings could be attributed to factors outside of the health care system, such as greater inequality, less social protection, population density, and societal preferences. Concerns have also been raised around the variability in data availability and quality within the United States and abroad, which can introduce bias or make international comparisons more difficult. Given the increasingly important question of what the United States is buying for its health expenditures and interest in understanding what factors of the health system contribute to better health outcomes, in 2020, the Immediate Office of the Secretary of the U.S. Department of Health and Human Services (HHS) sought to identify measures that could improve the ability of the United States and other countries to learn from international comparisons of health system performance specifically.

To inform the identification of measures for international comparison that could eventually be proposed to OECD, RAND Corporation researchers worked with a diverse group of 15 experts in quality measurement, clinical care, and health economics to generate and prioritize potential measure constructs that align with HHS priorities and are particularly promising for international comparisons. With few exceptions, the measure constructs identified as most promising will require additional development work to establish their operational definitions and specifications to ensure that any measure developed is valid and feasible for international comparisons of health system performance. This report is designed to lay a strong foundation for these future refinements by noting the degree of consensus among experts about the importance, scientific acceptability, feasibility, and usability of measure constructs; summarizing the strengths and limitations of the measure constructs; and providing additional context that can be useful for informing the selection of measure constructs that might ultimately be developed into measures and proposed to OECD for consideration.

This research was funded by the Office of the Assistant Secretary for Planning and Evaluation, HHS (Contract No. HHSP233201500038I), and carried out within the Payment, Cost, and Coverage Program in RAND Health Care.

RAND Health Care, a division of the RAND Corporation, promotes healthier societies by improving health care systems in the United States and other countries. We do this by providing health care decisionmakers, practitioners, and consumers with actionable, rigorous, objective evidence to support their most complex decisions. For more information, see www.rand.org/health-care, or contact
Abstract

In 2020, the Immediate Office of the Secretary of the U.S. Department of Health and Human Services (HHS) sought to identify measures that could improve the ability of the United States and other countries to learn from international comparisons of health system performance. To inform the identification of measures for international comparison that could eventually be proposed to the Organisation for Economic Co-operation and Development (OECD), RAND Corporation researchers worked with a diverse group of 15 experts in quality measurement, clinical care, and health economics to generate and prioritize potential measure constructs that align with HHS priorities and are particularly promising for international comparisons. The authors implemented a modified Delphi panel approach using RAND’s ExpertLens platform. This process occurred in three phases. Phase 1 involved the identification of measure constructs that should be considered for inclusion in the expert rating process. Phase 2 was the initial assessment of each measure construct by experts against prespecified rating criteria. Phase 3 was a virtual convening of experts to discuss the ratings, explore areas of disagreement, and allow experts to adjust ratings, if desired. Eight measure constructs were identified as having the most promise for international comparison, but they will require additional development work to establish their operational definitions and specifications to ensure that any measure developed is valid and feasible for international comparisons of health system performance. This report is designed to lay a strong foundation for these future refinements by noting the degree of consensus among experts about the importance, scientific acceptability, perceived feasibility, and usability of measure constructs; summarizing the strengths and limitations of the measure constructs; and providing additional context that can be useful for informing the selection of measure constructs that might ultimately be developed into measures and proposed to OECD for consideration.
# Contents

Preface ........................................................................................................................................... iii  
Abstract ............................................................................................................................................ v  
Figures ............................................................................................................................................. x  
Tables ............................................................................................................................................. xi  
Summary ...................................................................................................................................... xiii  
Acknowledgments ........................................................................................................................ xix  
Abbreviations .......................................................................................................................... xx  
1. Introduction ................................................................................................................................. 1  
   Methods ..................................................................................................................................................... 2  
      Phase 1: Identification of Measure Constructs That Should Be Considered ................................. 2  
      Phase 2: Evaluation of Each Measure Construct Against Prespecified Criteria ......................... 3  
      Phase 3: Convene, Discuss, and Adjust Ratings ........................................................................... 4  
   Organization of This Report ...................................................................................................................... 6  
2. Measure Constructs That Hold the Most Promise for International Comparisons ..................... 7  
   Treatment and Control of Hypertension ............................................................................................... 7  
      Description Provided in Evidence Table ....................................................................................... 7  
      Strengths and Limitations ............................................................................................................. 8  
      What Is Needed to Move This Measure Forward ......................................................................... 9  
   Access to and Coverage for Telehealth .............................................................................................. 9  
      Description Provided in Evidence Table ....................................................................................... 9  
      Strengths and Limitations ............................................................................................................. 9  
      What Is Needed to Move This Measure Forward ....................................................................... 11  
   Quality-Adjusted Life Expectancy ..................................................................................................... 12  
      Description Provided in Evidence Table ....................................................................................... 12  
      Strengths and Limitations ........................................................................................................... 12  
      What Is Needed to Move This Measure Forward ....................................................................... 13  
   Insurance Coverage for Mental Health, Behavioral Health, and Substance Abuse Services .......... 13  
      Description Provided in Evidence Table ....................................................................................... 13  
      Strengths and Limitations ........................................................................................................... 14  
      What Is Needed to Move This Measure Forward ....................................................................... 15  
   Receipt of Preference-Concordant End-of-Life Care ......................................................................... 15  
      Description Provided in Evidence Table ....................................................................................... 15  
      Strengths and Limitations ........................................................................................................... 15  
      What Is Needed to Move This Measure Forward ....................................................................... 16  
   Care Continuity or Consistent Provider .......................................................................................... 17  
      Description Provided in Evidence Table ....................................................................................... 17  
      Strengths and Limitations ........................................................................................................... 17
What Is Needed to Move This Measure Forward

Access to Mental Health Providers

Description Provided in Evidence Table

Strengths and Limitations

What Is Needed to Move This Measure Forward

Data Transfer and Interoperability

Strengths and Limitations

What Is Needed to Move This Measure Forward

Conclusion

3. Measure Constructs That Received Lower or Inconsistent Ratings After Expert

Discussion

Self-Reported Pain

Description Provided in Evidence Table

Strengths and Limitations

What Is Needed to Move This Measure Forward

Access to Primary Palliative Care

Description Provided in Evidence Table

Strengths and Limitations

What Is Needed to Move This Measure Forward

Prices for Brand-Name and Generic Drugs

Description Provided in Evidence Table

Strengths and Limitations

What Is Needed to Move This Measure Forward

Diffusion of and Access to New Prescription Drugs

Description Provided in Evidence Table

Strengths and Limitations

What Is Needed to Move This Measure Forward

Avoidable Emergency Department Use

Description Provided in Evidence Table

Strengths and Limitations

What Is Needed to Move This Measure Forward

The Percentage of Patients with an Opioid Use Disorder Who Were Referred to or Prescribed

Medication-Assisted Treatment

Description Provided in Evidence Table

Strengths and Limitations

What Is Needed to Move This Measure Forward

Estimates of Administrative Complexity and Cost

Description Provided in Evidence Table

Strengths and Limitations

What Is Needed to Move This Measure Forward

Disadoption of Ineffective Medical Services

Description Provided in Evidence Table
5. Recommended Modifications to Existing OECD Measures ..................................................... 51

4. Measure Constructs That Were Not Discussed by Experts ..................................................... 41
   Clinician Workforce Who Can Prescribe Medication-Assisted Treatment or Naloxone ............ 41
      Description Provided in Evidence Table ................................................................................. 41
      Strengths and Limitations ........................................................................................................ 41
   Access to Opioid Treatment Centers.......................................................................................... 42
      Description Provided in Evidence Table ................................................................................. 42
      Strengths and Limitations ........................................................................................................ 42
   The Percentage of Patients with a Follow-Up Visit Within Four Weeks of Starting an Opioid for
      Chronic Pain .......................................................................................................................... 44
      Description Provided in Evidence Table ................................................................................. 44
      Strengths and Limitations ........................................................................................................ 44
   Time to Regulatory Approval for New Prescription Drugs......................................................... 45
      Description Provided in Evidence Table ................................................................................. 45
      Strengths and Limitations ........................................................................................................ 45
   Travel Time to Provider Office..................................................................................................... 46
      Description Provided in Evidence Table ................................................................................. 46
      Strengths and Limitations ........................................................................................................ 46
   Health Care Spending in the Last Year of Life............................................................................ 48
      Description Provided in Evidence Table ................................................................................. 48
      Strengths and Limitations ........................................................................................................ 48
   Spending on Mental Health (Percentage of Total Health Spending)........................................... 49
      Description Provided in Evidence Table ................................................................................. 49
      Strengths and Limitations ........................................................................................................ 49
   Conclusion ................................................................................................................................ 50

5. Recommended Modifications to Existing OECD Measures ..................................................... 51
   OECD Health Statistics.................................................................................................................. 51
   Health Outcomes for Common Diseases ..................................................................................... 51
   Health Expenditure ....................................................................................................................... 52
   Health Workforce ......................................................................................................................... 53
   Waiting Times ............................................................................................................................... 53
   Allocation of Social Spending by Age Group ............................................................................. 54
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equity</td>
<td>54</td>
</tr>
<tr>
<td>Other Expert Recommendations</td>
<td>54</td>
</tr>
<tr>
<td>Critiques of OECD Measures</td>
<td>55</td>
</tr>
<tr>
<td><strong>6. Measure Constructs Requiring Development</strong></td>
<td>56</td>
</tr>
<tr>
<td>Important Measure Constructs Requiring Significant Development</td>
<td>56</td>
</tr>
<tr>
<td>Measure Constructs Related to Health System Preparedness, the Public Health System, and/or the COVID-19 Pandemic</td>
<td>58</td>
</tr>
<tr>
<td>Measure Constructs Not Clearly Related to Health Care System Performance or HHS Priorities</td>
<td>59</td>
</tr>
<tr>
<td>Measure Constructs for Which Experts and/or Federal Advisors Raised Questions About Directionality</td>
<td>60</td>
</tr>
<tr>
<td><strong>7. Recommendations and Conclusion</strong></td>
<td>62</td>
</tr>
<tr>
<td>Limitations</td>
<td>63</td>
</tr>
<tr>
<td>Conclusion</td>
<td>63</td>
</tr>
<tr>
<td>References</td>
<td>65</td>
</tr>
<tr>
<td>Appendix A. Evidence Tables</td>
<td>73</td>
</tr>
</tbody>
</table>
Figures

Figure 2.1. Treatment and Control of Hypertension ................................................................. 8
Figure 2.2. Access to and Coverage for Telehealth ................................................................. 10
Figure 2.3. Quality-Adjusted Life Expectancy ...................................................................... 12
Figure 2.4. Insurance Coverage for Mental Health, Behavioral Health, and Substance Abuse Services ............................................................................................................. 14
Figure 2.5. Receipt of Preference-Concordant End-of-Life Care ............................................ 16
Figure 2.6. Care Continuity or Consistent Provider ............................................................... 17
Figure 2.7. Access to Mental Health Providers .................................................................... 19
Figure 2.8. Data Transfer and Interoperability .................................................................... 21
Figure 3.1. Self-Reported Pain ............................................................................................. 24
Figure 3.2. Access to Primary Palliative Care ..................................................................... 26
Figure 3.3. Prices for Brand-Name and Generic Drugs ......................................................... 27
Figure 3.4. Diffusion of and Access to New Prescription Drugs ........................................... 29
Figure 3.5. Avoidable Emergency Department Use ............................................................... 31
Figure 3.6. The Percentage of Patients with an Opioid Use Disorder Who Were Referred to or Prescribed Medication-Assisted Treatment ......................................................................................... 32
Figure 3.7. Estimates of Administrative Complexity and Cost ........................................... 34
Figure 3.8. Disadoption of Ineffective Medical Services ..................................................... 36
Figure 3.9. Healthy Days at Home ....................................................................................... 38
Figure 3.10. Availability of Emergency Medical Services to Prevent Opioid Death ............ 39
Figure 4.1. Clinician Workforce Who Can Prescribe Medication-Assisted Treatment or Naloxone .................................................................................................................................................................. 42
Figure 4.2. Access to Opioid Treatment Centers ................................................................. 43
Figure 4.3. The Percentage of Patients with a Follow-Up Visit Within Four Weeks of Starting an Opioid for Chronic Pain ........................................................................................................... 44
Figure 4.4. Time to Regulatory Approval for New Prescription Drugs .................................. 45
Figure 4.5. Travel Time to Provider Office .......................................................................... 47
Figure 4.6. Health Care Spending in the Last Year of Life .................................................... 48
Figure 4.7. Spending on Mental Health (Percentage of Total Health Spending) .................. 49
Tables

Table S.1. Expert Reviews of Measure Constructs ................................................................. xvi
Table 1.1. Measure Constructs Included in Expert Ratings .................................................. 5
Table 6.1. Measure Constructs Requiring Significant Development ...................................... 56
Table 6.2. Measure Constructs Related to Health System Preparedness, Public Health, or COVID-19 .................................................................................................................. 58
Table 6.3. Measure Constructs Not Clearly Related to Health Care System Performance or HHS Priorities .................................................................................................................. 60
Table 6.4. Measure Constructs with Questions About Directionality .................................... 61
Table A.1. Treatment and Control of Hypertension .................................................................. 74
Table A.2. Access to and Coverage for Telehealth ................................................................. 78
Table A.3. Quality-Adjusted Life Expectancy ........................................................................ 84
Table A.4. Insurance Coverage for Mental Health, Behavioral Health, and Substance Abuse Services .......................................................................................................................... 90
Table A.5. Receipt of Preference-Concordant End-of-Life Care ............................................ 96
Table A.6. Care Continuity or Consistent Provider ............................................................... 103
Table A.7. Access to Mental Health Providers ...................................................................... 109
Table A.8. Data Transfer and Interoperability ..................................................................... 114
Table A.9. Self-Reported Pain ............................................................................................. 119
Table A.10. Access to Primary Palliative Care ..................................................................... 124
Table A.11. Prices for Brand-Name and Generic Drugs ........................................................ 129
Table A.12. Diffusion of and Access to New Prescription Drugs ........................................ 135
Table A.13. Avoidable Emergency Department Use ............................................................. 140
Table A.14. The Percentage of Patients with an Opioid Use Disorder Who Were Referred to or Prescribed Medication-Assisted Treatment ...................................................................... 144
Table A.15. Estimates of Administrative Complexity and Cost ........................................... 150
Table A.16. Disadoption of Ineffective Medical Services .................................................... 157
Table A.17. Healthy Days at Home ..................................................................................... 162
Table A.18. Availability of Emergency Medical Services to Prevent Opioid Death ............. 167
Table A.19. Clinician Workforce Who Can Prescribe Medication-Assisted Treatment or Naloxone ......................................................................................................................... 171
Table A.20. Access to Opioid Treatment Centers ................................................................. 177
Table A.21. The Percentage of Patients with a Follow-Up Visit Within Four Weeks of Starting an Opioid for Chronic Pain ....................................................................................... 181
Table A.22. Time to Regulatory Approval for New Prescription Drugs ............................... 185
Table A.23. Travel Time to Provider Office ........................................................................ 190
Table A.24. Health Care Spending in the Last Year of Life .......................................................196
Table A.25. Spending on Mental Health (Percentage of Total Health Spending) .......................200
Summary

Compared with other Organisation for Economic Co-operation and Development (OECD) member countries, the United States ranks lower on several notable health-related outcomes, such as infant mortality, that tend to be overemphasized by stakeholders and used as a marker of health care system performance. However, the driving factors contributing to many of these rankings could be attributed to factors outside of the health care system, such as greater inequality, less social protection, population density, and societal preferences. Concerns have also been raised around the variability in data availability and quality within the United States and abroad, which can introduce bias or make international comparisons more difficult. Given the increasingly important question of what the United States is buying for its health expenditures and interest in understanding what factors of the health system contribute to better health outcomes, in 2020, the Immediate Office of the Secretary (IOS) of the U.S. Department of Health and Human Services (HHS) sought to identify measures that align with HHS priorities and could improve the ability of the United States and other countries to learn from international comparisons of health system performance. The IOS sought to identify new measures that are important, valid, feasible for cross-country comparisons, and useful for policymakers and are also aligned with HHS priorities.

To inform the identification of measures for international comparison, RAND Corporation researchers worked with a diverse group of 15 experts in quality measurement, clinical care, and health economics to generate and prioritize potential measure constructs that are particularly promising for international comparisons. This report is designed to lay a strong foundation for decisions around these measure constructs by noting the degree of consensus among experts about their importance, scientific acceptability, perceived feasibility, and usability; summarizing the strengths and limitations of the measure constructs; and providing additional context that could be useful for informing the selection of measure constructs that might ultimately be developed into measures and proposed to OECD for consideration.

Methods

We used a modified Delphi panel approach implemented using RAND’s ExpertLens platform, an online approach to expert elicitation and stakeholder engagement. This process occurred in three phases.

Phase 1 involved the identification of measure constructs that should be considered for inclusion in the expert rating process. Experts submitted 142 recommendations that represented 80 unique measure constructs. RAND researchers then conducted a review of these measure constructs to identify and exclude sets of measure constructs that, while important, either were
not yet specific enough for experts to provide ratings (e.g., innovation) or could warrant a separate discussion or expert review panel, such as in the case of public health preparedness. A summary of these measure constructs is presented in Chapter 6. In addition, some of the measure constructs provided by experts in Phase 1 were closely related to currently reported OECD measures. In many cases, the expert suggestions indicated potential refinements or extensions to the existing OECD measures. Because these measures are already being used by OECD in some form, offering these refinements could help to strengthen international comparisons of health system performance in the near term. Details on current OECD measures and these measure constructs are provided in Chapter 5. In total, 25 measure constructs were identified for the Phase 2 rating process.

Phase 2 included an initial assessment of each of the 25 measure constructs by experts against prespecified rating criteria. Experts were provided evidence tables for each measure and asked to draw on the findings in the evidence tables and their own knowledge and expertise to evaluate each measure construct against the following four criteria: importance, scientific acceptability, feasibility for international comparisons, and usability. Experts were asked to provide a numeric rating from 1 (lowest) to 9 (highest) for each criterion, along with a brief written rationale for their rating using RAND’s ExpertLens platform. After the rating period closed, the RAND team summarized the numeric and written assessments for each measure construct. These data, coupled with the distribution of responses, were used to identify the measure constructs that would be discussed in the Phase 3 meeting.

In Phase 3, the RAND team hosted a virtual convening of experts to discuss the ratings, explore areas of disagreement, and allow experts to adjust ratings, if desired. In total, of the 25 measure constructs, ratings for ten measure constructs were unambiguous, and no further discussion was warranted. However, experts asked to move two of these (self-reported pain and estimates of administrative complexity and cost) to discussion. One measure construct (treatment and control of hypertension) received such consistently high ratings by experts in Phase 2 that it was not necessary to discuss it in Phase 3. The remaining seven measure constructs were not discussed by experts in Phase 3, due to their lower initial ratings. In total, 17 measures were discussed in the Phase 3 virtual convening.

Results

Of the 25 measure constructs included in the expert rating process, eight (Table S.1, column 1) were rated by experts as having the most promise for international comparisons, meaning that experts scored the measure high (median 7 or higher) on three of the four rating criteria. In many cases, experts came to an agreement on key assumptions or revisions to the measure construct that strengthened its potential. All of these, however, will require additional development work to establish their operational definitions and specifications to ensure that any measure developed is valid and feasible for international comparisons of health system performance. Despite these
concerns, experts felt that overcoming these challenges was achievable and, in many cases, pointed to measures that are currently in use or to work that is underway in developing or validating potential measures that could be leveraged.

Ten measure constructs (Table S.1, column 2) were determined by experts to hold promise but will require additional refinements to the measure construct prior to moving toward the development of an operational definition. Experts suggested ways to refine, focus, or split the measure constructs, providing a helpful direction for how the measure construct may be strengthened. Many of the perceived limitations of these measure constructs centered around two major concerns: (1) measure constructs that might be impacted by factors outside of the health system, making them less useful for measuring health system performance, and (2) variability in data availability and quality within the United States and abroad, which could introduce bias or make international comparisons more difficult.

As noted earlier, seven measure constructs (Table S.1, column 3) were not discussed by experts in Phase 3, due to their lower initial ratings by experts. For some of the measure constructs, experts were split, with some providing strong support and others calling out what they considered to be critical flaws or challenges with the measure construct that would be difficult to overcome. Other measure constructs were not met with strong opposition, but they also did not have a champion among experts, with ratings squarely in the middle of the range. While these measure constructs hold merit and could provide useful information for strengthening the knowledge base of important areas of inquiry, particularly within the United States, they require significant work before they can be used for international comparisons of health system performance.
Table S.1. Expert Reviews of Measure Constructs

<table>
<thead>
<tr>
<th>Measure Constructs That Hold the Most Promise</th>
<th>Measure Constructs That Received Lower or Inconsistent Ratings</th>
<th>Measure Constructs That Were Not Discussed by Experts</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Treatment and control of hypertension</td>
<td>• Self-reported pain</td>
<td>• Clinician workforce who can prescribe medication-assisted treatment or naloxone</td>
</tr>
<tr>
<td>• Access to and coverage for telehealth</td>
<td>• Access to primary palliative care</td>
<td>• Access to opioid treatment centers</td>
</tr>
<tr>
<td>• Quality-adjusted life expectancy</td>
<td>• Prices for brand-name and generic drugs</td>
<td>• The percentage of patients with a follow-up visit within four weeks of starting an opioid for chronic pain</td>
</tr>
<tr>
<td>• Insurance coverage for mental health, behavioral health, and substance abuse services</td>
<td>• Diffusion of and access to new prescription drugs</td>
<td>• Time to regulatory approval for new prescription drugs</td>
</tr>
<tr>
<td>• Receipt of preference-concordant end-of-life care</td>
<td>• Avoidable emergency department use</td>
<td>• Travel time to provider office</td>
</tr>
<tr>
<td>• Care continuity or consistent provider</td>
<td>• The percentage of patients with an opioid use disorder who were referred to or prescribed medication-assisted treatment</td>
<td>• Health care spending in the last year of life</td>
</tr>
<tr>
<td>• Access to mental health providers</td>
<td>• Estimates of administrative complexity and cost</td>
<td>• Spending on mental health (percentage of total health spending)</td>
</tr>
<tr>
<td>• Data transfer and interoperability</td>
<td>• Disadoption of ineffective medical services</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Healthy days at home</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Availability of emergency medical services to prevent opioid death</td>
<td></td>
</tr>
</tbody>
</table>

Recommendations

Detailed recommendations for what is needed to move each measure construct forward are provided in the body of the report. In addition, a number of cross-cutting recommendations and key decision points emerged from this effort that could help to shape next steps.

**Pursue measure specification for promising measure constructs.** Experts rated eight of the measure constructs as having the most promise for international comparisons. Next steps will require additional development work to establish the operational definitions and specifications of these measure constructs to ensure that any measure developed is valid and feasible for international comparisons of health system performance. In many cases, experts pointed to existing measures that could be leveraged for this purpose.

**Develop guidelines for assessing “health care system performance.”** Some experts felt that if a measure construct could be influenced by factors outside of the health system (e.g., in the case of the construct “healthy days at home”), that construct was less useful for international comparisons of health care system performance given differences in populations, cultural preferences, and social conditions across countries. Other experts noted, however, that there
could be ways to account for these differences in analysis (e.g., using regression analysis) to isolate the effect of health system performance on the outcome of interest. This ambiguity was particularly challenging for measure constructs that were both highly relevant for the health system and influenced by external factors. These measure constructs include quality-adjusted life expectancy, which is influenced by social, economic, and geographic factors, and the approval and diffusion of new drugs, which are shaped by regulatory guidelines and industry processes. Developing clear guidelines to assess the alignment of measure constructs with health system performance will help ensure that such decisions are being applied consistently and help determine which measure constructs that were flagged with this concern may warrant further development.

**Disentangle and prioritize related constructs of access, coverage, utilization, and outcomes.** Several measure constructs focused on issues of access or coverage (e.g., “Access to mental health providers”), but experts were quick to point out that coverage for or access to services does not necessarily equate to appropriate utilization of those services or result in improvements in health outcomes. As a result, in some cases, experts recommended strengthening the measure construct by focusing on utilization or the expected health outcomes. However, experts also noted that coverage of and access to specific types of health care providers, services, and treatments could be important intermediate outcomes that could shed valuable insight into key outcomes of interest. As the measure constructs undergo further development work, it will be beneficial to provide additional specificity around these related constructs to help ensure that what is being measured will reflect the intended health outcome.

**Invest in developing novel measure constructs.** As outlined in Chapter 6, experts provided a number of ideas for measure constructs that are important but were deemed not specific enough to advance to the rating stage. Experts who suggested these concepts often mentioned that measurement would be difficult but still recommended panel consideration because of their importance. One example is the idea of innovation, which was offered by ten experts in Phase 1. Although experts nominated different areas for innovation (e.g., drugs, diagnostics, health information technology), this and other proposed measure constructs may be good candidates for measurement development work by HHS or other federal agencies in the near future.

**Offer refinements to existing OECD measures.** In Phase 1, experts nominated measure concepts for consideration. Many of these were refinements to or expansions of existing OECD measures, with suggestions for additional specification or recommendations to further parse out the data by key subgroups that would facilitate subgroup analyses and more direct, and appropriate, cross-country comparisons (see Chapter 5). Because these measures are already being used by OECD in some form, offering these refinements for OECD consideration may help to strengthen international comparisons of health system performance in the near term.
Limitations

All of the experts were based in the United States. While many of the project experts do have expertise in international comparisons of health care quality, they may not be in the best position to judge how relevant, useful, and feasible a measure construct is for other countries. There may be benefits to engaging an international audience in the future development of these measure constructs.

Experts raised the limitation that this project’s U.S.-centric approach to construct selection might result in identification of measure constructs that are less important or relevant for other countries. For example, the opioid crisis is a more significant concern in the United States and is not necessarily relevant for international benchmarking. Where experts raised this as a concern in their comments or discussion, it was noted.

Conclusion

Collectively, experts provided a wealth of information on potential measure constructs that could strengthen international comparisons of health care system performance. Eight measure constructs were identified as having the most promise, although all require additional work to strengthen the measure specification and address the feasibility of data collection in an international setting. Experts felt that many other measure constructs were important and held promise but would require additional investment to develop. In the near term, experts pointed to a number of opportunities to strengthen the existing OECD measures set, providing important refinements to measure specifications and how data can be collected and presented to facilitate more appropriate and accurate international comparisons. These suggestions could be explored with OECD sooner, while the new measure constructs suggested by experts are developed into feasible, reliable, and valid measures suitable for international comparisons.
We gratefully acknowledge the support and assistance of several people in writing this report. First, we thank the following experts for their continued involvement and participation in this effort: Scott Atlas (Hoover Institution), Jay Bhattacharya (Stanford University), Helen Burstin (Council of Medical Specialty Societies), David Cutler (Harvard University), Patricia Danzon (University of Pennsylvania), Bianca Frogner (University of Washington), Craig Garthwaite (Northwestern University), Jeremiah (Awori) Hayanga (West Virginia University), Regina Herzlinger (Harvard University), Jonathan Ketcham (Arizona State University), Mary Naylor (University of Pennsylvania), Stephen Parente (University of Minnesota), Avik Roy (Foundation for Research on Equal Opportunity), and Eric Schneider (the Commonwealth Fund). We thank Patrick Romano (University of California, Davis), who graciously agreed to chair and facilitate the June 2020 meeting. We would also like to thank four federal experts for their advice and guidance: Irma Arispe (Centers for Disease Control and Prevention), Jeff Brady (Agency for Health Research and Quality [AHRQ]), Joel Cohen (AHRQ), and Steven Sheingold (HHS). We thank our HHS and Office of the Assistant Secretary for Planning and Evaluation project leads, Mamatha Pancholi, Pamela Owens, Christie Peters, and Rose Chu, for their input and guidance.

We also thank several colleagues at RAND for their assistance: Christine Eibner for serving in a senior advisory role on this project; Hamad Al Ibrahim, Erika Bloom, Joan Chang, and Lilian Perez for help with creating the evidence tables; and Karishma Patel for administrative assistance. We are also grateful to Christine Eibner, Paul Koegel, and Rachel Reid (RAND) and Soeren Mattke (University of Southern California) for their thoughtful review of this work.
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACA</td>
<td>Affordable Care Act</td>
</tr>
<tr>
<td>ACSC</td>
<td>ambulatory care–sensitive conditions</td>
</tr>
<tr>
<td>AOD</td>
<td>acute opioid disorder</td>
</tr>
<tr>
<td>BPI</td>
<td>Short-Form Brief Pain Inventory</td>
</tr>
<tr>
<td>CAHPS</td>
<td>Consumer Assessment of Healthcare Providers and Systems</td>
</tr>
<tr>
<td>CDC</td>
<td>Centers for Disease Control and Prevention</td>
</tr>
<tr>
<td>CHF</td>
<td>congestive heart failure</td>
</tr>
<tr>
<td>CMS</td>
<td>Centers for Medicare &amp; Medicaid Services</td>
</tr>
<tr>
<td>COPD</td>
<td>chronic obstructive pulmonary disease</td>
</tr>
<tr>
<td>COVID-19</td>
<td>coronavirus disease 2019</td>
</tr>
<tr>
<td>ED</td>
<td>emergency department</td>
</tr>
<tr>
<td>EHR</td>
<td>electronic health record</td>
</tr>
<tr>
<td>EMS</td>
<td>emergency medical services</td>
</tr>
<tr>
<td>FDA</td>
<td>U.S. Food and Drug Administration</td>
</tr>
<tr>
<td>GDP</td>
<td>gross domestic product</td>
</tr>
<tr>
<td>GIS</td>
<td>geographic information systems</td>
</tr>
<tr>
<td>GNP</td>
<td>gross national product</td>
</tr>
<tr>
<td>GP</td>
<td>general practitioner</td>
</tr>
<tr>
<td>HALE</td>
<td>healthy life expectancy</td>
</tr>
<tr>
<td>HDAH</td>
<td>healthy days at home</td>
</tr>
<tr>
<td>HHS</td>
<td>U.S. Department of Health and Human Services</td>
</tr>
<tr>
<td>HIC</td>
<td>high-income country</td>
</tr>
<tr>
<td>HRQoL</td>
<td>health-related quality of life</td>
</tr>
<tr>
<td>HRSA</td>
<td>Health Resources and Services Administration</td>
</tr>
<tr>
<td>ICT</td>
<td>information and communications technology</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>IOS</td>
<td>Immediate Office of the Secretary</td>
</tr>
<tr>
<td>IQR</td>
<td>interquartile range</td>
</tr>
<tr>
<td>IT</td>
<td>information technology</td>
</tr>
<tr>
<td>LIC</td>
<td>low-income country</td>
</tr>
<tr>
<td>LMICs</td>
<td>low- and middle-income countries</td>
</tr>
<tr>
<td>MAT</td>
<td>medication-assisted treatment</td>
</tr>
<tr>
<td>MIC</td>
<td>middle-income country</td>
</tr>
<tr>
<td>MSRP</td>
<td>manufacturer’s suggested retail price</td>
</tr>
<tr>
<td>NBER</td>
<td>National Bureau of Economic Research</td>
</tr>
<tr>
<td>NHATS</td>
<td>National Health and Aging Trends Study</td>
</tr>
<tr>
<td>NHEA</td>
<td>U.S. National Health Expenditure Accounts</td>
</tr>
<tr>
<td>NHHSUQ</td>
<td>National Health and Health Services Use Questionnaire</td>
</tr>
<tr>
<td>NRS</td>
<td>Numerical Rating Scale</td>
</tr>
<tr>
<td>NSDUH</td>
<td>National Survey on Drug Use and Health</td>
</tr>
<tr>
<td>N-SSATS</td>
<td>National Survey of Substance Abuse Treatment Services</td>
</tr>
<tr>
<td>OECD</td>
<td>Organisation for Economic Co-operation and Development</td>
</tr>
<tr>
<td>OTP</td>
<td>opioid treatment program</td>
</tr>
<tr>
<td>OUD</td>
<td>opioid use disorder</td>
</tr>
<tr>
<td>PPC</td>
<td>primary palliative care</td>
</tr>
<tr>
<td>QALE</td>
<td>quality-adjusted life expectancy</td>
</tr>
<tr>
<td>QALY</td>
<td>quality-adjusted life year</td>
</tr>
<tr>
<td>QUALICOPC</td>
<td>Quality and Costs of Primary Care in Europe</td>
</tr>
<tr>
<td>R&amp;D</td>
<td>research and development</td>
</tr>
<tr>
<td>RCT</td>
<td>randomized controlled trial</td>
</tr>
<tr>
<td>SAMHSA</td>
<td>Substance Abuse and Mental Health Services Administration</td>
</tr>
<tr>
<td>SHS</td>
<td>serious health-related suffering</td>
</tr>
<tr>
<td>SUD</td>
<td>substance use disorder</td>
</tr>
<tr>
<td>VBP</td>
<td>value-based payment</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
WHOAIMS  World Health Organization Assessment Instrument for Mental Health Systems
1. Introduction

Project Overview

Differences in per capita spending and health outcomes between the United States and other countries have been widely reported. Compared with other Organisation for Economic Co-operation and Development (OECD) member countries, the United States ranks in the bottom third for several health outcomes, such as infant mortality,\(^1\) that tend to be overemphasized by stakeholders and used as a marker of health care system performance. However, the driving factors contributing to many of these rankings are commonly attributed to factors outside of the health care system, such as greater inequality and less social protection. Health behaviors, social determinants, structural factors, measurement issues, and data differences likely contribute to international differences in health care spending and outcomes. Identifying a core set of metrics built on comparable data and aligned with health care system priorities can help researchers and policymakers better understand the impact of different structural and policy features on international health spending and outcomes. Given the increasingly important question of what the United States is buying for its health expenditures, and interest in understanding what factors of the health care system contribute to better health outcomes, an examination of the factors contributing to these differences is warranted.

To shed light on this issue, in 2020, the Immediate Office of the Secretary (IOS) of the U.S. Department of Health and Human Services (HHS) sought to identify measures that could improve the ability of the United States and other countries to learn from international comparisons of health care system performance. The IOS sought to identify new measures that are important, valid, feasible for cross-country comparisons, and useful for policymakers and are also aligned with HHS priorities. HHS Secretary Alex Azar identified four priorities for HHS: the opioid crisis, health insurance reform, drug pricing, and value-based care. These priorities were developed before the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus was identified. During the course of this project, coronavirus disease 2019 (COVID-19) developed into a pandemic and was declared a national emergency by President Donald Trump on March 13, 2020. Slowing the spread and responding to the immediate effects of COVID-19 has been a central focus of HHS. Although the COVID-19 pandemic provided important context as the experts identified potential measures for international comparison, this report does not focus on COVID-19 measurement.

To inform the identification of measures for international comparison, the RAND Corporation team worked with a diverse group of 15 experts in quality measurement, clinical care, and health economics to generate and prioritize potential measure constructs that are particularly promising for international comparisons. Experts were selected in partnership with
HHS, and we sought to achieve a balance of expertise relevant to this effort. A list of experts and their affiliations is provided in the acknowledgments section of this report.

With few exceptions, the measure constructs identified as most promising will require additional development work to establish their operational definitions and specifications to ensure that any measure developed is valid and feasible for international comparisons of health system performance. This report is designed to lay a strong foundation for these future refinements by noting the degree of consensus among experts about the importance, scientific acceptability, perceived feasibility, and usability of measure constructs; summarizing the strengths and limitations of the measure constructs; and providing additional context that may be useful to inform the selection of measure constructs that could ultimately be developed into measures and proposed to OECD for consideration. A secondary purpose of this report is to share information on promising measure constructs that are ripe for additional development work.

Methods

We implemented a modified Delphi panel approach based on the RAND-UCLA Delphi method implemented via RAND’s ExpertLens platform, an online tool for expert elicitation. This process occurred in three phases:

- Phase 1: Identification of measure constructs that should be considered for inclusion in the expert rating process
- Phase 2: Initial assessment of each measure construct by experts against prespecified rating criteria
- Phase 3: Virtual convening of experts to discuss the ratings, explore areas of disagreement, and then provide final ratings of each measure construct against the criteria.

**Phase 1: Identification of Measure Constructs That Should Be Considered**

In April 2020, the RAND team asked experts to submit measure constructs for consideration that held promise for international comparisons. To support the submission of measure constructs, RAND researchers prepared a source document with relevant background information and quick reference tables to facilitate the experts’ input to this project. This document included, for example, a summary of HHS priorities and potential measure concepts related to these topics. Given that our goal was to identify measures that could be considered for adoption by OECD, it was important that experts had a strong and complete understanding of what is and is not currently measured by OECD (see Chapter 5 for additional detail on current OECD measures); whether and how current measures map to HHS priorities; and known critiques of OECD measures, particularly with regard to those that are commonly used for international comparisons. The document also included a high-level summary of existing
measurement frameworks and health-related domains within these frameworks for use by experts considering current gaps and opportunities for measurement.

Collectively, experts submitted 142 measure constructs. To identify those that would move forward to the expert rating process, we conducted a multistep process. First, we identified duplicate or overlapping submissions, resulting in 80 unique measure constructs. We then conducted a review of these measure constructs to identify and exclude sets of measure constructs that, while important, require additional work before they are specific enough for experts to provide ratings (e.g., innovation), reflect refinements to existing OECD measures, or could warrant a separate discussion or expert review panel, such as in the case of public health preparedness and measure constructs related to COVID-19. Although this group of excluded measure constructs was not included in the Phase 2 expert rating process, these constructs are important and provide valuable insights around opportunities to strengthen existing OECD measures (described in more detail in Chapter 5) or emerging areas that might warrant additional consideration to be shaped into a more specified measure construct (described in more detail in Chapter 6). Twenty-five measure constructs were identified for the Phase 2 rating process.

**Phase 2: Evaluation of Each Measure Construct Against Prespecified Criteria**

To support experts in the Phase 2 rating process, we developed evidence tables for each of the measure constructs (see Appendix A). Experts were asked to draw on the findings in the evidence tables and from their own knowledge and expertise to evaluate each measure construct against the following criteria. These criteria were adapted from similar criteria used by the National Quality Forum and in other measure assessment activities:

- **Importance:** The extent to which the measure construct represents an issue with high health or social impact both within the United States and globally, is aligned with HHS priorities, and provides opportunities for countries to make significant gains in health care performance
- **Scientific acceptability:** The extent to which adequate scientific evidence or professional consensus exists to support that the concept measured is related to health care system performance
- **Feasibility for international comparisons:** The extent to which data are likely to be readily available in OECD countries and could be collected in multiple countries without undue burden
- **Usability:** The extent to which policymakers and other potential audiences are using or could use data from this measure construct to inform policy.

Experts provided ratings through RAND’s ExpertLens platform. Experts were asked to provide a numeric rating from 1 to 9 for each criterion, along with a brief written rationale for their rating. Experts were provided with the following guidance: Scores of 1–3 indicate low ratings, scores of 4–6 indicate moderate or unsure ratings, and scores of 7–9 indicate high ratings. Experts completed their ratings between June 8 and June 18, 2020. After the rating
period closed, the RAND team summarized the numeric and written assessments for each measure construct. ExpertLens generated the median and interquartile range (IQR) of ratings for each criterion within each measure construct. Although absolute distance to the median is another commonly used measure of dispersion, ExpertLens draws on both level of agreement and the value of the median to determine the group decision (i.e., strong, moderate, or low support of a criterion), which is more readily understandable by participants. These data, coupled with the distribution of responses, were used to identify the measure constructs that would be discussed in the Phase 3 meeting. We sought to limit the number of measure constructs that were discussed to ensure that adequate time was available for discussion and to ensure that discussion was limited to those measures that held the most promise for international comparisons based on Phase 2 ratings. In total, ratings for ten measure constructs were unambiguous, and no further discussion of these was warranted.

**Phase 3: Convene, Discuss, and Adjust Ratings**

On June 29, 2020, the RAND team convened experts for a five-hour videoconference to discuss the measure constructs rated in Phase 2 and to provide updated ratings, if they chose, based on the discussion. In advance of the meeting, experts were sent the results of the Phase 2 ratings through the ExpertLens platform, where they were able to see the distribution of ratings for each measure construct and criteria, along with the median and IQR; their own Phase 2 ratings were also superimposed on the distributions.

At the beginning of the meeting, experts were given an opportunity to review the measure constructs that were not slated for discussion based on the criteria noted above and to nominate any for discussion. Two measure constructs were moved to the agenda as a result, and a total of 17 measure constructs were discussed during the meeting (see Table 1.1). The meeting was facilitated by Patrick Romano, an experienced facilitator with expertise in quality measurement and experience with the use and development of OECD measures. For each measure construct, experts first reviewed the distribution of scores and a summary of comments at the lower and higher ends of the rating distribution and were then provided an opportunity to discuss ratings and areas of disagreement. Because many of the measure constructs could be broadly interpreted, in many cases experts agreed on a refinement to the measure construct that helped to clarify or strengthen it, which resulted in increased convergence of expert ratings.

After discussion of each measure, experts re-rated the measure construct through ExpertLens. Experts who did not provide Phase 3 ratings to one or more measure constructs confirmed with us that they would like their Phase 2 ratings to stand. Two of our 15 experts were not able to participate in the Phase 2 discussion and did not provide updated ratings. Unless otherwise noted, this report summarizes the results from the Phase 3 ratings. For each measure construct described in Chapters 2–4, we provide a figure that summarizes the final distribution of expert ratings (yellow bars), median rating (blue line), and IQR (gray shaded rectangle) for each of the four
domains: importance, scientific acceptability, feasibility for international comparisons, and usability.

Summaries of measure constructs are derived from RAND’s analysis of expert commentary during the Phase 3 meeting and comments that experts provided through the ExpertLens platform to justify their numeric ratings. To provide additional context for each measure construct, we included findings from the literature, drawn from the respective evidence table for that measure construct (Appendix A).

Table 1.1. Measure Constructs Included in Expert Ratings

<table>
<thead>
<tr>
<th>Discussed in Phase 3</th>
<th>Not Discussed in Phase 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Availability of emergency medical services to prevent opioid death</td>
<td>1. Treatment and control of hypertension (consistently high expert ratings)</td>
</tr>
<tr>
<td>2. The percentage of patients with an opioid use disorder who were referred to or</td>
<td>2. Clinician workforce who can prescribe medication-assisted treatment or naloxone</td>
</tr>
<tr>
<td>prescribed medication-assisted treatment</td>
<td></td>
</tr>
<tr>
<td>3. Access to and coverage for telehealth</td>
<td>3. Access to opioid treatment centers</td>
</tr>
<tr>
<td>4. Insurance coverage for mental health, behavioral health, and substance abuse</td>
<td>4. The percentage of patients with a follow-up visit within four weeks of starting an opioid for chronic pain</td>
</tr>
<tr>
<td>services</td>
<td></td>
</tr>
<tr>
<td>5. Avoidable emergency department use</td>
<td>5. Time to regulatory approval for new prescription drugs</td>
</tr>
<tr>
<td>6. Prices for brand-name and generic drugs</td>
<td>6. Travel time to provider office</td>
</tr>
<tr>
<td>7. Estimates of administrative complexity and cost</td>
<td>7. Health care spending in the last year of life</td>
</tr>
<tr>
<td>8. Quality-adjusted life expectancy</td>
<td>8. Spending on mental health (percentage of total health spending)</td>
</tr>
<tr>
<td>9. Diffusion of and access to new prescription drugs</td>
<td></td>
</tr>
<tr>
<td>10. Access to primary palliative care</td>
<td></td>
</tr>
<tr>
<td>11. Receipt of preference-concordant end-of-life care</td>
<td></td>
</tr>
<tr>
<td>12. Healthy days at home</td>
<td></td>
</tr>
<tr>
<td>13. Care continuity or consistent provider</td>
<td></td>
</tr>
<tr>
<td>14. Access to mental health providers</td>
<td></td>
</tr>
<tr>
<td>15. Data transfer and interoperability</td>
<td></td>
</tr>
<tr>
<td>16. Disadoption of ineffective medical services</td>
<td></td>
</tr>
<tr>
<td>17. Self-reported pain</td>
<td></td>
</tr>
</tbody>
</table>
Organization of This Report

Chapter 2 provides a summary of the eight measure constructs that received the highest ratings from experts after Phase 3. Chapter 3 provides a summary of the measure constructs that were discussed by experts but received lower ratings after Phase 3, suggesting that more refinement to these measure constructs could be warranted prior to offering them as suitable measures for international comparisons. Chapter 4 presents measure constructs that received lower ratings in Phase 2 and were not discussed by experts. In Chapter 5, we discuss measure constructs suggested by experts that offered important refinements and extensions of existing OECD measures to strengthen international comparisons. In Chapter 6, we discuss the remaining measure constructs that were excluded from expert ratings for other reasons, including emerging ideas that may warrant additional consideration to be shaped into a more specified measure construct. The evidence tables that the RAND team developed to support expert ratings are included in Appendix A.
2. Measure Constructs That Hold the Most Promise for International Comparisons

Eight measure constructs were identified that received the highest ratings in the expert elicitation process:

- treatment and control of hypertension
- access to and coverage for telehealth
- quality-adjusted life expectancy
- insurance coverage for mental health, behavioral health, and substance abuse services
- receipt of preference-concordant end-of-life care
- care continuity or consistent provider
- access to mental health providers
- data transfer and interoperability.

For these eight measure constructs, experts rated three out of four criteria (importance, scientific acceptability, feasibility for international comparisons, usability) as high, meaning that the median of the distribution for that criteria was a 7 or higher on a scale of 1 to 9. In this chapter, we show the distribution of expert ratings, summarize key strengths of the measure constructs, and note where disagreement still exists. Although each of these measure constructs holds promise, for each measure we also provide a summary of what experts felt was needed to move this measure construct forward for international comparisons.

Treatment and Control of Hypertension

Description Provided in Evidence Table

Hypertension is blood pressure that is higher than normal. In the United States, the threshold for a hypertension diagnosis is defined as systolic blood pressure consistently at or greater than either 130 or 140 mm Hg and diastolic blood pressure consistently at or greater than either 80 or 90 mm Hg. The population of individuals tested and treated for hypertension typically includes adults aged 18 and older.

- Treatment of hypertension: There are two components of treatment that can be used in isolation or in combination: (1) antihypertensive medications and (2) lifestyle changes (e.g., physical activity and diet).
- Control of hypertension: Control is achieved when the individual’s blood pressure measures are consistently lower than 140 mm Hg (systolic) and lower than 90 mm Hg (diastolic).
**Strengths and Limitations**

As shown in Figure 2.1, experts provided consistently high ratings for importance, scientific acceptability, feasibility for international comparisons, and usability. Because this measure construct received consistently high scores in Phase 2, it was not discussed at the Phase 3 meeting and was not re-rated by experts. As a result, data in Figure 2.1 reflect Phase 2 expert ratings. In comments provided by experts to justify their ratings, most agreed that this was an important measure, noting that “it is one of the biggest health problems we have” and that it is “one of the key chronic diseases and a risk for many serious conditions that decrease life expectancy.” Data suggest that an estimated 1.13 billion people have hypertension, at a global cost of $370 billion annually. Experts also pointed to the availability of currently used measures in the United States and abroad that assess hypertension treatment and control.

Two experts cautioned, however, against measuring “treatment” of hypertension, noting that “we are too quick to prescribe” and that “assessing only ‘treatment’ is not sufficient.” They went on to note that the CDC [Centers for Disease Control and Prevention], ACC [American College of Cardiology], AHA [American Heart Association], and others routinely assess ‘control’ as the more meaningful indicator.” With respect to the feasibility for international comparisons, our evidence table reported that there is international variation in diagnostic thresholds, guidelines for the initiation of treatment (e.g., lifestyle changes, medications, no explicit guidelines), and the focus on total cardiovascular risk rather than only blood pressure when initiating treatment. Related to this latter point, one expert noted that the measure “would be of greater value [if it included] glycemic control and obesity.”
What Is Needed to Move This Measure Forward

Discrepancies in the threshold for hypertension treatment both within and across high-income nations would need to be resolved. A key decision would be whether to collect data on hypertension treatment, control, or both. How the data are collected would also need to be resolved because data describing treatment (e.g., medication utilization, lifestyle changes) are typically self-reported but have potential biases due to differences in the structure of the survey questions and cross-cultural differences in responses. Hypertension control, however, could be described by administrative data that document those who have been diagnosed with hypertension and their recent blood pressure readings. Among the input sources for the OECD data, diagnosed diseases are typically documented by national disease surveillance systems. However, describing control of hypertension would require the integration of advanced health information technology systems with repeated measures of blood pressure.

Access to and Coverage for Telehealth

Description Provided in Evidence Table

There is no widely agreed-upon definition of telehealth; however, the Health Resources and Services Administration (HRSA) of HHS defines telehealth as “the use of electronic information and telecommunication technologies to support long-distance clinical health care, patient and professional health-related education, public health, and health administration. Technologies include video conferencing, the internet, store-and-forward imaging, streaming media, and terrestrial and wireless communications.”

HHS currently describes four modalities of telehealth: (1) live video interaction between a patient and provider, (2) store and forward (electronic transmission of videos and digital images to a provider in a different location), (3) remote patient monitoring (electronic transmission of personal health and medical data, such as blood pressure readings from an individual in one location to a provider in another location), and (4) mobile health (mHealth, which includes smartphone apps intended to promote health and well-being).

Strengths and Limitations

After discussion, experts rated this measure concept high on importance, scientific acceptability, and usability, but they scored it in the moderate range on feasibility for international comparisons (Figure 2.2). Experts noted that telehealth is an important measure construct but one that is in “dynamic evolution” in the current COVID-19 pandemic. One expert noted that the recent expansion of telehealth services “created a lot of enthusiasm on the ground, particularly for rural populations and those who previously had not been able to travel to a provider office.” While experts pointed to the value of telehealth as a measure construct, for it to
be successful, experts noted a need to disentangle the two separate concepts of access and insurance coverage and recommended that the measure focus on coverage and payment for telehealth.

Several concerns were raised with regard to a measure of access to telehealth. First, it is not yet clear what the “right” level of access to telehealth is, given the value of in-person interactions and a lack of data on the longer-term outcomes of telehealth services. Experts also cautioned that the lack of a clear definition of telehealth and variability in broadband access could lead to skewed data. While some experts raised the concern that obtaining access to data would require the use of surveys that could have limitations for international comparisons, others pointed to examples where surveys are used successfully for international comparisons (e.g., the Commonwealth Fund International Health Policy Survey). Insurance coverage for telehealth, however, was viewed by experts as “the most important and the place where CMS [the Centers for Medicare & Medicaid Services] has been most engaged” and offering “an important policy lens” where “international comparisons could be helpful.” Experts cautioned that coverage does not necessarily translate to access and use. Our evidence table included findings that reimbursement and regulations vary by state, and, during the COVID-19 pandemic, Medicare and Medicaid expanded coverage of telehealth services and lifted requirements that previously restricted telehealth services, which may make estimates of coverage for telehealth services somewhat challenging in the near term.
What Is Needed to Move This Measure Forward

The definition of telehealth (specifically, which types of services count as telehealth) would need to be agreed on first in order to design a measure that can accurately reflect international differences in access to and coverage for telehealth. Experts felt that this measure construct should be refined to focus on coverage and payment for telehealth, specifically the percentage of the population with health care coverage for patient-to-provider telehealth services. Experts also noted that there may be benefits to further bounding the measure to populations where coverage for telehealth may be particularly beneficial, such as senior or rural populations or chronically ill patients.
Quality-Adjusted Life Expectancy

Description Provided in Evidence Table

Quality-adjusted life expectancy (QALE) is a summary measure of population health that combines information on fatal and nonfatal health outcomes. Data describing (1) health-related quality of life (HRQoL) and (2) life expectancy (by characteristics of interest, such as age, sex, race, and ethnicity) are required to calculate QALE. QALE has been used to monitor changes in population health over time, compare population health across countries, investigate health inequalities, and quantify the benefits of health interventions in cost effectiveness analysis.27

Strengths and Limitations

After discussion, experts rated this measure construct high on importance, scientific acceptability, and feasibility, although there was some disagreement (Figure 2.3). Experts varied widely in their perception of the usability of this measure construct.

Experts considered QALE an important measure construct that provided an improvement over the more commonly used measure of life expectancy. As noted in our evidence table, comparing a nation’s life expectancy and its QALE can help stakeholders understand the additional toll that morbidity takes on society. Moreover, comparing QALE for different subgroups within one nation (e.g., men versus women) can lend insight into how different subgroups experience outcomes disproportionately. Experts noted that the current widespread use of QALE in research and the availability of existing tools to measure it were considered strengths of the measure construct. Of note, the HHS Healthy People 2020 goals include increasing the quality of life and years of
healthy life, both of which make up QALE. Experts, however, raised concern about its use for international comparisons of health care system performance. They noted that while QALE is a good measure to assess the health of a community, given the numerous factors that influence life expectancy (e.g., race/ethnicity, socioeconomic factors), it is not a good measure of health care system performance. One expert noted, for example, that “it tends to be major economic events that change measures of this type.”

**What Is Needed to Move This Measure Forward**

Experts noted that although it may be possible to isolate the impact of the health care system by controlling for other factors that influence life expectancy, it would “require advanced analysis to make sense of the health care system’s contribution.” Another expert noted that QALE could be adjusted in a manner similar to how adjustments are made for purchasing power in cost and pricing data, but experts also noted that more work would be needed to figure out how best to normalize income, the distribution of income, or socioeconomic factors outside of income that influence life expectancy.

It is also worth noting that calculation of QALE requires three sources of data: the probability of death at various ages; data on HRQoL for people in the sample; and a value set (i.e., how individual countries weight the impact of various nonfatal illnesses on health), which is likely to vary across nations due to differences in sociocultural and demographic characteristics. Standardization of this value set would be a necessary step in moving this measure forward. OECD does not currently measure QALE, but it does collect and report data on healthy life years at age 65. Although this is a slightly different measure from QALE, the United States currently does not participate in this measure; one expert noted that submitting data on this measure might be worth considering.

**Insurance Coverage for Mental Health, Behavioral Health, and Substance Abuse Services**

*Description Provided in Evidence Table*

Behavioral health conditions can include mental illnesses, such as anxiety disorders, major depression, schizophrenia, and posttraumatic stress disorder, and substance use disorders (SUDs), such as opioid addiction and alcoholism. Mental health, behavioral health, and substance abuse services can include care in inpatient and outpatient settings, residential treatment centers, day care facilities, and detox centers. Treatments can include pharmaceutical and psychological interventions. There are several measures of insurance coverage to consider, including the following:
• Effective coverage: “People who need health services obtain them in a timely manner and at a level of quality necessary to obtain the desired effect and potential health gains.”

• Universal health coverage: “Universal health coverage means all people receiving the health services they need . . . of sufficient quality to be effective while at the same time ensuring that the use of these services does not expose the user to financial hardship.”

Strengths and Limitations

After discussion, experts rated this measure construct high on importance, scientific acceptability, and usability and in the moderate range for feasibility (Figure 2.4).

Experts rated this measure construct as important, citing data included in RAND’s evidence table that almost one in five American adults live with a mental illness, but only 41 percent of people with mental disorders access any mental health services. According to the World Health Organization (WHO), “The cost of mental health problems in developed countries is estimated to be between 3% and 4% of GNP [gross national product]. Approximately 14% of the global burden of disease can be attributed to these disorders.”

Experts raised several challenges and concerns with how this measure is operationalized, however. First, they noted that insurance coverage is not a useful proxy for access; several experts noted anecdotally that they “know of many psychiatrists who will not take Medicare, Medicaid, or commercial insurance.” Many agreed that access is important to consider because coverage is not helpful if you cannot be seen or it is cost-prohibitive. Second, some experts recommended splitting apart mental, behavioral, and substance abuse services, noting that “there is ambiguity in lumping together three different services”; another noted that
the “most ambiguous is behavioral health, which can mean different things for different people.” However, others countered that although splitting apart the measure could be useful, this creates sample-size issues for survey measures, which are likely needed if the measure intends to capture availability of or access to services.

**What Is Needed to Move This Measure Forward**

To improve the international comparability and precision of the measure, a list of mental health, behavioral health, and substance use conditions and treatments would need to be included in any measurement definition. Experts disagree, for example, on which programs constitute behavioral health treatment, such as self-help programs like Alcoholics Anonymous. Given the different types of insurance mechanisms across countries (e.g., universal, nationally funded coverage versus more fragmented, private coverage), comparison of coverage might be more difficult and the presence of coverage might not have the same implications for access given different payment structures and cost-sharing requirements (e.g., deductibles, copays). Furthermore, because some U.S. psychiatrists judge public-sector payments as inadequate and only take private-pay patients, a key decision point for this measure construct will be determining whether access to such services should be used in place of, or in combination with, a measure of insurance coverage alone.

**Receipt of Preference-Concordant End-of-Life Care**

*Description Provided in Evidence Table*

*Receipt of preference-concordant end-of-life care* refers to the level of agreement between a patient’s desired end-of-life care and the actual end-of-life care received. These preferences are supposed to be translated into actionable medical orders that are maintained in the patient’s medical record and designed to accompany the patient wherever he or she resides.

**Strengths and Limitations**

After discussion, experts rated this measure construct high on importance, scientific acceptability, and usability but scored it in the moderate range for feasibility (Figure 2.5). One expert commented that “it is more powerful than it looks. . . . when patients come in with advance directives . . . it really changes the expectation, comprehension, and entire patient-provider encounter. . . . having a detailed knowledge of what a patient desires before the end helps whatever we do as providers.” Yet several experts struggled with the broad nature of the measure construct, and specifically the notion of “preference-concordant,” noting that this is “very ambiguous and an amalgam of preference of care, legality of what the patient is requesting, and an understanding of [those preferences].” Experts noted that although there is a literature on advanced directives and following them, development of a measure construct
“would take work on understanding data resources, [the] legal structure for patients to give their preferences, and ways to measure concordance.” Experts also noted more fundamentally that “we need to have advanced directive baseline information before we can measure receipt of it” and that “the issue is that we’re not even eliciting what people want.” In discussion, experts agreed that this measure construct should be refined to focus on **documenting goals of care through advanced directives or similar tools.** It was through this lens that experts provided Phase 3 ratings.

**What Is Needed to Move This Measure Forward**

Experts recommended focusing on the documentation of patient goals of care. Additional work will be needed to flesh out how documentation is defined, as RAND’s evidence table notes that comparisons of different methods of communicating end-of-life preferences are limited by the inconsistent use of valid and reliable measures across studies.\(^{37}\) Specific issues identified that can affect validity include response shift (i.e., when people’s values and choices change over time), a lack of specificity and consistency of documentation, and biases related to chart reviews (e.g., preferences may be “judged invalid” or assessed as concordant when they do not align with care received).\(^{35}\) There are a number of survey measures that assess preference of end-of-life care among older adults (see the Commonwealth Fund International Health Policy Survey,\(^{38}\) for example), which could be helpful starting points for further operationalization of this measure construct.
Care Continuity or Consistent Provider

Description Provided in Evidence Table

WHO defines *continuity of care* as “the degree to which a series of discrete health care events is experienced by people as coherent and interconnected over time and consistent with their health needs and preferences.” Continuity of care can include the quality of care over time, which includes the patient’s experience of a continuous caring relationship with an identified health care professional (i.e., a consistent provider), as well as the provider’s delivery of seamless service through integration, coordination, and the sharing of information between different providers. Three ways of providing care continuity are (1) personal continuity (via an ongoing patient-provider relationship), (2) information continuity (present care takes into account past information), and (3) management continuity (management of a health condition is consistent but also flexible to a patient’s changing needs).

Strengths and Limitations

After discussion, experts rated this measure construct high on importance, scientific acceptability, and usability but moderate for feasibility for international comparisons (Figure 2.6). In discussing this measure construct, experts felt it was particularly important for patients with chronic illnesses like diabetes. Experts also pointed to literature included in RAND’s evidence table that found that continuity of care allows for better support of individuals and a better understanding of their conditions, which in turn increases patient satisfaction, patient-
clinician trust, and quality of communication.\textsuperscript{42–46} About 20 percent of Americans report having no usual source of health care.\textsuperscript{47}

Experts discussed the various ways this measure construct could be operationalized, noting that it could “refer to coordination by a primary care provider or coordination among members of a team who cares for someone with a chronic disease and its many comorbidities.” Experts generally agreed that it was better to focus this measure on a team conceptualization, with one expert noting that we “need to leave space for innovation in this area. Focusing too much on a single provider might be against where we see the system evolving in general. There are opportunities we haven’t exploited yet that might lead to this type of continuity.” Experts also discussed potential data sources, including health insurance claims; administrative data, such as those aggregated by care collaboration companies like PatientPing; and survey data. With respect to claims data, one expert raised the questions of “which entity is submitting the claim” and how we might ensure that claims data can be used to assess continuity across a team of providers. Experts agreed that survey data may be the most feasible data source for constructing a continuity measure that is comparable across countries but noted that one of the challenges is in understanding “expectations around a measure like this. What do people expect [with respect to care coordination]? This varies across countries and can vary within the U.S. by insurance type, age, and other characteristics.” Experts noted that capturing data for this measure construct requires complementary approaches: “The patient and caregiver’s perspective are crucial to understand if there is continuity.”

\textit{What Is Needed to Move This Measure Forward}

Experts noted that fully capturing continuity of care would require utilizing multiple data sources to get at various dimensions and perspectives of care continuity. How to combine those sources and ensure feasibility for international comparisons will need to be determined. RAND’s evidence table, for example, documented that claims-based measures of care continuity, while valuable, have several limitations, including an inability to capture information about interpersonal continuity among clinicians and patients and coordination activities that might occur outside of visits between clinicians or between clinicians and patients. Also, some continuity of care analyses examine continuity during episodes of care (e.g., congestive heart failure [CHF], chronic obstructive pulmonary disease [COPD]) and could exclude a large proportion of patients and claims data; results might therefore not be generalizable.\textsuperscript{48} There are several examples of validated scales (e.g., National Health and Health Services Use Questionnaire for patient-reported continuity of care\textsuperscript{49}) and existing measures of continuity of care used to facilitate international comparisons (the Commonwealth Fund International Health Policy Survey\textsuperscript{50,51}) that might be useful for the development of this measure.
Access to Mental Health Providers

Description Provided in Evidence Table

This construct refers to the availability of mental health providers to meet population behavioral health needs and is one dimension of overall access to mental health care. However, which providers are counted as mental health providers can vary. The term mental health provider broadly includes psychiatrists, psychologists, licensed clinical social workers, counselors, marriage and family therapists, and mental health providers who treat alcohol and other drug abuse. HRSA only includes those with a graduate degree.

Strengths and Limitations

Experts rated this measure construct high on importance, scientific acceptability, and feasibility but moderate on usability (Figure 2.7). RAND’s evidence table documented that there are about 5,565 designated mental health professional shortage areas in the United States. Provider shortages result in lower access to care, long waits for necessary treatment, and provider burnout. Low access to mental health providers can also contribute to unmet mental health needs. In 2019, about 18 percent of United States adults (about 44 million) had a mental illness, but 56 percent of those received no treatment, and 20.6 percent of those who sought treatment reported not being able to receive treatment.

Experts discussed some of the inherent challenges with this measure construct, notably the need to define mental health provider. One expert noted that “there are data in the OECD on primary care providers versus specialty care. The challenge I can see is [that] some primary care physicians are trained in mental health,
addition to psychiatrists.” Another expert noted that “if this measure is framed as a ratio of available providers to the population, it should recognize that a broad workforce can contribute to the mental health workforce.” To this point, experts also discussed whether this measure construct should be one of provider availability or access. “I think access is important, but it tends to be narrow in its scope whenever it’s discussed. If we do talk about access, which is the more critical point, think about the broader set of people who can provide the mental health services we need.”

**What Is Needed to Move This Measure Forward**

To ensure the feasibility of international comparisons, it will be necessary to standardize what types of providers are mental health providers and how to handle the vast amount of mental health care provided by those who are not medical providers. RAND’s evidence table noted, for example, that a substantial amount of mental health care is administered by professionals outside the health sector—e.g., criminal justice, education, and social services. More work is needed to determine which measure construct—one of provider counts or ratios versus one of access to mental health—will be most valuable for international comparisons and policy decisions and to then align the inclusion criteria for mental health provider accordingly.

**Data Transfer and Interoperability**

The transfer and interoperability of health data involves the hardware and software that deal with the storage, retrieval, sharing, and use of health care information. Appropriately exchanging health information among health care professionals and patients allows the doctors, nurses, pharmacists, and other providers to obtain vital medical information. The objectives of health information technology and exchange are to improve the speed, quality, safety, and cost of patient care.

**Strengths and Limitations**

After discussion, experts rated this measure construct high on importance, scientific acceptability, and usability and moderate on feasibility (Figure 2.8). In justifying the higher rating for importance, experts pointed to the potential for health information exchanges to lead to higher-quality care. RAND’s evidence table noted that the transfer and interoperability of health information can reduce medical errors and improve compliance with clinical guidelines and reduce overindicated and contraindicated prescribing. Health information exchange systems, while not yet widespread, have also been shown to increase timely communication, efficient provider-patient interactions, and patient engagement.
Experts discussed numerous efforts that are currently underway to operationalize this measure construct, noting that there may be several promising avenues on which to build, including a pilot study conducted by OECD.\textsuperscript{65} Experts also pointed to the Healthcare Information and Management Systems Society Interoperability & Health Information Exchange Committee, which has proposed eight metric categories as a “starting point for consensus around the measurement for market suppliers, providers and government agencies” and has developed a three-part definition.\textsuperscript{66,67} The Commonwealth Fund International Health Policy Survey of Primary Care Physicians also includes questions on the electronic exchange of information.\textsuperscript{68} One expert noted that “defining those categories of exchange, using the data to determine how much exchange is occurring, and getting from physicians an assessment of whether they are reaching their full potential is another complementary way to look at this.” Experts also pointed to a number of issues that should be decided upon as the measure becomes operationalized: the value of measuring electronic data transfers that are mandated in the United States (e.g., e-prescribing), the usability of the data being exchanged for clinical care (e.g., some are more or less relevant), and “how broad must the interoperability be to count? In the United States it is within a system, but in other countries it is national or regional. In the United States it’s a much narrower concept.”

\textit{What Is Needed to Move This Measure Forward}

Experts noted that the next step is “to identify the most important domains of data transfer and interoperability, and how to measure those in a valid way.” As noted in RAND’s evidence table, the OECD pilot study on this measure construct included measures of (1) the utilization of
electronic health records (EHRs); (2) whether laboratory, radiology, and other results are shared with outside organizations; (3) e-prescription utilization; and (4) telehealth capacities. Given that the goal of this effort is to encourage OECD adoption of a set of promising measures for international comparison, the United States may wish to engage with OECD to suggest refinements based on findings of its pilot study and other measurement work happening in the United States on this measure construct.

Conclusion

These eight measures hold the most promise for international comparisons of health care system performance. Overall, experts scored them high across multiple dimensions of importance, scientific acceptability, feasibility, and usability. Each of them, however, requires additional work to further specify the measure, strengthen feasibility for international comparisons, or continue to build the use case for understanding health care system performance. Despite these concerns, experts felt that overcoming these challenges was achievable and, in many cases, pointed to work underway that is developing or validating potential measures that could be leveraged.
3. Measure Constructs That Received Lower or Inconsistent Ratings After Expert Discussion

Of the 25 measure constructs discussed, ten received lower or inconsistent ratings from experts. In most cases, experts felt that these measure constructs were important but identified a number of challenges and limitations with respect to scientific acceptability, feasibility, and usability that must be addressed if these measure constructs are to move forward. These measure constructs are

- self-reported pain
- access to primary palliative care
- prices for brand-name and generic drugs
- diffusion of and access to new prescription drugs
- avoidable emergency department use
- the percentage of patients with an opioid use disorder who were referred to or prescribed medication-assisted treatment
- estimates of administrative complexity and cost
- disadoption of ineffective medical services
- healthy days at home
- availability of emergency medical services to prevent opioid death.
Self-Reported Pain

*Description Provided in Evidence Table*

*Pain* is physical suffering or discomfort caused by illness or injury. Because pain is subjective, the existence and intensity of pain is often measured by patient self-reports. Elements of self-report could include relevant pain; physical, psychological, emotional, and social domains of functioning; and patient-reported outcomes and defined goals. Pain assessment can also occur through physical examination or behavioral observation when self-report is not possible.

*Strengths and Limitations*

Self-reported pain was briefly discussed by experts in Phase 3 both alone and in combination with other measures, such as access to primary palliative care. After discussion, experts rated it higher on importance and usability, but the scores for scientific acceptability were widely distributed, resulting in median scores in the moderate range (Figure 3.1). Experts felt that this is an important measure construct because it is a function of appropriate medical care that relates to two HHS top priorities: the opioid crisis and value-based care. As one expert succinctly stated, “Few health problems are more important than pain.” Experts widely disagreed on the scientific acceptability and feasibility of this measure, with some pointing to existing measures that are “reasonably well validated, despite subjectivity” that could be used for international comparisons, while others made the point that this measure construct is “very subjective [and] highly dependent on personal, international, and cultural differences.”
What Is Needed to Move This Measure Forward

Experts were split in their interpretation of this measure construct and what it would need to move forward. A few felt that it had potential as is, given existing validated measures, while others felt that it should not move forward at all, given limitations. A few experts, however, recommended a potential approach to strengthening the measure construct that involves pairing it with questions “about pain management medicines or tools” to help understand how well those are working or how well they work in relation to one another (e.g., chronic use of opioids versus other non-narcotic medications). Similarly, another expert noted that this measure construct would be “most useful as an outcome if linked to treatment modalities and their use across countries.” However, more work would be needed to identify the set of pain management approaches that would be paired with self-reported pain to make this a viable and useful measure for international comparisons.

Access to Primary Palliative Care

Description Provided in Evidence Table

Primary palliative care (PPC) is all of the palliative care competencies expected of the primary care physician, including but not limited to assessing a patient’s physical and nonphysical symptoms, assisting the patient and his or her family with establishing appropriate goals of care, communicating the patient’s illness and its trajectory, and affirming preference-concordant end-of-life care. In the United States, palliative care is provided within and outside hospice programs. Primary care focuses on continuity of care and seeing the whole person, which is why the majority of palliative care is administered in that setting.

Providing PPC requires the following: involving an interdisciplinary team of providers to meet patient and family needs; creating systems for routine, reliable assessments of symptoms and care needs; creating pathways to triage patients who screen positive for symptoms or distress; and committing to continuing education and skills training for staff.

Strengths and Limitations

After discussion, experts rated access to PPC higher on importance and usability but in the moderate range for scientific acceptability and feasibility (Figure 3.2). According to RAND’s evidence table, WHO estimates that more than 20 million people require palliative care worldwide; most are adults over 60 years of age (69 percent), and 6 percent are children. Other researchers have estimated that 45 percent of all people who died in 2016 worldwide experienced serious health-related suffering (SHS) requiring palliative care (25.5 million deaths). This percentage is estimated to increase to 47 percent by the year 2060, accounting for more than 48 million deaths.
Experts felt that this measure construct was important, but several struggled with how to operationalize it. “The biggest problem with this measure is that [it] is so vague that I don’t know what we are actually measuring.” Experts discussed that a measure of pain and its treatment would be an important component of this measure. “I agree that . . . linking to self-reported pain is an important way to understand what is going on.” But experts also discussed the importance of the patient experience and perception of options to manage symptoms. Experts cautioned about placing too much emphasis on pain management because they “don’t want a situation where countries rush to prescribe opioids to get a higher score.” One expert “read this measure construct as a decision [about where to seek care] and resources to care for patients outside of the hospital” and cautioned that “community-based care may not be understood in the same way in every country.” Experts agreed that a family of measures would be necessary to adequately capture this measure construct. “I like this measure, but I think looking at it as a family of measures would be useful. There may be some opposition to an overarching measure. If we can do it in a culturally sensitive way to recognize different norms on how to do this, I think the effort is worthwhile, especially as populations age around the world.”

What Is Needed to Move This Measure Forward

This measure construct is multidimensional, and a necessary next step is to operationalize the relevant domains. Experts noted that pain is a sentinel component of this measure but that it should also include measures of the patient experience. One expert noted that the Consumer Assessment of Healthcare Providers and Systems (CAHPS) hospice survey might have some relevant, validated questions that could be
leveraged. Work is also needed to ensure that the setting for PPC is well defined to facilitate international comparisons.

**Prices for Brand-Name and Generic Drugs**

*Description Provided in Evidence Table*

Pharmaceuticals have enabled a multitude of improvements in health outcomes for patients. One in every five health care dollars is spent on pharmaceuticals in OECD countries.\(^76\) While overall spending has not changed in recent years, it has become increasingly skewed toward expensive specialty medicines.\(^77\) Measuring drug pricing at the international level could lead nations to make decisions that impact prices, making them available more broadly and/or to specific subpopulations of patients.

Brand-name and generic prescription drug pricing is dependent on quantity of services, mix of compounds, prices, and other factors. International comparisons are best done with sales of single-molecule products.\(^78\)

**Strengths and Limitations**

After discussion, experts felt that the measure of drug prices was important, particularly given findings summarized in RAND’s evidence table that annual expenditures on pharmaceuticals exceed half a trillion dollars in the United States and account for 17 percent of national spending on health care.\(^79\) Americans have reported skipping prescriptions because of cost, and this phenomenon is more common for those who have lower income and more chronic conditions.\(^80\) Drug prices are also one of HHS’s top priorities. Experts also rated this construct
higher on feasibility, although the distribution for scientific acceptability and usability was wider, resulting in median ratings in the moderate range (Figure 3.3).

Experts pointed to a number of issues that make capturing pricing difficult, including cost-sharing and bargaining power that large insurers or pharmacy benefit managers have that could affect drug prices. These challenges caused some experts to question how this measure is linked to health care system performance. Experts also discussed the pros and cons of various prices, including the manufacturer’s suggested retail price (MSRP), list price, post-rebate price, negotiated price (for many OECD countries), and net price. Although several experts agreed that the post-rebate price would be useful conceptually, they also were quick to note that this was unrealistic given the lack of data. One expert further noted that the post-rebate price “isn’t reflective of the net price, because rebates flow to payers, not consumers.” One expert proposed using the list price, or MSRP, because it can be collected. Another proposal was made to utilize the list price because it “is the least worst option [and] the most informative.”

However, other experts questioned the value of “studying a price that we know nobody is paying and using it as a metric to assess the quality of the health care system. It is more dangerous to put out a price you know is false than not have information on pricing at all.”

**What Is Needed to Move This Measure Forward**

This measure construct would benefit from more in-depth study and discussion of health economists and experts in drug pricing to identify the optimal approach for operationalizing this measure for international comparisons. OECD published a report on prescription drug pricing in 2008; however, this could be updated to include all participating OECD countries and more-recent drug prices.81 Experts did not reach an agreement on how best to measure pricing for drugs, and RAND’s evidence table indicates that comparing prices within and between countries is complicated by differences in product availability and formulations between countries, the proprietary nature of the rebates, different manufacturer charges for different payers, inclusion of different market sectors (retail versus inpatient versus outpatient), variability in definition of price, use of data describing price versus spending, distribution costs, value-added tax rates, and subpopulations of patients included in the evaluation (if inclusion varies based on insurance coverage).1,80 Standardization of these issues will be necessary.

**Diffusion of and Access to New Prescription Drugs**

*Description Provided in Evidence Table*

Pharmaceutical innovations have led to reductions in morbidity and mortality due to conditions such as infectious diseases, cardiovascular diseases, and cancer. However, barriers to widespread diffusion of and access to new medications exist. For example, some health care systems could be slower to adopt new medications despite good evidence of benefits because of
poor communication about such benefits or because of policies such as price regulation. In the literature, the following definitions of diffusion of new drugs have been identified:

- the share of daily doses and of expenditures in a therapeutic class that is captured by new medicines\textsuperscript{82}
- the percentage of the population that has access to new products, by drug or class, reported by years since first approval (or use) anywhere in the world.\textsuperscript{83}

Strengths and Limitations

After discussion, experts rated the diffusion of and access to new prescription drugs higher on importance, but ratings were varied for scientific acceptability, feasibility, and usability (Figure 3.4). Experts offered a number of caveats and a recommendation to split this measure construct into two or more distinct measures. Several experts noted that the emphasis of this measure should be on “new prescription drugs that have the maximum public health impact.” They noted that there are a lot of medications coming to market with modest incremental benefits: “Simply measuring the total number of medications approved in a country says nothing about their scientific or public health importance.” Experts also discussed whether this measure would be further strengthened by assessing what share of people who need the medication are getting it, and one noted that it is not the approval of a drug or where it was developed that is important but “what utilization management is there” to ensure that patients get access to the drug. At least one expert advocated for a measure construct that was more focused on the development of drugs created in
“novel ways like gene therapy or other techniques to spur more competition in what is a vast market.”

What Is Needed to Move This Measure Forward

Experts recommended splitting this measure construct into several distinct measures:
1. a measure of access that captures whether the people who need a new medication can access it
2. a measure that captures diffusion or delivery of a new drug to the populations who need the drug
3. a measure that captures innovation in the marketplace and the timing of bringing new drugs to market.

RAND’s evidence table also cited a challenge for international comparisons that should be addressed, which is that access to novel but low-value medications is not necessarily an indicator of good health system performance, particularly if those drugs are expensive. Some countries, such as the United Kingdom, conduct a cost-benefit analysis and might not allow certain drugs to be covered if their effectiveness is deemed low.84 This lack of access to lower-value drugs runs counter to the policy of many countries and could count against the United Kingdom in assessing health care system performance, which might bias international comparisons.

Avoidable Emergency Department Use

Description Provided in Evidence Table

Avoidable (or “preventable”) emergency department (ED) visits are those in which people use hospital EDs for nonurgent care and for conditions that could have been treated in a primary care setting.85 These are discharged ED visits not requiring any diagnostic or screening services, procedures, or medications.86 ED visits for ambulatory care–sensitive conditions (ACSCs) are considered avoidable.87 Avoidable ED visits can be experienced by all populations, regardless of sociodemographic or health insurance factors.87

Strengths and Limitations

Experts noted that avoidable ED use is an important measure construct in the United States (Figure 3.5) because “we view [ED visits] as a high-cost visit” and because it “reflects underlying primary care capacity.” Data included in RAND’s evidence table suggest that in the United States, ED visits account for 12.5 percent of national health expenditures ($328.1 billion U.S. dollars in 2010). Furthermore, avoidable ED visits accounted for $64.4 billion or 2.4 percent of national health expenditures.88
However, experts provided lower and more-diverse ratings for scientific acceptability, feasibility, and usability. Experts raised a number of concerns with using this measure to make international comparisons. First, “different health care systems around the world have different policies for what gets treated where. What might be a failure in the U.S. might not be a failure in other countries and vice versa.” One expert also noted that ED visits “are contingent upon the availability of alternative resources.” In areas where there are few alternatives for care in both the United States and in other countries, what we consider to be medically avoidable may actually be necessary to ensure receipt of high-quality care.

Experts also cautioned that this measure is commonly analyzed after the incident and diagnosis, which can be problematic. “We have to be careful about visits that are ex-post deemed to be avoidable as opposed to those [considered avoidable] ex-ante, or what a reasonable individual would think [did not need to be treated in a hospital].” RAND’s evidence table similarly noted that there is wide variability in the data sources used across studies to define avoidable ED visits (e.g., retrospective diagnoses, hospital admissions, triage scores, patient self-reported data), and validity of these measures is not always reported.

What Is Needed to Move This Measure Forward

Experts felt that the success of this measure would require clearer definitions and measurement specifications around the term avoidable. Analyses of avoidable hospitalizations commonly use claims analyses, but alternative data sources are likely necessary to more accurately reflect avoidable visits when the diagnosis is
not yet known. Additional measure specifications could require the exclusion of populations (e.g., in rural areas) for whom the ED is the only source of after-hours medical care.

The Percentage of Patients with an Opioid Use Disorder Who Were Referred to or Prescribed Medication-Assisted Treatment

*Description Provided in Evidence Table*

Medication-assisted treatment (MAT) is the most effective medical treatment for opioid use disorder (OUD). Despite the prescribing options available, estimates suggest that only 20–40 percent of people with OUD receive MAT. The referral treatment process consists of assisting a patient with accessing MAT, selecting treatment facilities, and helping to navigate any barriers, such as treatment cost or lack of transportation, that could hinder accessing treatment in a specialty setting. Measuring the number of people who were or were not able to access MAT requires measuring the scope of the problem, treatment providers available, and education of those providers.

*Strengths and Limitations*

Experts rated this measure construct as higher on importance, with wider distributions for scientific acceptability, feasibility, and usability (Figure 3.6). This measure construct aligns with a top HHS priority of addressing the opioid crisis, and RAND’s evidence table notes that the misuse of and addiction to opioids is a major public health crisis in the United States, causing over 1.6 million years of life lost annually. While experts considered this
measure construct to be important, they expressed a number of concerns about the readiness of this measure construct for international comparisons. First, experts noted that receiving a prescription for MAT does not necessarily translate to access. “You may be referred to MAT but may not be able to get it. You can wait months after a referral [to access treatment].” Experts also pointed out that there are good data available on the number of people who receive MAT, but the denominator is more difficult to capture. Experts questioned whether data were available on how many people have an OUD and whether that number is available internationally. One expert noted that “other countries are taking a different approach [to treating OUD], such as safe injection sites,” which may make this measure less comparable across countries. Although experts agreed that a lack of international data should not preclude pursuing this measure construct, they also noted that collecting this information is likely to require survey data. While RAND’s evidence table notes that many European countries and other OECD countries regularly field surveys like the National Survey on Drug Use and Health (NSDUH) that attempt to describe licit and illicit drug use in their populations, one expert noted that “as a rule, quality measures based on potentially criminal activity tend to be difficult to implement. And if the legal environment is different across countries, it may introduce new sources of bias.”

What Is Needed to Move This Measure Forward

For this measure construct to move forward, significant challenges with respect to the quality and comparability of data need to be addressed. Given the challenges identified with respect to referral versus receipt, experts agreed and re-rated this measure construct with the assumption that it would be a measure of receipt of MAT. However, even with that refinement, there still could be challenges with international comparisons; RAND’s evidence table noted that there are important differences in how receipt of MAT might be documented across countries because some have national prescription registries while others collect separate data describing reimbursement and prescription. Other countries have good prescription data for specific subpopulations (e.g., those covered by certain insurance programs) but lack data for other subpopulations. Challenges identifying the number of people with an OUD necessary to standardize the measure also remain and would need to be addressed.

Estimates of Administrative Complexity and Cost

Description Provided in Evidence Table

In the clinical setting, administration refers to the nonclinical activities that regulate or control the behavior of individuals in a health care organization, enabling them to implement policy decisions and achieve their goals. Administrative tasks must be carried out at all levels of the health care system to ensure high-quality health care. While more resources invested in
administration can be productive, especially high administrative complexity and cost might be a sign of wasteful spending that could be curtailed without worsening patient health outcomes.97

- **Administrative complexity** is the complicated workflows and fragmented financing procedures that can waste clinicians’ time and interfere with their caring for patients.98 Estimates can include, but are not limited to, time spent preparing paperwork (e.g., processing and billing of claims) and contacting payers.99

- **Administrative costs** are the costs to carry out administrative services that help ensure against illness and deliver medical care.100 Estimates can include, but are not limited to, transaction-related costs, benefits management, selling and marketing, and regulatory and compliance costs.100

**Strengths and Limitations**

Although the median score for importance of administrative complexity and cost was in the high range, several experts rated its importance as very low (Figure 3.7). Similarly, several experts provided low ratings for scientific acceptability, feasibility, and usability, although the median scores were in the moderate range. RAND’s evidence table provided recent estimates of the percentage of health spending dedicated to administration in the United States, which ranged from 8.3 to 25.3 percent,101,102 depending on how costs were calculated and what was included. Experts noted that figuring out how to compare this in a meaningful way across countries is challenging because each country’s health care system is set up differently. Health economist members of our panel noted that “the cost accounting to accurately capture administrative complexity is exceedingly difficult,” in part because “the basis for accounting and cost accounting systems differ among countries and between governments and private
institutions.” Experts also discussed whether labor force surveys and the number of people in administrative positions could be used as a proxy for administrative complexity. While some found this a feasible option, others noted that it was more important to assess whether the “people produced useful output” as opposed to using the number of staff alone.

Even though this is a challenging measure construct to develop from an economic standpoint, experts argued that perceived difficulty should not be the reason not to pursue development. “We know that we are not doing a strictly apples to apples measure, but starting to collect the data will . . . shed better light and improve comparability. It’s a big enough topic and controversial enough that it is important.”

**What Is Needed to Move This Measure Forward**

Given the inherent challenges with this measure construct, developing a measure that is feasible and valid for international comparisons will likely require significant development. It will be important to develop a clear definition of administrative complexity and costs that can be universally applied across countries and that addresses some of the challenges with different accounting and health financing systems. As starting points, in RAND’s evidence table, we note that the system of health accounts described by OECD provides an international framework for the definition, demarcation, and categorization of health expenditures. Under this system, transactions are classified with regard to who pays (financing), who provides (provision), and what the purpose is (function). In the United States, administrative costs have previously been subdivided into costs attributable to private insurers (e.g., claims processing, profits), the government (e.g., eligibility determination, revenue collection), providers (e.g., billing, collections, collection and reporting of quality data), and employers (e.g., choosing plans, designing benefits). However, a more fundamental problem is that specific accounting approaches differ between private and public insurers and between countries. One of our experts noted that although a measure of precise costs could be quite challenging for the above-mentioned reasons, a survey approach to capture some of this information might be a feasible alternative.

**Disadoption of Ineffective Medical Services**

*Description Provided in Evidence Table*

This measure construct, depending on its precise specification, could describe the rate at which the health care system identifies ineffective treatments and reduces their availability or utilization in the health care system. A disproven treatment is a treatment that has fallen out of favor due to evidence overwhelmingly demonstrating that it is more harmful than beneficial. In the literature, this is often referred to as a medical reversal. Medical reversals occur when new
evidence (better designed, controlled, or powered than predecessors) contradicts the current standard of care.103

**Strengths and Limitations**

Experts rated this measure concept higher on importance, but ratings were more spread out for scientific acceptability, feasibility, and usability (Figure 3.8). Experts voiced concern over the use of the term *ineffective* in this measure construct, noting that it is hard to define “services that do more harm than good.” One expert noted that this measure could be more palatable and actionable if the definition were flipped to specify “very low or no benefit” based on published evidence. Experts noted that Choosing Wisely, an initiative of the ABIM Foundation that seeks to advance a national dialogue on avoiding unnecessary medical tests, treatments, and procedures, has a series of reports that are “trying to find consensus around low-value services. Some may be inappropriate, some questionable, but the idea is that through claims you can identify situations where the service is likely to be of low value.”104 Another expert liked the idea of this measure and pointed out that, although it is difficult to grade the evidence of medical interventions, infrastructure to do so is in place. One expert pointed to the RAND appropriateness criteria developed in the 1980s, although this person noted that those determinations are made on an individual basis and that “this [measure construct] is very challenging to implement if it is defined based on the implications for individuals.” Another expert agreed that a lot of debate will be around individual-level decisionmaking but noted that a national-level measure may alleviate some of these concerns. “I like the framing of countries persuading physicians to avoid problematic treatments.”
What Is Needed to Move the Measure Forward

Experts cautioned that there will be opposition to a measure of this type from stakeholders who currently benefit from the continued implementation of low-value or ineffective medical services (e.g., providers, drug manufacturers), regardless of how it is operationalized. Although a change in the framing could be useful for increasing acceptance, experts asked themselves whether “the benefits are worth the costs in terms of the blowback [we may get].” Several experts pointed to a potential solution to reduce the contentiousness of this measure that involved identifying a limited set of low-value (or no-value) services that were not controversial and that the majority of physicians would agree should not be utilized: “If we select a group of [services] or stand up a group like the U.S. Preventive Services Task Force to identify them, this would address a lot and improve health care systems.” Although in some cases different groups might read or interpret data differently, it could be possible to restrict the procedures included in this measure to those with wide agreement. Ensuring clear delineation between the related concepts of negative clinical benefit and poor value for the money spent will be important in the operationalization of this measure moving forward.

Healthy Days at Home

Description Provided in Evidence Table

Healthy days at home (HDAH) is a population-based measure that attempts to quantify how well health care organizations keep people alive, healthy, and not utilizing health care services.\textsuperscript{105,106} This measure is one of a select few population-based measures of the health care system that attempts to quantify health, as opposed to the lack thereof.\textsuperscript{106} The measure is calculated by subtracting the total number of days out of the year that the individual was not alive or was utilizing acute or post-acute health care services from 365 days and then calculating the mean number of HDAH for the population of interest.\textsuperscript{106}

Strengths and Limitations

After discussion, experts remained divided in their perceptions of the importance, scientific acceptability, feasibility, and usability of this measure construct, resulting in median scores in the moderate range for all domains (Figure 3.9). This may have been due, in part, to experts interpreting this measure differently and lack of consensus around whether it was a useful measure of the health care system specifically. Some commented that they “do not see how the health care system has a strong connection to this [measure construct] compared to the public health system, SES [socioeconomic] characteristics, etc.” Another noted that to connect this
measure construct to the health care system “is a long stretch.” Others, however, noted that the “idea behind the measure is to indicate whether the health care system is effective at keeping patients at home, where they generally want to be.” Another expert commented that this measure construct can reflect “a reduction of unnecessary hospital days, a worthy goal that can be achieved. . . . It is an improvement on measures like readmissions or counting hospital days.” Experts also noted that a measure like this could also help to inform our understanding of transitions to home-based care and assess, to some extent, “how hospital versus community-based care stacks up.” Experts noted, however, that with this measure one would want to stratify patients according to the risks they are facing, which is difficult even within the United States and particularly challenging for international comparisons.

**What Is Needed to Move This Measure Forward**

A lot of the discussion of this measure centered around whether it was one that reflected health care system performance. Given that the objective of this project was to identify measures to strengthen international comparisons of health care system performance, this is a critical decision point that could shape next steps. One expert questioned whether this measure could be offered to OECD as part of a broader set of measures rather than ones specific to the health care system, although there was concern that it would be misinterpreted by the user as a measure of health care system performance. Experts noted, however, that there are a number of measures that capture similar ideas that might be useful to draw on if this measure construct
moves forward, including one recently developed by Burke et al. in conjunction with the Medicare Payment Advisory Commission.¹⁰⁶

Availability of Emergency Medical Services to Prevent Opioid Death

Description Provided in Evidence Table

Emergency medical services (EMS) are a potential mechanism for provision of naloxone, an opioid antagonist that reverses the effects of opioid overdose. Naloxone is increasingly being used by police officers, emergency medical technicians, and nonemergency first responders to reverse opioid overdoses.¹⁰⁷ EMS include out-of-hospital EMS practitioners, such as emergency medical technicians and paramedics (i.e., first responders); initial out-of-hospital treatment; and transport by air or by ground to a hospital.¹⁰⁸

Strengths and Limitations

Although some experts rated this measure construct high for importance and scientific acceptability, overall, this measure construct received low to moderate scores for importance, scientific acceptability, feasibility, and usability (Figure 3.10). Experts noted that, while they liked this measure construct in theory, they questioned its added value and its utility for international comparisons. One expert commented that a more useful measure construct would capture the availability of naloxone to prevent opioid death across several settings: “This is an important concept, but . . . why are we focusing on the EMS response?” Another noted that it was hard to interpret the measure without knowing more about the
Building on the discussion of international comparisons, one expert questioned whether this was a measure we wanted to propose, noting that “it’s a critically important issue in the U.S., but is [this] a similar issue around the world that would give us a peek into health care systems that we could learn from?” Even within the United States, one expert commented that this information might not be readily available: “The one data source I’m aware of is a proprietary source.” One expert felt that “measuring opioid consumption and naloxone consumption in other countries is probably as far as you will get.”

**What Is Needed to Move This Measure Forward**

Of the measures that were discussed by experts in the Phase 2 rating process, this measure construct received the lowest ratings. Should the decision be made to move this measure forward, it will be important to take into account that emergency services are structured differently in OECD countries. For example, in some countries emergency physicians who have the ability to administer naloxone are sent out to emergencies. RAND’s evidence table also discussed potential ways to operationalize this measure construct. They include

- the percentage of opioid overdoses that were reversed with naloxone administered by EMS
- the percentage of EMS personnel who are trained to administer and are equipped with naloxone
- the percentage of people who could be reached by EMS trained to administer and equipped with naloxone within a specified amount of time
- the volume of ED admissions for opioid overdose with transport to the ED via EMS.

**Conclusion**

The ten measures included in this chapter hold promise but require significant work before they are ready to be implemented for the purpose of comparing health care system performance across countries. Experts suggested ways to refine, focus, or split the measure constructs, providing a helpful direction for how they could be strengthened. Many of the perceived limitations of these measure constructs centered around three major concerns: (1) Measure constructs could be impacted by factors outside of the health care system, (2) data both within the United States and abroad might not be available, and (3) cultural or organizational differences across countries that shape where and how care is delivered could pose a barrier to meaningful comparison. If these measure constructs are to move forward, these and other measure-specific challenges will need to be overcome.
4. Measure Constructs That Were Not Discussed by Experts

The following measure constructs were rated by experts in Phase 2 but did not score highly enough to be discussed at the Phase 3 teleconference. Experts were given a chance to move any of these measure constructs to the Phase 3 agenda, but they felt that other measure constructs held more promise overall and did not request that these constructs be discussed. The results presented in this chapter, therefore, summarize the Phase 2 ratings and provide additional context for some of the perceived limitations offered by experts through comments provided as justifications for their ratings in the ExpertLens platform. Given that experts were not as supportive of these measure constructs, we do not include a discussion of what is needed to move each measure construct forward. Measure constructs include:

- clinician workforce who can prescribe medication-assisted treatment or naloxone
- access to opioid treatment centers
- the percentage of patients with a follow-up visit within four weeks of starting an opioid for chronic pain
- time to regulatory approval for new prescription drugs
- travel time to provider office
- health care spending in the last year of life
- spending on mental health (percentage of total health spending).

Clinician Workforce Who Can Prescribe Medication-Assisted Treatment or Naloxone

Description Provided in Evidence Table

MAT is the most effective medical treatment for OUD and other SUDs, including alcohol. Three medications have been approved by the U.S. Food and Drug Administration (FDA) to treat OUD: methadone, buprenorphine, and naltrexone. Additionally, the FDA has approved naloxone, an injectable or intranasal drug used to prevent opioid overdose. Historically, the primary treatment for OUD was methadone, which was only administered by accredited opioid treatment programs (OTPs). However, there have been and remain barriers to methadone treatment, including locations of OTPs, provider shortages, and treatment waitlists.

Strengths and Limitations

In Phase 2, experts provided a wide distribution of responses for importance, scientific acceptability, feasibility, and usability of this measure construct, resulting in median ratings in the moderate range for all domains (Figure 4.1). Experts noted that this measure aligns with the
HHS priority to address the opioid crisis in the United States, and data included in RAND’s evidence table suggest that only 20–40 percent of people with OUD receive MAT, with much of this treatment gap attributed to the workforce, specifically due to (1) insufficient prescriber training, (2) lack of institutional support, (3) poor care coordination, (4) provider stigma, (5) inadequate reimbursement, and (6) burdensome regulatory procedures. Despite the importance of this measure in the United States, several experts questioned whether this was an appropriate measure for international comparisons or health care system performance in countries where OUD may be less prevalent. As one expert noted, “The appropriate policy for the U.S. on this issue depends on the need in the U.S. and the cost-effectiveness of this [clinician workforce who can prescribe MAT] versus other approaches. International comparisons are not useful to this decisionmaking.” Another commented, “Given major differences in health professional licensure and credentialing across countries, [this measure construct] may be meaningless.” Experts also noted that counts of prescribers who have the ability to prescribe do not necessarily equate to access to MAT, which experts considered the more important measure construct, or the ability to meet the demand. One expert commented, for example, that it was unclear whether “higher rates actually lead to better substance use disorder outcomes.”

Access to Opioid Treatment Centers

*Description Provided in Evidence Table*

*Access to treatment facilities* (for SUDs, specifically OUD) is defined as the ability to
offer same-day or walk-in appointments in the outpatient setting and the ability to optimize resources, refer patients, and support them if the necessary level of care is not immediately available.113

**Strengths and Limitations**

Experts rated this measure construct high on importance given its alignment with the HHS priority to address the opioid crisis (Figure 4.2). RAND’s evidence table also noted that increasing access to medications for OUD treatment is widely acknowledged to be a key strategy for addressing the opioid epidemic114–116 and that outpatient treatment facilities are particularly important for the treatment of SUDs, since only 13 percent of individuals who receive treatment do so in a private physician’s office.117 Despite higher ratings for importance, expert ratings for scientific acceptability, feasibility, and usability were more varied.

Experts noted that this measure construct is “key to treating addiction,” but some experts cautioned that “physical access does not assure patients will use these facilities, due to information and financial barriers.” One expert noted that a “more valid measure may be the percent of patients receiving this treatment.”

Experts felt that this measure construct would benefit from increased specificity on what constitutes a treatment center. One commented, “It would be important to carefully define these centers and to capture information on capacity of the facility.” Other experts offered cautions around potential data sources, noting, “If data are based on subjective opinions about availability, then it . . . would be misleading. If it is data from objective, scientifically sound evidence sources, then it will be useful.” Some experts also questioned the value of this measure construct for international comparisons: “[It is] not clear how this would be implemented across countries with different regulatory systems and different
infrastructures for substance use disorders.” Despite these challenges, one expert commented that “if a meaningful measure could be developed, it could be a useful comparison across countries, much like we compare number of hospital beds across countries.”

The Percentage of Patients with a Follow-Up Visit Within Four Weeks of Starting an Opioid for Chronic Pain

Chronic pain is typically defined as pain that lasts more than three to six months and is present on most days or every day. In 2016, the CDC released its guidelines for prescribing opioids for chronic pain in the United States to encourage careful and selective use of opioid therapy for pain management. The CDC advised that “clinicians should evaluate benefits and harms within 1 to 4 weeks of starting opioid therapy for chronic pain or of dose escalation.” The CDC also developed quality improvement measures to incentivize providers to follow the new guidelines, including an assessment of the recommendation for follow-up within four weeks of starting opioid therapy for chronic pain in order to evaluate the benefits and harms of continued opioid therapy.

Strengths and Limitations

Experts had several concerns about this measure construct overall, which resulted in lower ratings (Figure 4.3). While experts acknowledged that addressing the opioid crisis is a top HHS priority, they noted that the evidence base for the timing of this particular visit was not very strong and that this measure “could be met without much impact on patient outcomes.” One expert noted that this measure “feels incomplete, as it does not tell you whether the patient continued with the opioid or not,” and others were skeptical about the usefulness of this measure if “follow-up visits
become ‘tick the box’ exercises, especially if telehealth visits count.” Experts noted that although this may be a useful measure “to see how compliant our clinicians are in providing appropriate patient follow-up . . . it is not clear that this recommendation is the same across all countries.” RAND’s evidence table raised additional considerations about data access for this measure construct, mainly that EHRs can differ in their ability to capture opioid prescriptions, and countries with EHRs that do not capture all of the available opioid medications could have biased results. If survey data are used to collect these data, a minimum standard for a follow-up visit (e.g., duration, content) should be established.

**Time to Regulatory Approval for New Prescription Drugs**

*Description Provided in Evidence Table*

When new prescription medications are discovered, it can take different amounts of time for the same biologic agent to be approved for sale in different nations. An international measure of time to regulatory approval could encourage nations or regulatory bodies with slower approval processes to consider streamlining them, particularly for medications that have demonstrable benefits, such as improved health outcomes or cost-effectiveness.

**Strengths and Limitations**

Experts rated this measure construct higher on feasibility for international comparisons, with more disagreement among experts on the importance, scientific acceptability, and usability of this measure construct; several experts rated it very high, and others rated it very low (Figure 4.4). In terms of importance, experts giving high
ratings explained that rapid approval was especially important for cancer patients and also during a pandemic. For example, one expert commented, “Obviously very important for cancer, where time is of the essence . . . and cancer drugs make up most new drugs.” Those giving importance a low rating responded that “faster is not always better” and pointed out that “countries may have differing priorities related to their economic situation, capacity to spend on new therapies, views of safety, epidemiology of disease, and alternative treatments.”

Experts on both ends of the rating spectrum recommended more investigation to clarify how increasing the pace of regulatory approval would alter benefits and risks. For example, one expert commented, “The validity of a ‘timing’ measure should be supported by evidence on the marginal difference in benefits and risk, which will clearly differ among newly introduced drugs and is likely to evolve over time as more data become available.”

There was relative consensus among experts that data collection would be feasible. In fact, one expert commented that the Foundation for Research on Equal Opportunity already collects similar data, which are incorporated into its World Index of Healthcare Innovation. However, multiple experts expressed concern that market characteristics would affect the usability of the measure construct, noting that “this [concept] reflects both corporate strategy and regulatory practices,” which would make results hard to interpret. Others were concerned that health care systems would have little ability to influence regulatory approval times: “Not sure how health care systems can change this per se. [It] seems like a reflection of our regulatory process via the FDA.”

**Travel Time to Provider Office**

*Description Provided in Evidence Table*

The burden of travel from a patient’s residence to his or her health care provider can be an important measure of access to preventive and treatment services. Travel time can be estimated directly (e.g., using geographic information systems [GIS], online routing websites, or self-report measures), but when data are unavailable, travel distance is often used as a proxy measure of travel time. However, although time and distance are highly correlated, they can differ depending on several factors (travel mode, time of day, travel speed, etc.).

*Strengths and Limitations*

In Phase 2, experts rated this measure construct in the moderate range across all domains (Figure 4.5). Experts commented that a measure construct quantifying travel time could play an important role in identifying access problems, particularly for underserved patients. However, one expert responded that it was unclear “how this aligns with HHS priorities.” Four different experts mentioned telehealth in relation to this concept. The overall implication of these
comments was that the importance of optimizing travel time is decreasing with the current expansion of telehealth.

Another common theme among experts was that travel time could be used as part of a set of measures characterizing barriers to access of care. For example, one expert wrote, “Travel time can be useful locally when included in a larger suite of measures of access.” Multiple experts reported that it would be necessary to clarify aspects of measurement methodology before they could comment on dimensions of this concept, such as scientific acceptability. For example, “Low travel time to providers is important in some settings and less important in others. [On the one hand,] short time to the emergency room during a heart attack is really important. By contrast, travel time to an oncologist for an initial cancer workup is much less important.” Another expert commented, “There are limitations to how we calculate distance and time traveled. The available tools to do this calculation have improved over time. A key issue is agreeing upon to what provider we are calculating a patient’s travel.” Similarly, there was concern that international comparability would be challenging given differences in transportation infrastructure, population density, and availability of telehealth. However, some experts were more optimistic that the measure could be acceptable, meaningful, and useful: “Data are more readily available than in the past. If technical challenges can be addressed, measures based on this concept can add useful information to other measures of access.”

![Figure 4.5. Travel Time to Provider Office](image)

NOTES. blue: median; gray: IQR
Health Care Spending in the Last Year of Life

*Description Provided in Evidence Table*

Health care spending in the last year of life is thought to be an indicator of wasteful spending on care for terminally ill patients with poor prognoses. After a patient’s death, the amount spent on that patient’s care can be calculated retrospectively. If the same calculation is performed for all patients who died in a one-year period, the total amount spent can be averaged over the number of persons who died.

*Strengths and Limitations*

Experts were split on all dimensions of this measure construct, with some scoring it very high and others scoring it very low (Figure 4.6). Several experts commented that the measure construct was well aligned with the HHS priority to advance value-based care: “Reducing unnecessary spending at the end of life would improve the value of care delivered.” However, a common concern was that peer-reviewed literature has questioned the validity of using health care spending at the end of life as a measure of health care system performance. Reasoning against the measure construct centered on the idea that death is difficult to predict prospectively, which means that much high-intensity spending is for care of patients who ultimately survive—sometimes for years after a period of high spending. Multiple experts cited literature opposing the measure construct, and one expert recommended that an improved end-of-life spending measure would “somehow need to account for the uncertainty of [prospectively] predicting whether it will be the last year of life.”
A criticism shared by several experts was that international differences would be driven by cultural variation among OECD nations rather than health care system performance. One expert commented, “Every person has a different value placed on care in end stages of life. It is simply wrong to put this together and then extrapolate to some sort of meaningful policy difference.” An additional criticism was that nations with high spending might experience consequential political pressure to decrease costs that would lead to harmful rationing of care.

In terms of feasibility, some experts highlighted the simplicity of the measure construct, saying that it is “easy to define and measure (in theory).” Others cautioned that data collection in the United States might be limited to Medicare beneficiaries, thereby compromising international comparability.

Spending on Mental Health
(Percentage of Total Health Spending)

_Description Provided in Evidence Table_

_National health spending_ describes both individual needs and population health as a whole. Health spending includes the consumption of health goods and services and includes all types of financing arrangements.\(^1\) _Percentage of total health spending_ is the ratio of health spending allocated to mental health services compared with total national health spending.\(^{125}\)
**Strengths and Limitations**

Experts were divided on all dimensions of this measure construct (Figure 4.7). Overall, the highest median rating was for importance. Experts giving high ratings for this dimension explained that mental health had historically been underprioritized: “[Mental health] is an important concept, especially in the U.S. due to longstanding differences in parity.” Experts giving low importance ratings commented that the directionality of the concept was unclear and that it was not clear whether we should be aiming for a high or a low percentage. One expert noted, “This measure punishes countries that have cost-efficient mental health systems.” Another expert felt that “absolute spending per capita is more important,” since the percentage of mental health spending is affected by the amount of total health spending in the country.

Multiple experts commented that a measure encouraging nations to increase spending might not lead to improved outcomes: “Not sure spending (inputs) is as important as outcomes (outputs).” One expert expressed concern that prioritizing mental health spending was “politically charged due to debate about efficacy of intervention.”

A common criticism of the measure construct was that definitions of mental health might vary among nations, leading to challenges with measurement and international comparability: “Organization, payment, workforce, and pricing differences are among several challenges.” One expert was optimistic that measurement could be harmonized because “we have had similar discussions around defining primary care spending.”

Suggestions for strengthening the measure construct included (1) using existing “specifications from the Substance Abuse and Mental Health Services Administration (SAMHSA) or OECD tools” to facilitate common definitions across nations and (2) reporting stratified results by income, race, and ethnicity to ensure validity.

**Conclusion**

The seven measure constructs included in this chapter scored the lowest, overall, in Phase 2 ratings. For some of the measure constructs, experts were split, with some providing strong support and others calling out what they considered to be critical flaws or challenges with the measure construct that would be difficult to overcome. Other measure constructs were not met with such strong opposition, but they also did not have a strong champion among experts. While the measure constructs included in this chapter hold merit and could provide useful information for strengthening the knowledge base, they require significant work before they can be used for international comparisons of health care system performance.
5. Recommended Modifications to Existing OECD Measures

Some of the measure constructs provided by experts in Phase 1 were closely related to currently reported OECD measures. In many cases, the expert suggestions indicate potential refinements or extensions to the existing OECD measures. In this chapter, we first provide a brief overview of OECD Health Statistics, then summarize suggestions for refining or extending those measures collected during Phase 1 of our expert elicitation process.

OECD Health Statistics

The OECD Health Statistics database includes the latest comparable data for the OECD indicators across the 36 OECD member countries, as well as candidate and partner countries. These indicators reflect differences across countries in

- health status—e.g., life expectancy, main causes of mortality, disease incidence
- risk factors for health—e.g., smoking, alcohol, obesity, opioid use
- access to care—e.g., the extent to which people can access needed services, population coverage
- quality and outcomes of care—e.g., patient safety, clinical effectiveness, prescribing practices
- health expenditure and financing—e.g., how much countries spend per person, prices paid
- health workforce—e.g., supply and remuneration of doctors and nurses
- health care activities—e.g., use and supply of hospital services, medical technologies
- pharmaceutical sector—e.g., pharmaceutical spending, research and development (R&D), use of generics
- aging and long-term care—e.g., demand for long term care, quality of dementia care
- patient-reported outcomes and experiences—e.g., preliminary results from limited countries in three areas: elective hip and knee replacement, breast cancer care, and mental health.

Experts provided suggestions for expanding and enhancing OECD measures in four of these areas: quality and outcomes of care (focusing on health outcomes for common diseases), health expenditure, health workforce, and access to care (focusing on waiting times). Experts also suggested allocation of social spending by age group and ways that equity could be incorporated into OECD measurement across these measurement topics.

Health Outcomes for Common Diseases

Health outcomes for the most prevalent and burdensome diseases, such as cancer, ischemic heart disease, stroke, and diabetes, are an important dimension of health care system
performance in OECD countries. OECD Health Statistics currently includes a variety of outcome measures for these diseases, including mortality rates, cancer incidence, five-year survival rates for selected types of cancer, breast cancer stage distribution, and 30-day mortality following hospitalization for acute myocardial infarction. Given the importance of these diseases in contributing to death and disability in OECD countries, continuing to expand upon these measures may be warranted. For cancer, reporting stage distribution and survival rates for additional types of cancer could be considered. A number of outcome measures, including such intermediate outcomes as blood glucose control, are routinely measured and reported in the United States and other countries and would expand the possibilities of international comparisons.

Health Expenditure

OECD reports a variety of measures of health expenditure, including expenditure by type of service and by source of financing. Health expenditure is reported in currency units (local and adjusted for purchasing power or exchange rates), per capita, and as a percentage of gross domestic product (GDP).

One participating expert suggested exploring denominators other than GDP. GDP is used as a proxy for income and is a strong predictor of national health spending. However, GDP is a measure of production and can systematically vary from measures of income or consumption. Alternatively, international comparisons could focus on health expenditure as a share of consumption as measured by existing measures of consumption (such as actual individual income) or measures of income (such as household net adjusted disposable income). The latter measure is reported as part of the OECD Better Life Index.\textsuperscript{126}

Another suggestion to improve international comparisons of health expenditure is to stratify by household income. This type of measure provides a measure of the burden of health care costs among population groups and would allow for international comparisons of the degree of redistribution of costs among groups. For example, Carman et al. reported that United States households in the lowest income quintile pay an average of 34 percent of income toward health care, compared with 16 percent of income among households in the highest income quintile.\textsuperscript{127}

Another suggestion was to increase transparency in accounting methods that can affect international comparisons of health expenditure and take additional steps to increase comparability where possible. Experts noted that accounting differences among countries, and between private- and public-sector organizations within countries, can profoundly affect the measurement of health expenditure. Specific areas of difference that were called out include the chart of accounts (an accounting term for the list of the revenues, expenses, assets, and liabilities that activities and sites included); measurement of amortized items, such as pensions, retiree health care costs, contingent liabilities, and depreciation expenses; use of accrual or encumbrance accounting; alignment with Generally Accepted Accounting Principles; and the
extent of debt financing, among others. OECD publishes health accounting guidance in the publication *A System of Health Accounts*, and countries submit details on their accounting methodology in a “note on data sources and comparability” that is linked from the country’s health expenditure data entry in OECD Health Statistics.17

**Health Workforce**

OECD currently reports on the number of health care workers in each country by various job categories, such as doctors, nurses, dentists, and pharmacists, among others. However, with limited exceptions (e.g., hospital employment, long-term care workers), the currently reported measures do not allow for analysis of employment by setting of care. Additional data on workers employed by setting would allow for greater understanding of how resources are distributed within health care systems and productivity.

A further type of stratification of employment statistics would be between clinical and nonclinical staff working in each setting. This would enable more-accurate estimates of the labor component of administration of health care.

OECD currently reports a measure of “total health and social employment.” One expert recommended breaking out the reporting of health care worker employment and social work employment because in the United States social work is mostly not funded through the health care system.

Another suggestion for health workforce statistics is to expand the reporting of remuneration of health care providers. OECD currently reports on remuneration of doctors and nurses, but data are missing for many countries. An increased focus on data completeness and expansion to different types of health professions would enable more-detailed analyses of the sources of differences in health expenditure among countries and over time.

**Waiting Times**

OECD Health Statistics currently includes waiting times for elective surgery and specialist appointments. In addition, a May 2020 OECD report provided more detail on waiting times in OECD countries. As noted in that report, waiting times are a persistent challenge in many countries and “are a reflection of the functioning of the health care system as a whole and provide an opportunity for policy makers to trigger changes to improve the appropriateness, responsiveness and efficiency in health service delivery.”128

The OECD data on waiting times for elective surgery and specialist appointments currently do not include data for the United States. Developing a data source that would allow for U.S. submissions of waiting time data would permit comparisons with other countries.

In addition, OECD should consider expansion of the types of waiting times included in OECD Health Statistics. The OECD report on waiting times noted that OECD countries increasingly measure waiting times for primary care, cancer care, and mental health services.
Diagnostic screening is another important type of service for which waiting times comparisons could be useful to policymakers.

Allocation of Social Spending by Age Group

OECD reports total social spending, which comprises cash benefits, direct in-kind provision of goods and services, and tax breaks with social purposes (to be considered “social,” the spending must involve either compulsory participation or redistribution across households). One expert recommended that the OECD report social spending in categories of benefits targeted to different age groups, such as programs for older people (e.g., pensions, home help) and family benefits (e.g., early childhood education, parental leave).

Equity

In all phases of this project, experts raised equity as an important dimension of the measure constructs proposed, rated, and discussed. Through the lens of measurement, equity can be operationalized through ensuring that data are available on race and ethnicity, gender, income, and other characteristics, such as residing in a rural or urban community. Encouraging countries to begin collecting demographic data on their populations that could be linked to existing measures could create a more nuanced understanding of the OECD measures within countries and help to elucidate important disparities and gaps in performance by country that might be masked when such subgroup analyses are not conducted. As such, experts noted that a stronger focus on equity and disparities is needed not only in the current OECD measures set but should also be considered in refinements to the measure constructs identified in this project.

Other Expert Recommendations

Experts included a number of measure construct nominations that are currently reported by OECD. Unlike the measure constructs discussed in the previous section, the experts did not provide recommendations for substantial changes to OECD methodology for these measure constructs. Although experts were not explicitly asked to identify existing OECD measures that they believe are more useful for international comparisons than others, the nominations might provide some insight as to which OECD measures could be highlighted in future comparisons. These nominations include

- opioid-related or drug-related death
- out-of-pocket expenses
- private health insurance coverage
- skipping or being unable to access prescribed medication due to cost, formularies
- availability of generics and biosimilar measures as share of unit volume and share of sales
- total health employees per 1,000 capita
- health care provider compensation and earnings
- percentage of potential years of life lost, separated by cause
- quality of outcomes of top mortality diseases, key chronic diseases
- average length of stay in acute care
- self-reported health
- total doctor visits per capita
- utilization of cancer screening
- ambulatory care–sensitive admissions
- acute care beds per 1,000 capita
- health care consumption, price, and utilization indexes
- obesity, calorie consumption
- alcohol use
- smoking
- patient satisfaction.

Critiques of OECD Measures

The background research provided to experts in Phase 1 included a number of critiques of OECD measures that were identified in the literature. Although OECD has worked to strengthen international comparison of quality metrics by improving the definition and availability of quality indicators and supporting information infrastructure to produce more-reliable indicators in an increasing number of countries, critiques of OECD measures for international comparisons remain. International comparisons using existing OECD measures, as well as efforts to develop new measures, should consider these critiques, which include

- inclusion of measures that are heavily influenced by nonclinical factors, such as life expectancy (e.g., impacted by accidents, homicides)\textsuperscript{129}
- use of survey data, which can vary significantly by country in terms of completeness and quality if not standardized\textsuperscript{15}
- challenges with harmonizing internationally recognized units of measurement, such as for spending: different currencies, fluctuating exchange rates, and different accounting methods\textsuperscript{15}
- differences in gestational age and/or birthweight cutoffs for reporting live births\textsuperscript{129–131}
- differences in coding practices, particularly due to variation in the number of diagnoses coded, the procedure classification system in use, and the presence or absence of diagnosis timing flags, which can impact the accuracy of indicators and potentially introduce bias\textsuperscript{132}
- differences in level of detail captured with respect to time of clinical events, which preclude accurate comparisons of outcomes within a specified time period (e.g., 48 hours)\textsuperscript{132}
Experts recommended dozens of measure constructs that required further development before they could be discussed and rated. In the sections that follow, we present groups of measure constructs as well as our recommendations for future directions in development.

Important Measure Constructs Requiring Significant Development

The ten measure constructs in this category are important but were deemed not specific enough to advance to the rating stage. Experts who suggested these concepts often mentioned that measurement would be difficult but still recommended panel consideration because of importance. Thus, these measure constructs could be good candidates for measurement development work by HHS or other federal agencies in the near future.

In Table 6.1, we list the measure constructs, the number of experts who suggested each concept, and our recommendations about future steps that can be taken to develop and strengthen each concept.

Table 6.1. Measure Constructs Requiring Significant Development

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Number of Experts Who Suggested This Measure Construct</th>
<th>Recommendations for Future Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Innovation around basic science, drugs, diagnostics, medical devices, health</td>
<td>10</td>
<td>Determine which aspect(s) of innovation should be prioritized (e.g., pace, impact, accessibility, equity, value, progress in important sectors).</td>
</tr>
<tr>
<td>information technology, and health care delivery</td>
<td></td>
<td>Determine how international measurement would be feasible and meaningful given that OECD member countries might specialize in specific sectors.</td>
</tr>
<tr>
<td>Regulations pertaining to generic and trade drug variants</td>
<td>1</td>
<td>Envision how an ideal regulatory environment would operate so that pharmaceutical R&amp;D would be productive while also ensuring that patients have access to essential medications.</td>
</tr>
<tr>
<td>Family caregiving</td>
<td>1</td>
<td>Identify and characterize models of home care used in OECD nations.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Compare family caregiving with professional home care services to determine how costs to the family, costs to society, and health outcomes differ.</td>
</tr>
<tr>
<td>Measure Construct</td>
<td>Number of Experts Who Suggested This Measure Construct</td>
<td>Recommendations for Future Development</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------------------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>Determine whether OECD nations consider family caregiving to be within the purview of the health care system or whether it is an external social service.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall survival for various cancers after standards of care change because of new innovations</td>
<td>1</td>
<td>Determine how to select impactful innovations that are relevant to cancer care across OECD nations.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Determine how to identify target time periods during which changes in survival could be detected.</td>
</tr>
<tr>
<td>Fraud</td>
<td>1</td>
<td>Determine which types of fraud are relevant and important across OECD (i.e., health care provider overcharging, patient sharing of identities to defraud insurance companies, patient claims for services they have not received).&lt;sup&gt;133&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Determine whether detection of fraud is feasible across OECD nations and whether measurement could be comparable.</td>
</tr>
<tr>
<td>Choice</td>
<td>1</td>
<td>Determine what type of choice should be prioritized (e.g., insurance product, insurance company, provider, health system).</td>
</tr>
<tr>
<td>Safety and quality of behavioral health services</td>
<td>1</td>
<td>Identify important aspects of behavioral health that would be relevant and important across OECD nations (e.g., screening for depression, adherence to medications, safety of prescribing, follow-up care after emergency visits, institutionally acquired infections).</td>
</tr>
<tr>
<td>Launched and diffusion of new technology</td>
<td>1</td>
<td>Determine how to select candidate technologies for assessment (e.g., high-impact technology, technology with high volume of sales, “orphan” status).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Determine how to measure launch (e.g., time between patent and regulatory approval, time between patent and availability for purchase).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Conceptualize ideal diffusion (e.g., availability to high-need patients; availability in all geographic areas).</td>
</tr>
<tr>
<td>Actual costs of medical care</td>
<td>1</td>
<td>Envision a strategy for measuring cost of care that would be comparable across OECD nations. For example, administer surveys to derive cost estimates for key clinical conditions or clinical episodes (myocardial infarction, appendicitis, ED visit for asthma). Determine whose costs should be captured (e.g., patient, system as a whole).</td>
</tr>
<tr>
<td>Quality of long-term care</td>
<td>1</td>
<td>Identify and prioritize measures of the quality of long-term care (e.g., avoidance of infections, licensed staffing hours per resident day, percentage of residents reporting moderate to severe pain).&lt;sup&gt;134&lt;/sup&gt;</td>
</tr>
</tbody>
</table>
Measure Constructs Related to Health System Preparedness, the Public Health System, and/or the COVID-19 Pandemic

Nine experts suggested measure constructs were related to health system preparedness for a disaster, pandemic, or similar threat. Four of these concepts referred specifically to COVID-19. Because the objective of this project was to identify measure constructs of health system performance, and the public health system can be considered to be separate, these were not discussed in the expert panel meetings. However, these public health measures were numerous and important enough that they should be considered for evaluation in a future evaluation activity.

Table 6.2. Measure Constructs Related to Health System Preparedness, Public Health, or COVID-19

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Number of Experts Who Suggested This Measure Construct</th>
<th>Recommendations for Future Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to respond to a pandemic</td>
<td>4</td>
<td>Determine which aspects of response should be prioritized and measured across OECD nations. Specific development options are numerous and include characterizing the federal stockpile, ability to distribute supplies from a federal stockpile, ability to expand the workforce, ability to expand hospital capacity, capability to develop and distribute testing for novel pathogens, and implementation of policies to spare morbidity and mortality.</td>
</tr>
<tr>
<td>Investment in public health preparedness</td>
<td>1</td>
<td>Determine which aspects of investment are most important across OECD nations (e.g., expenditure, development of national stockpile of critical supplies, numbers and expertise of public health workers).</td>
</tr>
<tr>
<td>Ability to provide vaccines</td>
<td>1</td>
<td>Develop an improved understanding of which barriers to provision of vaccines are important across OECD nations (e.g., vaccine development, production of large quantities, administration across the nation, administration in vulnerable areas, population vaccine refusal rate).</td>
</tr>
<tr>
<td>Number of public health workers</td>
<td>1</td>
<td>Determine how the roles of public health workers vary across OECD nations and which types of public health workers are most important to member nations (e.g., physician, nurse, researcher, contact tracer, environmental safety expert, infectious disease expert, injury prevention expert).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Develop an improved understanding of current data sources characterizing the public health workforce domestically and internationally.</td>
</tr>
</tbody>
</table>

135
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Number of Experts Who Suggested This Measure Construct</th>
<th>Recommendations for Future Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of ventilators, intensive care unit beds, respiratory therapists, and nursing home beds</td>
<td>1</td>
<td>Explore the availability of and develop new data sources domestically and internationally to characterize total supply and percentage of availability.</td>
</tr>
<tr>
<td>COVID-19 mortality rate</td>
<td>1</td>
<td>Determine how to calculate mortality rate in light of the occurrence of asymptomatic spread of infection, which makes choosing a denominator challenging. For example, mortality rate could be calculated among hospitalized patients or patients on a ventilator with confirmed infection. Alternatively, mortality rate could be calculated at the population level.</td>
</tr>
<tr>
<td>COVID-19 testing rate</td>
<td>1</td>
<td>Investigate how COVID-19 testing priorities vary among OECD nations (e.g., population-level testing, testing for symptomatic patients, testing for patients with known exposure, testing for health care workers, testing in nursing homes).</td>
</tr>
<tr>
<td>COVID-19 waiting time for testing and results</td>
<td>1</td>
<td>Determine how to measure waiting time for testing to reflect important policy goals (e.g., waiting time for testing at the population level, for symptomatic patients, for patients with exposure, for health care workers, for nursing home patients).</td>
</tr>
<tr>
<td>Number and proportion of COVID-19 infected frontline health care workers</td>
<td>1</td>
<td>Determine whether comparable data could be collected across OECD nations, given that testing strategies and capabilities vary.</td>
</tr>
</tbody>
</table>

Measure Constructs Not Clearly Related to Health Care System Performance or HHS Priorities

Four measure constructs reflected important concepts but were determined not to be directly related to health care system performance or HHS priority areas (the opioid crisis, health insurance reform, drug pricing, and value-based care).
### Table 6.3. Measure Constructs Not Clearly Related to Health Care System Performance or HHS Priorities

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Number of Experts Who Suggested This Measure Construct</th>
<th>Recommendations for Future Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gun violence</td>
<td>3</td>
<td>Determine how best to characterize the impact of gun violence (e.g., population-level mortality rates, years of lost life).</td>
</tr>
<tr>
<td>Access to licensed and unlicensed home health workers</td>
<td>1</td>
<td>Home health workers play a variety of roles, including providing health services and other types of support. Determine how the roles of home health workers vary across OECD nations and whether health outcomes are better when home health care is provided by licensed workers. Develop an improved understanding of current data sources characterizing the home health workforce domestically and internationally.</td>
</tr>
<tr>
<td>Age-adjusted dementia prevalence</td>
<td>1</td>
<td>This measure was excluded because it assesses disease prevalence, which is not likely to be a reflection of health care system performance unless there are cures or known prevention strategies. Future concept development could seek to characterize health outcomes for patients diagnosed with dementia (e.g., ability to perform activities of daily living).</td>
</tr>
<tr>
<td>Air pollution</td>
<td>1</td>
<td>Determine how to combine data on air quality with data on population density. Determine whether to use ambient particulate matter concentrations alone or whether to report additional air quality data.</td>
</tr>
</tbody>
</table>

**Measure Constructs for Which Experts and/or Federal Advisers Raised Questions About Directionality**

Two measure constructs were suggested but not discussed due to concerns by other experts or federal advisers about the directionality of the measure (i.e., it is not clear which direction is better). In Table 6.4, we list these measure constructs, the number of experts who suggested each concept, and the rationale for exclusion due to directionality, as well as recommendations for future development of a similar concept.
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Number of Experts Who Suggested This Measure Construct</th>
<th>Rationale for Exclusion and Recommendations for Future Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>People who went without care during the last year</td>
<td>1</td>
<td>This measure was excluded because people might forego care for a year by choice, which means that this measure construct might not accurately reflect health care system performance. Future concept development could seek to identify patients who require follow-up but fail to adhere (e.g., patients with prescriptions for antihypertensives who fail to fill prescriptions for a year).</td>
</tr>
<tr>
<td>Deaths at home versus deaths at a hospital</td>
<td>1</td>
<td>This measure was excluded because this concept would be driven by cultural or personal values rather than health care system performance, so it is possible that people in one nation would consider death in hospital to be preferred over death at home. Future concept development could focus on assessment of concordance of end-of-life beliefs with setting of death.</td>
</tr>
</tbody>
</table>
7. Recommendations and Conclusion

We identified a number of promising additions and improvements to measures for international health care system comparisons. In addition to recommendations provided above about steps needed to move each measure construct forward, a number of cross-cutting recommendations and key decision points emerged from this effort that could help to shape next steps.

**Pursue measure specification for promising measure constructs.** Experts rated eight of the measure constructs as having the most promise for international comparisons. Next steps will require additional development work to establish the operational definitions and specifications of these measure constructs to ensure that any measure developed is valid and feasible for international comparisons of health system performance. In many cases, experts pointed to existing measures that could be leveraged for this purpose.

**Develop guidelines for assessing health care system performance.** Some experts felt that if a measure construct could be influenced by factors outside of the health system (e.g., in the case of HDAH), it was less useful for international comparisons of health care system performance given differences in populations, cultural preferences, and social conditions across countries. Other experts noted, however, that there could be ways to account for these differences in analysis (e.g., using regression analysis) to isolate the effect of health system performance on the outcome of interest. This ambiguity was particularly challenging for measure constructs that were both highly relevant for the health system and influenced by external factors. This includes measure constructs, such as QALE, that are influenced by social factors, as well as the approval and diffusion of new drugs, which are shaped by regulatory guidelines and industry processes. Developing clear guidelines to assess the alignment of measure constructs with health system performance will help ensure that such decisions are being applied consistently and help determine which measure constructs that were flagged with this concern might warrant further development.

**Disentangle and prioritize related constructs of access, coverage, utilization, and outcomes.** Several measure constructs focused on issues of access or coverage (e.g., access to mental health providers), but experts were quick to point out that coverage for or access to services does not necessarily equate to appropriate utilization of those services or result in improvements in health outcomes. As a result, in some cases experts recommended strengthening the measure construct by focusing on utilization or the expected health outcomes. However, experts also noted that coverage of and access to specific types of health care providers, services, and treatments could be important intermediate outcomes that can shed valuable insight into key outcomes of interest. As the measure constructs undergo further
development work, it will be beneficial to provide additional specificity around these related
constructs to help ensure that what is being measured will reflect the intended health outcome.

**Invest in developing novel measure constructs.** As outlined in Chapter 6, experts provided
a number of ideas for measure constructs that are important but were deemed not specific enough
to advance to the rating stage. Experts who suggested these concepts often mentioned that
measurement would be difficult but still recommended panel consideration because of their
importance. One example is the idea of innovation, which was offered by ten experts in Phase 1.
Although experts nominated different areas for innovation (e.g., drugs, diagnostics, health
information technology), this and other proposed measure constructs could be good candidates
for measurement development work by HHS or other federal agencies in the near future.

**Offer refinements to existing OECD measures.** In Phase 1, experts nominated measure
concepts for consideration. Many of these were refinements to or expansions of existing OECD
measures, with suggestions for additional specification or recommendations to further parse out
the data by key subgroups that would facilitate subgroup analyses and more direct, and
appropriate, cross-country comparisons (see Chapter 5). Because these measures are already
being used by OECD in some form, offering these refinements for OECD consideration could
help to strengthen international comparisons of health system performance in the near term.

**Limitations**

All experts were based in the United States. Although many of the project experts do have
expertise in international comparisons of health care quality, they may not be in the best position
to judge how relevant, useful, and feasible a measure construct is for other countries. There could
be benefits to engaging an international audience in the future development of these measure
constructs.

Experts raised the limitation that our U.S.-centric approach to construct selection could result
in identification of measure constructs that are less important or relevant for other countries. For
example, the opioid crisis is a more significant concern in the United States and is not
necessarily relevant for international benchmarking. Where experts raised this as a concern in
their comments or discussion, it was noted.

**Conclusion**

Collectively, experts provided a wealth of information on potential measure constructs that
would strengthen international comparisons of health care system performance. Eight measure
constructs were identified as having the most promise, although all require additional work to
strengthen the measure specification and address the feasibility of data collection in an
international setting. Experts felt that many other measure constructs were important and held
promise but would require additional investment to develop. In the near term, experts pointed to
a number of opportunities to strengthen the existing OECD measures set, providing important
refinements to measure specifications and how data can be collected and presented to facilitate more appropriate and accurate international comparisons. These suggestions could be explored with OECD while the new measure constructs suggested by experts are developed into feasible, reliable, and valid measures suitable for international comparisons.
References


83. Expert-provided definition.


117. Wallack SS, Thomas CP, Martin TC, Chilingerian J, Reif S. Substance abuse treatment organizations as mediators of social policy: Slowing the adoption of a congressionally approved medication. J Behav Health Serv Res. 2010;37(1):64-78.
125. Vigo DV, Kestel D, Pendakur K, Thornicroft G, Atun R. Disease burden and government spending on mental, neurological, and substance use disorders, and self-harm: Cross-


Appendix A. Evidence Tables

The RAND research team developed the following evidence tables for experts to use and reference during their rating process. Experts were asked to draw on these findings, as well as their own expertise and experience, to provide their ratings. The evidence tables are organized by the order in which they appear in the report.

Measure constructs that hold the most promise for international comparisons are

- treatment and control of hypertension
- access to and coverage for telehealth
- quality-adjusted life expectancy
- insurance coverage for mental health, behavioral health, and substance abuse services
- receipt of preference-concordant end-of-life care
- care continuity or consistent provider
- access to mental health providers
- data transfer and interoperability.

Measure constructs that received lower or inconsistent ratings after expert discussion are

- self-reported pain
- access to primary palliative care
- prices for brand-name and generic drugs
- diffusion of and access to new prescription drugs
- avoidable emergency department use
- the percentage of patients with an opioid use disorder who were referred to or prescribed medication-assisted treatment
- estimates of administrative complexity and cost
- disadoption of ineffective medical services
- healthy days at home
- availability of emergency medical services to prevent opioid death.

Measure constructs that were not discussed by experts are

- clinician workforce who can prescribe medication-assisted treatment or naloxone
- access to opioid treatment centers
- the percentage of patients with a follow-up visit within four weeks of starting an opioid for chronic pain
- time to regulatory approval for new prescription drugs
- travel time to provider office
- health care spending in the last year of life
- spending on mental health (percentage of total health spending).
Hypertension, also called high blood pressure, is blood pressure that is higher than normal. The threshold for a hypertension diagnosis is defined as systolic blood pressure consistently at or greater than either 130 or 140 mm Hg and diastolic blood pressure consistently at or greater than either 80 or 90 mm Hg. The population of individuals tested and treated for hypertension typically includes adults aged 18 and up.

- **Treatment of hypertension:** There are two components of treatment that can be use in isolation or in combination: (1) antihypertensive medications and (2) lifestyle changes (e.g., physical activity and diet).
- **Control of hypertension:** Control is achieved when the individual's blood pressure measures are consistently less than 140 mm Hg (systolic) and less than 90 mm Hg (diastolic).

### Table A.1. Treatment and Control of Hypertension

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>Hypertension, also called high blood pressure, is blood pressure that is higher than normal. The threshold for a hypertension diagnosis is defined as systolic blood pressure consistently at or greater than either 130 or 140 mm Hg and diastolic blood pressure consistently at or greater than either 80 or 90 mm Hg. The population of individuals tested and treated for hypertension typically includes adults aged 18 and up.</td>
</tr>
<tr>
<td><strong>Importance: Health or social impact</strong></td>
<td>Population impact:</td>
</tr>
<tr>
<td>National prevalence: 108 million U.S. adults, or 45%, have hypertension. Of these, 24% have control of their hypertension. Many adults for whom hypertension medication is recommended are untreated.</td>
<td>Global prevalence: An estimated 1.13 billion people worldwide have hypertension, most (two-thirds) living in low- and middle-income countries. Of these, fewer than 20% have control of their hypertension.</td>
</tr>
<tr>
<td>Clinical impact:</td>
<td>Hypertension increases the risk of heart disease, kidney disease, and stroke. Hypertension is the leading preventable risk factor for premature death and disability worldwide.</td>
</tr>
<tr>
<td>Performance gap:</td>
<td>Adherence to prescribed lifestyle changes and medications is low.</td>
</tr>
<tr>
<td>Costs:</td>
<td>More deaths can be averted through improved treatment and control of hypertension than any other major risk factor.</td>
</tr>
<tr>
<td>Value-based care:</td>
<td>The HHS priority to transform the U.S. health care system from one that pays for procedures and sickness to one that pays for outcomes and health relates to the treatment and control of hypertension because of (1) the involvement of lifestyle changes, which involve counseling and discussion rather than billable procedures, in the treatment and control of hypertension and (2) the need for greater transparency around the price and quality of the range of antihypertensive medications.</td>
</tr>
<tr>
<td>Other related priorities:</td>
<td>Drug pricing: The HHS priority to lower the price of prescription drugs without discouraging innovation relates to the treatment and control of hypertension because of (1) the number of individuals diagnosed with hypertension and their need for antihypertensive prescription drugs and (2) the important role of antihypertensive medications in the treatment and control of hypertension.</td>
</tr>
<tr>
<td>Measure Construct</td>
<td>Treatment and Control of Hypertension</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------------</td>
</tr>
</tbody>
</table>
| **Validity**      | **Construct validity:**<br>• Definitional differences in the diagnosis and treatment of hypertension would need to be resolved first in order to design a measure that can accurately reflect international differences in health care system performance.  
**Discriminant validity:**<br>• Because the diagnosis of hypertension varies as a function of diagnosis thresholds\(^1\) and the frequency of patient-provider interactions,\(^{14}\) a measure of hypertension treatment and/or control could reflect health care system characteristics rather than performance.  
• Variation in treatment guidelines across countries\(^{15,16,17}\) could bias a measure of hypertension treatment and/or control against countries whose treatment guidelines are less aggressive.  
There are unmodifiable risk factors of hypertension, including age and family history.\(^{18}\) To the extent that countries’ demographic compositions are not accounted for, a measure of hypertension treatment and/or control might reflect international variation in these risk factors rather than health care system performance. |
| **Feasibility for international comparisons** | **Definitions:**<br>• There are discrepancies in the threshold for hypertension diagnosis both within and across high-income nations (e.g., 140/90 or 130/80).\(^1\)  
• There is international variation in diagnostic thresholds,\(^{14}\) the initiation of treatment (e.g., lifestyle changes\(^15\) only, medications\(^{16}\) only, both, or no explicit guidelines),\(^{17}\) and the focus on total cardiovascular risk rather than only blood pressure when initiating treatment.\(^{19}\)  
**Data availability:**<br>• Data describing hypertension treatment would need to be regularly updated, nationally representative survey data describing medication utilization and lifestyle changes. Often, these measures would be self-reported, with potential biases due to differences in the structure of the survey questions\(^{20}\) and cross-cultural differences in responses.\(^{21}\)  
Examples of appropriate surveys include the European Health Interview Survey, the Medical Expenditures Panel Survey, and the Canadian Community Health Survey.\(^{22}\)  
• Hypertension control could be described by administrative data that document those who have been diagnosed with hypertension and their recent blood pressure readings. Among the input sources for the OECD data, diagnosed diseases are typically documented by national disease surveillance systems.\(^{23}\) However, describing control of hypertension would require the integration of advanced health information technology systems with repeated measures of blood pressure. A comparable set of indicators are those involving cancer diagnoses and survival rates. The OECD source for internationally comparable cancer data is the International Agency for Research on Cancer.\(^{23}\) The cost to collect and unify data describing hypertension control similar to the internationally comparable cancer data would likely be high.  
**Data comparability:**<br>• Health system characteristics also play a role in the diagnosis, treatment, and control of hypertension. Because of variation in health care costs borne by the patient, the frequency of interactions with providers and, thus, the diagnosis, treatment, and control of
hypertension might relate to system characteristics rather than performance. • International definitional differences would also threaten the comparability of the data.

Usability: Current evidence of use

Current evidence of the use of a measure of hypertension treatment and/or control includes the following (not exhaustive):

- **U.S. National Quality Forum**: Measures the percentage of patients 18 to 85 years of age who had a diagnosis of hypertension and whose blood pressure was adequately controlled (less than 140/90) during the measurement year. • **UK National Institute for Health and Care Excellence**: Measures (1) the percentage of patients under 80 years old with hypertension in whom the last recorded blood pressure measure was 140/90 or less and (2) the percentage of patients with hypertension aged 16 to 74 years who scored "less than active" on their annual physical assessment.

Usability: Relationship to currently used OECD measures

Hypertension treatment and control is not among the current OECD measures.

Closest related measures:

- **Antihypertensives prescribed per unit population**: This measure describes the medication component of the treatment of hypertension. It does not describe the lifestyle change component of treatment, and it does not describe the successful control of hypertension. Because this measure describes part of hypertension treatment, its data would be useful. The data sources for this measure are national administrative databases describing claims or dispensed prescriptions. The United States does not have comparable data.
- **Hypertension hospital admission**: This measure is indicative of uncontrolled hypertension. It does not describe hypertension treatment. Because this measure is indicative of the opposite of control, the main potential use of the source data would be complementary. In the United States, the main data source for this measure is the National Hospital Discharge Survey, a stratified, probability-designed survey representing approximately 1% of hospitalizations.
- **Ischemic heart disease mortality**: This measure is another indicator of uncontrolled hypertension. It does not describe hypertension treatment. Because this measure is indicative of the opposite of control, the main potential use of the source data would be complementary. The data source for this measure is the World Health Organization Mortality Database describing mortalities and their causes.

References

5. Centers for Disease Control and Prevention. Hypertension Cascade: Hypertension Prevalence, Treatment and Control Estimates Among US Adults Aged 18 Years and Older Applying the Criteria from the
Treatment and Control of Hypertension

Access to and Coverage for Telehealth

Table A.2. Access to and Coverage for Telehealth

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Access to and Coverage for Telehealth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Access, defined as “the timely use of personal health services to achieve the best health outcomes,” includes four dimensions: coverage, services, timeliness, and workforce. Access could be operationalized as the proportion of patients in the population for whom telehealth services are available or the proportion of patients who have used telehealth services. Other relevant data points when considering access include (1) whether the required technology is available in the patient’s geographic location (e.g., broadband internet and/or cellular service), (2) the proportion of patients who have personal access to the required hardware (e.g., a computer or smartphone), (3) the number of or proportion of providers who offer telehealth services, and (4) the proportion of patients who are aware of and know how to contact telehealth providers. Coverage could be defined as the availability and cost of telehealth services within health insurance plans or covered health services. There is no widely used definition of telemedicine; however, the Health Resources and Services Administration (HRSA) of the U.S. Department of Health and Human Services (HHS) defines telehealth as “the use of electronic information and telecommunication technologies to support long-distance clinical health care, patient and professional health-related education, public health, and health administration. Technologies include video conferencing, the internet, store-and-forward imaging, streaming media, and terrestrial and wireless communications.” The World Health Organization (WHO) defines telehealth as “[t]he delivery of health care services, where distance is a critical factor, by all health care professionals using information and communication technologies for the exchange of valid information for diagnosis, treatment and prevention of disease and injuries, research and evaluation, and for the continuing education of health care providers, all in the interests of advancing the health of individuals and their communities.” HHS currently describes four modalities of telehealth: (1) live video interaction between a patient and provider, (2) store and forward (electronic transmission of videos and digital images to a provider in a different location), (3) remote patient monitoring (electronic transmission of personal health and medical data, such as blood pressure readings, from an individual in one location to a provider in another location), and (4) mobile health (mHealth, which includes smartphone apps intended to promote health and well-being). HRSA distinguishes between telehealth and telemedicine thusly: Telehealth “refers to a broader scope of remote health care services than telemedicine. While telemedicine refers specifically to remote clinical services, telehealth can refer to remote non-clinical services, such as provider training, administrative meetings, and continuing medical education, in addition to clinical services” (emphasis in original).</td>
</tr>
</tbody>
</table>

Access to and Coverage for Telehealth
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Access to and Coverage for Telehealth</th>
</tr>
</thead>
</table>
| Importance: Health or social impact | Population impact:  
|  | - Estimates of the proportion of the U.S. population with access to and/or coverage for telehealth, as well as the prevalence of actual use of telehealth services, are widely disparate, primarily because of inconsistencies in the definition of telehealth (e.g., which services are included, such as audio-only telephones; live video; asynchronous written communications, such as patient portals, emails, and text messages; and/or synchronous written communications, such as e-visits); in the definitions of access, coverage, and use; and in which population is studied (e.g., the entire population or a specific population, such as individuals who are covered by a specific type of health insurance). Additionally, reimbursement and regulations vary by state.  
|  | - In 2016, it was estimated that 61% of U.S. health care institutions and 40–50% of U.S. hospitals were using some form of telehealth.  
|  | - Estimates of global access and use are also widely disparate for the same reasons (inconsistent definitions of telehealth, access, and use). More than half of the 70 countries participating in the WHO survey indicated that their country had a specific national telehealth policy or a reference to telehealth in their national eHealth policy.  
|  | - The Organisation for Economic Co-operation and Development (OECD) reported that telehealth was available and used in the majority of OECD countries but few countries currently collect specific data on volume or rates of use, and telehealth services were a small proportion of total health care services provided. For example, “teleconsultations” were estimated to account for only 0.1–0.2% of face-to-face consultations in a number of countries, including Canada, Australia, Portugal, Argentina, and the United States.  
|  | Meaningful clinical impact: There is evidence that telehealth can be an effective way for patients to access health care providers for primary care, specialists, and mental health care. Telehealth has the potential to increase access to health care to people who reported being unable to make it to the doctor’s office due to barriers to access, such as waiting times, distance, and transportation (estimated to be between 11–65% of people across OECD countries). Patients report high satisfaction with telehealth.  
|  | Performance gap: There are many barriers to the use of telemedicine in OECD countries, including lack of funding and lack of clear reimbursement mechanisms, payment schemes that discourage telemedicine, lack of legislation pertaining to telemedicine, inadequate information and communications technology (ICT) infrastructure, issues with connectivity, access to broadband, coverage in rural areas, challenges with onboarding health care staff, and privacy and data security.  

During the COVID-19 pandemic, Medicare and Medicaid expanded coverage of telehealth services and lifted requirements that previously restricted telehealth services. The increased use of telehealth has the potential to prevent overcrowding and exposure while still allowing those who need medical care to consult with their health care providers. Telehealth also has the potential to aid with mental health concerns and isolation during the COVID-19 pandemic.
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Access to and Coverage for Telehealth</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Costs:</strong> OECD conducted expert interviews and a literature review and found that there was some agreement that telemedicine services can deliver more cost-effective health care and possibly provide cost savings, but the data are not sufficient to generalize and require constant updates to keep up with the changes in available telehealth services and technologies. Telehealth revenue was $9.6 billion in 2013 and was estimated to increase to $30 billion by 2020. Global: The global telemedicine market was valued at $45 billion in 2019 and expected to grow to $175 billion by 2026.</td>
<td></td>
</tr>
<tr>
<td><strong>Importance:</strong> Related U.S. Department of Health and Human Services (HHS) priority topics</td>
<td>Health insurance reform: Insurance coverage of telehealth has the potential to improve access to health care services. Telehealth coverage is not currently consistently included in health insurance plans, and, given evidence that telehealth is cost-effective, increasing telehealth coverage could increase access to health care and reduce overall health care costs. Other related priorities: Value-based care: Use of telehealth services falls under the HHS priority area of &quot;maximizing the promise of health IT [information technology].&quot; The HHS priority to transform our health care system from one that pays for procedures and sickness to one that pays for outcomes and health relates to access to/coverage for telehealth because of the potential for telehealth to increase access to and use of preventive health care services that can be delivered via telehealth more easily (versus procedures which cannot be done via telehealth). As a result, it could reduce the number of procedures that are necessary and reduce delays in seeking care.</td>
</tr>
<tr>
<td><strong>Validity</strong></td>
<td><strong>Construct validity:</strong> The definition of telehealth (specifically, which types of services count as telehealth) would need to be agreed on first in order to design a measure that can accurately reflect international differences in access to and coverage for telehealth. <strong>Internal validity:</strong> There is evidence to suggest that telehealth services are effective and provide the same quality of care as in-person services. <strong>Discriminant validity:</strong> Although insurance coverage of telehealth is generally straightforward, because access to telehealth requires access to high-speed internet and digital literacy, measures assessing access may be confounded by these other factors.</td>
</tr>
<tr>
<td><strong>Feasibility for international comparisons</strong></td>
<td><strong>Definitions:</strong> There is not a common or agreed-upon definition for telehealth, which encompasses a variety of services and modalities. Meaningful measures would likely need to define specific types of telehealth (e.g., teleradiology, telepathology, teledermatology, remote patient monitoring, telepsychiatry) or a specific service (interacting with a health care provider through a video call, using email to communicate with a provider, etc.), as well as account for differences in coverage of health services in general (e.g., universal, nationally funded coverage versus more-fragmented, private coverage) to accurately compare between countries. <strong>Data availability:</strong> Data on insurance coverage of various telehealth services in the United States are available and could be used to estimate the overall</td>
</tr>
<tr>
<td>Measure Construct</td>
<td>Access to and Coverage for Telehealth</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td></td>
<td>coverage in the United States. However, access more broadly would likely require survey data to capture reach and feasibility.</td>
</tr>
<tr>
<td></td>
<td>o Similarly, on a global level, high-level coverage data exist; however, information about access is more limited.</td>
</tr>
<tr>
<td></td>
<td>o If telehealth visits could be captured via electronic health records, administrative data could be used. However, survey data would be more likely to yield the information of interest.</td>
</tr>
<tr>
<td></td>
<td><strong>Subgroup analysis and equity:</strong> Data could be stratified by type of insurance coverage, age, health status, income, and rural location; these subgroup analyses would facilitate a comparison of how well telehealth visits are improving access to health care for individuals who already have problems with access and could face challenges with accessing telehealth services as well.</td>
</tr>
</tbody>
</table>

### Usability: Current evidence of use
- There is no currently accepted measure for access to or coverage for telehealth. However, other relevant data are available, such as total revenue generated by telehealth, the number of telemedicine visits per year in specific populations (e.g., Medicaid recipients), coverage for telehealth by different types of health insurance plans (public and private), and patient-reported surveys about use of various types of telehealth services.
- Most data collection pertaining to telehealth is focused on the effectiveness and cost-effectiveness of telehealth; there is less information about access to and coverage of telehealth services.
- This measurement concept does not appear to be widely used globally or nationally; however, there is discussion in the literature for a need to better understand and measure the structure, processes, and outcomes of telehealth. The National Quality Forum has developed several measurement concepts to inform future studies of various aspects of telehealth, including the effect on access to health care; the cost and financial impact on the patient, care team, health system or payer, and society; the experience of the patient, the care team, and the community; and effectiveness (system, clinical, operational, and technical).

### Usability: Relationship to currently used OECD measures
- Access to and coverage for telehealth do not exist in the OECD Health at a Glance series or OECD Health Statistics; however, there is interest in better understanding aspects of digital health, including telehealth.
- Additionally, OECD captures other related measures:
  - Household access to ICT at home (computer access, internet access, broadband internet access); ICT and internet use (use of a computer, use of the internet, internet activities [email, social networking, reading, telephone or video calling, buying and selling goods, finding information about goods and services, seeking health care information, online coursework, looking for a job, banking, etc.]); online interaction with government agencies (interacting with health services via public authorities’ websites); and ICT skills (use of email with attached files, posting of messages, transferring files, etc.). These would be helpful in gauging barriers to access to telehealth.
  - Population coverage for a core set of services (% of population): Given that telehealth may often be included in overall coverage, this measure may provide insight into data collection for coverage of telehealth.
  - Needs-adjusted probability of visiting a doctor (% of population aged 15+): This measure could be positively correlated with use of telehealth.
Geographic distribution of doctors: Countries within which there are greater disparities in physician density across localities suggest a greater potential need for and benefit of increasing telehealth services.

References

15. Uscher-Pines L, Mehrotra A. Analysis of teladoc use seems to indicate expanded access to care for patients without prior connection to a provider. *Health Aff (Millwood)*. 2014;33(2):258-264.
Quality-Adjusted Life Expectancy

Table A.3. Quality-Adjusted Life Expectancy

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Quality-Adjusted Life Expectancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td><em>Quality-adjusted life expectancy</em> (QALE) is a summary measure of population health that combines information on fatal and nonfatal health outcomes. Data describing (1) health-related quality of life (HRQoL) and (2) life expectancy (by characteristics of interest, such as age, sex, race, and ethnicity) are required to calculate QALE. QALE has been used to monitor changes in population health over time, compare population health across countries, investigate health inequalities, and quantify the benefits of health interventions in cost-effectiveness analysis.¹</td>
</tr>
</tbody>
</table>

**Importance: Health or social impact**

**Population impact:**

*National:*
In 2008, the QALE of a U.S. adult at age 18 years was 52.3 years.²

*Global:*
- Life expectancy has increased in all Organisation for Economic Co-operation and Development (OECD) countries over the last few decades, though gains have recently slowed. Meanwhile, the rates of chronic conditions, such as obesity and diabetes, have been rising, and these conditions can lead to nonfatal adverse health outcomes such as functional disabilities.³
- A 15-country analysis showed that the QALE at age 20 ranged from 33 years (males in Armenia) to 61 years (females in Japan).¹ The United States performed worse than other Western high-income countries in the data set, which the authors conjectured could be due to disparities in access to care, unhealthy diets, or other behavioral factors. However, these estimates were calculated using the UK value set, and estimates changed when different countries’ value sets were applied.

**Meaningful clinical impact:**

Individuals with particular health conditions can lose quality-adjusted life years (QALYs).⁴ Evidence of differential QALE loss across subgroups can identify groups most affected by specific health conditions and thus those who are in need of targeted clinical interventions.⁴ HRQoL alone has also been used in clinical research and practice and is deemed important to inform decisions about treatments.⁵

**Performance gap:**

QALE can be an important indicator of health systems’ performance because it can assess population health; when linked to health outcomes, it can help identify the drivers of changes in QALE over time, which is important for informed policymaking.⁶ ⁷

**Costs:**

- Cost per QALE gained or lost is used in cost-effectiveness evaluations of health interventions (e.g., cancer screening).⁸ ⁹ In the United States, improvements in key risk factors (smoking, motor vehicle accidents, and alcohol) from 1960 to 2010 were attributed to QALE gains of 1.82 years in adults at age 25, valued at $65,428 per person over the course of a person’s adult life.¹⁰ ¹¹
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Quality-Adjusted Life Expectancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>• The costs to society due to low performance in QALE are potentially great. For example, the loss of an average of 28 years of quality (and thus productive) years of life for the average Armenian person relative to the average Japanese person has significant economic implications. A quality-adjusted life year (QALY) is thought to be worth somewhere between $50,000 and $150,000. Thus, if we assume the worth of a QALY to be $100,000, Armenia is losing $8.4 trillion per year due to the difference in QALE relative to Japan (28 QALYs per person in Armenia multiplied by $100,000 per QALY per year multiplied by 3 million Armenians).</td>
<td></td>
</tr>
</tbody>
</table>

**Importance:**

**Related U.S. Department of Health and Human Services (HHS) priority topics**

**Value-based care**

The HHS Healthy People 2020 goals include increasing the quality of life and years of healthy life, both of which make up QALE. Given that value-based care focuses on patient values, which include increased life expectancy and improved HRQoL, assessing QALE is highly relevant to this HHS priority. Measuring HRQoL on a large scale could lead stakeholders to better understand cost-effectiveness of interventions, thus informing coverage decisions that can improve the value of care.

**Other related priorities**

**Opioid crisis:** Certain regions in the United States likely have worse QALEs due to disproportionate burdens of opioid addictions leading to disability and/or death. Opioid crisis hotspots might be identifiable due to subgroup analyses of QALEs at the regional levels or by demographic characteristics.

**Validity**

**Construct validity:** QALE, as a summary measure of population health for fatal and nonfatal health outcomes, is a measure that offers more insight into health than classic life expectancy. Comparing a nation’s life expectancy versus its QALE can help stakeholders understand the additional toll that morbidity takes on society. Moreover, comparing QALE for different subgroups within one nation (e.g., men versus women) can lend insight into how different subgroups experience outcomes disproportionately.

**Convergent validity:** Although no single global or national QALE measure exists, researchers have used a combination of data sources on quality of life and life expectancy. One of the most common measures of HRQoL is assessed by the EQ-5D population survey, which is used around the world; this measure has been standardized and validated. For example, the convergent validities of the EQ-5D 3-item (0.88–0.99) and 5-item (0.90–0.99) HRQoL scales are high and about equal. However, response bias in the measurement of self-reported data such as HRQoL could affect the validity of QALE.

**Internal validity:** As with life expectancy, QALE is dependent on social determinants of health, in addition to medical determinants. For example, violent crime could lead QALE to drop in countries with high crime rates.

**Discriminant validity:** For some purposes, such as cost-effectiveness analysis, measures of HRQoL must be based on utilities or preferences for health states that meet the conditions of welfare economics, which assumes that individuals seek to maximize utilities (preferences for particular outcomes) and that overall societal welfare is some function of these individual utilities. Additionally, HRQoL is subjective, and
Feasibility for international comparisons

Definitions: QALE is categorized as a summary measure of population health: a measure that combines information on fatal and nonfatal health outcomes. Calculation of QALE requires two sources of data: probability of death at various ages and data on HRQoL for people in the sample, as well as a value set, which is likely to vary across nations due to differences in sociocultural and demographic characteristics.\(^1\)

QALY is the unit of measurement of QALE.

Healthy life expectancy (HALE) is a similar measure to QALE, but it is dependent on a population's rankings of preferences for different health conditions (health utility index) rather than HRQoL scores. In samples of patients with multiple comorbidities, calculating HALE is challenging because the large number of combinations of comorbidities makes scoring the health utility index challenging.\(^{19}\)

Data availability:

National:
The United States does not routinely calculate or report QALE.\(^20\) However, QALE is often estimated by combining HRQoL data from the Behavioral Risk Factor Surveillance System\(^21\) and life table estimates from the National Center for Health Statistics national mortality files.\(^2,22\) Yet, to estimate QALE at various levels (e.g., national, state, local), expanded measurement of “preference weights” is needed—i.e., the data required to scale health outcomes into health-state utilities. Currently, only a handful of weights exist for the U.S. population, and this affects QALE estimates.\(^20\)

Other HRQoL measures are assessed in the National Health and Nutrition Examination Survey and the Medicare Health Outcome Survey. There is also a 14-item version called the Centers for Disease Control and Prevention HRQOL-14, though this is not routinely administered in national surveys.\(^23\)

Global

In terms of HRQoL data, the EQ-5D questionnaire is the most common globally. Population norm values using the EQ-5D have been established in many European countries, as well as the United States, Canada, the United Kingdom, China, Japan, Korea, Argentina, New Zealand, Thailand, and Zimbabwe.\(^24\) Life expectancy tables are available in most countries and are maintained by Eurostat and the World Health Organization.\(^25,26\) A report from Korea explains the step-by-step process taken to calculate QALE using data from Statistics Korea (probability of death at 5-year intervals by gender, year, educational status, and region from 1970 to 2014) and HRQoL data from the Korea National Health and Nutrition Examination Survey.\(^27\)

Data comparability:

Data on life expectancy have been used to compare health outcomes internationally for decades. Although they are limited in the setting of rapidly changing age-specific death rates (e.g., in time of war or epidemics), they are relatively comparable across nations. Data on
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Quality-Adjusted Life Expectancy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HRQoL</strong> could be less comparable for several reasons: self-ratings of quality of life can be systematically different across and within countries because of sociocultural differences and other characteristics, such as age distribution. For QALE in particular, the choice of value sets for quality of life scores may influence overall performance on QALE, which “may seriously affect cross country comparisons of health expectancy.”</td>
<td></td>
</tr>
<tr>
<td><strong>Subgroup analysis and equity:</strong></td>
<td></td>
</tr>
<tr>
<td>• Calculating QALEs for different subgroups has been used in multiple contexts to evaluate for disparities in health expectancy, including by age, sex, race/ethnicity, socioeconomic status, and geographic region.</td>
<td></td>
</tr>
<tr>
<td>• In the United States, geographic variations in QALE among adults have been observed, with some of the lowest estimates found in Southern states (e.g., Mississippi, Alabama, Kentucky, Oklahoma).</td>
<td></td>
</tr>
<tr>
<td>• In the United States, QALE is higher among White women (54.1 years in 2008), followed by White men (51.1), Black women (50.5), and Black men (46.1).</td>
<td></td>
</tr>
<tr>
<td>• An international study of 15 countries showed that QALE was generally higher in females than in males.</td>
<td></td>
</tr>
<tr>
<td><strong>Usability:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Current evidence of use</strong></td>
<td></td>
</tr>
<tr>
<td>QALE measures are used as a routine health measure for tracking population health, estimating lifetime health losses for people with a particular health condition compared with those without the condition and to compare the burden of disease for target populations, monitoring trends, and measuring health disparities.</td>
<td></td>
</tr>
<tr>
<td><strong>National:</strong></td>
<td></td>
</tr>
<tr>
<td>There is minimal evidence of QALE use within the United States outside of the academic studies reported above; however, components such as life expectancy and HRQoL are routinely collected by the federal government.</td>
<td></td>
</tr>
<tr>
<td><strong>Global:</strong></td>
<td></td>
</tr>
<tr>
<td>There is no evidence of routine reporting of QALE internationally. However, the World Health Organization (WHO) reports data on HALE (derived from all-cause years lost due to disability in the Global Burden of Disease study from 2016) and generates estimates at the regional and global levels. QALE has been used to assess inequities in health in several countries.</td>
<td></td>
</tr>
<tr>
<td><strong>Usability:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Relationship to currently used OECD measures</strong></td>
<td></td>
</tr>
<tr>
<td>QALE is not currently represented among OECD measures. However, OECD does routinely collect and report two related measures: life expectancy and self-rated health.</td>
<td></td>
</tr>
<tr>
<td>• <strong>Life expectancy:</strong> Life expectancy, which refers to years of life at birth, is one of two necessary components for calculating QALE. Collecting life table data (life expectancy across all ages) would be a necessary step toward calculating QALE at the national level.</td>
<td></td>
</tr>
<tr>
<td>• <strong>Self-rated health:</strong> Nations report survey-based measures of self-rated health. HRQoL is more detailed but is conceptually similar to self-rated health. A related measure collected by WHO and applied in analyses involving OECD countries is HALE at birth, which is defined as the average number of years that a person can expect to live in full health by taking</td>
<td></td>
</tr>
</tbody>
</table>
### Measure Construct | Quality-Adjusted Life Expectancy
---|---
| | into account years lived in less than full health due to disease and/or injury.\(^{35}\) HALE (previously termed *disability-free life expectancy* by WHO) is closely related to QALE, but it is not weighted for health-state preferences.\(^{20,36}\)

#### References
Insurance Coverage for Mental Health, Behavioral Health, and Substance Abuse Services

Table A.4. Insurance Coverage for Mental Health, Behavioral Health, and Substance Abuse Services

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Insurance Coverage for Mental Health, Behavioral Health, and Substance Abuse Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>The 2008 Mental Health Parity and Addiction Equity Act required large group health plans to place benefits for mental health problems on equal footing with physical illness. The Affordable Care Act required small-group health plans and individual plans sold on exchanges to cover mental health services at the same level as medical services. However, patient advocates complain that health plans are more restrictive with reimbursement for mental health.¹</td>
</tr>
<tr>
<td></td>
<td>Behavioral health conditions can include mental illnesses, such as anxiety disorders, major depression, schizophrenia, and posttraumatic stress disorder, and substance use disorders (SUDs), such as opioid addiction and alcoholism. People with behavioral health needs may benefit from a range of treatments, from outpatient counseling to prescription drugs or inpatient care.² Mental health, behavioral health, and substance abuse services can include care in inpatient and outpatient settings, residential treatment centers, day care facilities, detox centers, etc. Treatments can include pharmaceutical and psychological interventions.³</td>
</tr>
<tr>
<td></td>
<td>• Medicaid plays a key role in financing behavioral health care. In 2015, Medicaid covered 21% of adults with mental illness, 26% of adults with serious mental illness, and 17% of adults with SUDs.⁴</td>
</tr>
<tr>
<td></td>
<td>• There are several measures of insurance coverage to consider:</td>
</tr>
<tr>
<td></td>
<td>o Effective coverage: “People who need health services obtain them in a timely manner and at a level of quality necessary to obtain the desired effect and potential health gains.”⁵</td>
</tr>
<tr>
<td></td>
<td>o Universal health coverage: “Universal health coverage means all people receiving the health services they need, including health initiatives designed to promote better health (such as anti-tobacco policies), prevent illness (such as vaccinations), and to provide treatment, rehabilitation, and palliative care (such as end-of-life care) of sufficient quality to be effective while at the same time ensuring that the use of these services does not expose the user to financial hardship.”⁵</td>
</tr>
</tbody>
</table>

Importance: Health or social impact

<table>
<thead>
<tr>
<th>Importance: Health or social impact</th>
<th>Population impact:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• National incidence: Almost one in five American adults lives with a mental illness. Only 41% of people with mental disorders access any mental health services.⁷</td>
</tr>
<tr>
<td></td>
<td>• Global incidence: Mental, behavioral, and substance use disorders are prevalent throughout the world, with 970 million people experiencing a mental disorder and 175 million experiencing an SUD in 2017.⁸</td>
</tr>
<tr>
<td></td>
<td>• Approximately 14% of the global burden of disease can be attributed to these disorders; however, most people do not have access to the treatment they need, including 75% of those in several low-income countries.⁹</td>
</tr>
</tbody>
</table>
Insurance Coverage for Mental Health, Behavioral Health, and Substance Abuse Services

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Clinical impact:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Living in a Medicaid expansion state is associated with greater decline in cost-related access problems, especially for young adults. However, there are racial and ethnic disparities in increases in access attributable to insurance.</td>
</tr>
<tr>
<td></td>
<td>• More than 70,000 Americans died of overdoses in 2017, yet only 1% of total health care dollars went to treatment for SUDs.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Performance gap:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• There is no generally accepted estimate of the health care costs associated with untreated behavioral health conditions. Data on untreated mental health conditions and associated health costs come from multiple sources and are difficult to link. Not knowing how many people are suffering from a specific behavioral health condition limits the ability to measure coverage rates.</td>
</tr>
<tr>
<td>• Nearly three-quarters of 169 World Health Organization (WHO) member states reported that care and treatment of persons with severe mental health disorders were included in their national health insurance or reimbursement schemes; 27% indicated that this care was not included, and 19% indicated that these conditions were explicitly excluded from the plans.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Costs:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• National: Annual U.S. mental health spending is more than $200 billion. Insufficient coverage of mental health and addiction treatments is forcing people to go out of network and incur higher out-of-pocket costs.</td>
</tr>
<tr>
<td>• Global: According to WHO, “The cost of mental health problems in developed countries is estimated to be between 3% and 4% of GNP [gross national product]. However, mental disorders cost national economies several billion dollars, both in terms of expenditures incurred and loss of productivity. The average annual costs, including medical, pharmaceutical and disability costs, for employees with depression may be 4.2 times higher than those incurred by a typical beneficiary. However, the cost of treatment is often completely offset by a reduction in the number of days of absenteeism and productivity lost while at work.”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Importance: Related U.S. Department of Health and Human Services (HHS) priority topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health insurance reform:</td>
</tr>
<tr>
<td>The HHS priority to improve the availability and affordability of health insurance seeks to help all Americans access health care that meets individual needs and budgets. Mental and behavioral health services are considered essential health benefits. Expanding Medicaid access may reduce opioid-related hospitalizations owing to better management of SUDs in outpatient settings.</td>
</tr>
</tbody>
</table>

| Other related priorities: |
| Opioid crisis: Expanding Medicaid may reduce opioid-related hospitalizations owing to management of opioid use disorder in outpatient settings. |

<table>
<thead>
<tr>
<th>Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construct validity: Individual patient measures and census-level data have shown the face validity of this measure. To improve comparability and precision of the measure, a list of mental health, behavioral health,</td>
</tr>
<tr>
<td>Measure Construct</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>and substance use conditions and treatments would need to be included in any measurement definition.</td>
</tr>
</tbody>
</table>

**Internal validity:** There is evidence that those with insurance are more likely to receive care for mental health disorders through the medical system than those without.\(^4\) Insurance coverage gains improve mental health and substance use treatment rates.\(^7,12\)

**Discriminant validity:** Health insurance coverage is not a perfect indicator of access to services. Countries with similar rates of insurance coverage for mental or behavioral health and SUDs but differences in cost-sharing could also have very different access to mental health treatment. Measures should account for differences in coverage of health services in general (e.g., universal, nationally funded coverage versus more-fragmented, private coverage). Furthermore, access to services does not imply that the services are of high quality.\(^25\)

### Feasibility for international comparisons

**Definitions:**
- There is no commonly used or standard definition of insurance coverage of mental health, behavioral health, or substance abuse measures, but there are international comparisons of insurance coverage overall and for other conditions, which could be leveraged to create a metric for this measurement concept.
- Given the different types of insurance mechanisms across countries (e.g., universal, nationally funded coverage versus more-fragmented, private coverage), comparison of coverage could be more difficult, and the presence of coverage might not have the same implications for access given different payment structures and cost-sharing requirements (e.g., deductibles, copays).

**Data availability:**
- Behavioral health conditions often go undiagnosed or untreated, which limits the ability to measure unmet need for insurance coverage.\(^26\)
- Review of payer claims data identifies which mental and behavioral conditions are or are not covered by insurance.

**Data comparability:**
- Experts disagree which programs constitute behavioral health treatment, such as self-help programs like Alcoholics Anonymous.\(^15\)
- The United States currently captures data insurance via several sources: the National Survey of Drug Use and Health (Substance Abuse and Mental Health Services Administration [SAMHSA]),\(^3\) the Current Population Survey (Census Bureau), the Medical Expenditure Panel Survey (Agency for Healthcare Research and Quality), the National Health Interview Study (National Center for Health Statistics/Centers for Disease Control and Prevention), and the Survey of Income and Program Participation (Census Bureau).\(^27\)
- WHO has collected data on coverage of mental health services by country via survey; however, these are high-level data and do not account for differences within countries with fragmented health care payment systems.\(^16\)
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Insurance Coverage for Mental Health, Behavioral Health, and Substance Abuse Services</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subgroup analysis and equity:</strong></td>
<td>Given the racial and ethnic disparities in insurance coverage and mental health treatment(^{28}), this measure should be stratified to assess whether these disparities continue to exist.</td>
</tr>
<tr>
<td><strong>Usability:</strong></td>
<td>Several countries currently provide insurance coverage for mental and behavioral health services. Evidence includes the following (not exhaustive):</td>
</tr>
<tr>
<td><strong>Current evidence of use</strong></td>
<td></td>
</tr>
</tbody>
</table>
- **National:** Financial burden of mental health conditions is recorded in the Medical Expenditure Panel Survey.\(^{29}\)  
- Access to behavioral health services for insured and noninsured populations is recorded annually in the SAMHSA Behavioral Health Barometer.\(^{30}\)  
- **International:** WHO collects data on coverage of mental health services by country via survey.\(^{16}\)  
- Australia uses national hospital data, community mental health data, and outcomes measures to measure publicly funded coverage. There are limited data on utilization or coverage of private psychiatry.\(^{31}\)  
- England’s Improving Access to Psychological Therapies initiative supports 1.1 million people accessing psychological support annually, with the plan to increase to 1.9 million by 2024. Because England has a national health care system, primary and specialist mental health care is paid for by the state.\(^{31}\)  
- In Japan, patients who need psychiatric services can apply for a municipal system where copayments for outpatient and inpatient care are 10%, with a monthly threshold based on income. Statistics are reported in the 630 Survey, the Patient Survey, and the National Database of Medical Fee Statements.\(^{31}\) |
| **Usability:** | Insurance coverage for mental health is not among the current OECD measures. |
| **Relationship to currently used Organisation for Economic Co-operation and Development (OECD) measures** | Closest related measures:  
- OECD collects data from the European Health Interview Survey on self-reported disability and substance abuse.\(^{32}\)  
- OECD reports on personal costs of mental ill health on a country level but has not updated its data since 2018.\(^{33}\)  
- In the 2019 report *Health at a Glance*, there are two measures related to mental health: suicide rates and responses to the question “Does your health keep you from working full-time or limit your ability to do housework or other daily activities?” which are stratified by responses to the following question from the Commonwealth Fund 2016 International Health Policy Survey of Adults: “Have you ever been told by a doctor that you have depression, anxiety, or other mental health problems?”\(^{34}\) |

**References**


29. Xu WY, Retchin SM, Seiber EE, Li Y. Income-based disparities in financial burdens of medical spending under the Affordable Care Act in families with individuals having chronic conditions. *Inquiry.* 2019;56.


Receipt of Preference-Concordant End-of-Life Care

Table A.5. Receipt of Preference-Concordant End-of-Life Care

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Receipt of Preference-Concordant End-of-Life Care</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>Receipt of preference–concordant end-of-life care refers to the level of agreement between a patient’s desired end-of-life care and the actual end-of-life care received.¹ These preferences are supposed to be translated into actionable medical orders that are designed to accompany the patient wherever he or she resides and in the patient’s medical record.²</td>
</tr>
</tbody>
</table>

Definitions for the various components of this construct are as follows:

- **End-of-life care** is the support and medical care (e.g., palliative care) given to people who are near the end of life; this care can include palliative care, hospice care, and social support from family or religious communities.³ There is no standard definition for the end-of-life period, but it is usually the period preceding an individual’s natural death from a process that is unlikely to be arrested by medical care; some insurance plans operationalize this period based on specific criteria for life expectancy (e.g., 6 months or less).⁴

- **Preference for end-of-life care** is the care the patient desires and can be assessed by various methods such as questionnaire, postmortem chart review (e.g., do-not-resuscitate forms, advance directives, other standardized legal forms), medical charts, and in-person interviews with the patient or the patient’s designated proxy.¹ This can also include preferences for site of care.³ Dying patients’ preferences for life-sustaining treatment are often unidentified due to physical deterioration and cognitive impairment; thus, proxies are often used to make decisions about end-of-life care on their behalf.⁵

- **Receipt of end-of-life care** is the actual care provided and is usually confirmed using postmortem chart review.¹

- **Concordance** represents the level of agreement between the preferred and received end-of-life care. Some studies also assess agreement between nonpreferred care and the care that was received.¹ When a patient’s preferences are not known, surrogate decisions (e.g., from a doctor or other decisionmaker) and medical orders have been used to determine concordance.

**Importance: Health or social impact**

**Population impact:**

**National:**
- Among a national sample of bereaved family members and friends of deceased older adults (65 years or older), about 87.4% reported that the end-of-life care received was consistent with decedents’ preferences (when weighted, this represents about 4.1 million decedents whose desired and received end-of-life care aligned).⁶ When end-of-life care was inconsistent with the decedent’s preferences, informants were more likely to report poor quality of care in the last month of life (e.g., unmet needs for pain management, decisions were made without enough input from the decedent or family, the patient was not always treated with respect).

**Global:**
- In Singapore, among adults who completed advance care plans as part of a national program, concordance for preferred care was high (98%
• Among the types of care used at the end of life, palliative care has received increasing attention worldwide. The goal of palliative care at the end of life is to relieve suffering of patients and their families by placing attention to physical, psychosocial, and spiritual symptoms as death approaches. It should be noted, though, that palliative care can be implemented at any stage of disease and is not limited to the end of life. Per a 2014 World Health Organization (WHO) report, globally, an estimated 19 million adults each year need palliative care at the end of life, with the majority being older adults (69%). Low- and middle-income countries (LMICs) have the highest prevalence of adults in need of palliative care at the end of life (70% in middle-income countries [MICs] plus 8% in low-income countries [LICs]) compared with high-income countries (HICs; 22%), but HICs have the highest rate (475 adults per 100,000). Among children globally, about 1.2 million per year need palliative care at the end of life, with the vast majority living in LMICs (62.9% in MICs plus 35% in LICs) as opposed to HICs (2.1%). LICs have the highest rate: 129 children per 100,000.

Clinical impact:
• Lack of concordance between a patient’s preferred and received end-of-life care could contribute to receipt of unnecessary care, potentially reducing a patient’s quality of life in the end-of-life period.
• Furthermore, patients who do not receive concordant care can suffer from depression and their families could endure additional psychological burden—e.g., decisional conflict, anxiety, depression, and posttraumatic stress disorder—before and after the patient’s death.
• Among older adults who received end-of-life care that was not consistent with their preferences, a higher proportion died in a hospital (30.3%) compared with their own home (27.9%) or a nursing home (22.9%), which could reflect trends of increasing intensive care unit use in the last 30 days of life and hospital readmissions in the last 90 days of life among Medicare beneficiaries.
• Discussing expectations for a patient’s remaining length of life is associated with higher hospice enrollment.

Performance gap:
• Changes in concordance have been not studied longitudinally, potentially distorting the extent of concordance.
• Automating assessment of concordant end-of-life care, with a chart review to verify findings, could be a valuable screening tool to identify potential overtreatment and undertreatment.

Costs:
National:
• Chronic care management and intensive interventions during the last three years of life are tied to high aggregate medical spending, even though such interventions do not always align with patient preferences.
• Among Medicare beneficiaries, receiving end-of-life care that is not consistent with a patient’s preferences is associated with higher health care costs. About one-fourth of all Medicare spending goes to end-of-

### Receipt of Preference-Concordant End-of-Life Care

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Receipt of Preference-Concordant End-of-Life Care</th>
</tr>
</thead>
<tbody>
<tr>
<td>desired and received comfort measures) but low for place of care (65% desired and received care at home).</td>
<td></td>
</tr>
<tr>
<td>• Among the types of care used at the end of life, palliative care has received increasing attention worldwide. The goal of palliative care at the end of life is to relieve suffering of patients and their families by placing attention to physical, psychosocial, and spiritual symptoms as death approaches. It should be noted, though, that palliative care can be implemented at any stage of disease and is not limited to the end of life. Per a 2014 World Health Organization (WHO) report, globally, an estimated 19 million adults each year need palliative care at the end of life, with the majority being older adults (69%). Low- and middle-income countries (LMICs) have the highest prevalence of adults in need of palliative care at the end of life (70% in middle-income countries [MICs] plus 8% in low-income countries [LICs]) compared with high-income countries (HICs; 22%), but HICs have the highest rate (475 adults per 100,000). Among children globally, about 1.2 million per year need palliative care at the end of life, with the vast majority living in LMICs (62.9% in MICs plus 35% in LICs) as opposed to HICs (2.1%). LICs have the highest rate: 129 children per 100,000.</td>
<td></td>
</tr>
<tr>
<td>Clinical impact:</td>
<td></td>
</tr>
<tr>
<td>• Lack of concordance between a patient’s preferred and received end-of-life care could contribute to receipt of unnecessary care, potentially reducing a patient’s quality of life in the end-of-life period.</td>
<td></td>
</tr>
<tr>
<td>• Furthermore, patients who do not receive concordant care can suffer from depression and their families could endure additional psychological burden—e.g., decisional conflict, anxiety, depression, and posttraumatic stress disorder—before and after the patient’s death.</td>
<td></td>
</tr>
<tr>
<td>• Among older adults who received end-of-life care that was not consistent with their preferences, a higher proportion died in a hospital (30.3%) compared with their own home (27.9%) or a nursing home (22.9%), which could reflect trends of increasing intensive care unit use in the last 30 days of life and hospital readmissions in the last 90 days of life among Medicare beneficiaries.</td>
<td></td>
</tr>
<tr>
<td>• Discussing expectations for a patient’s remaining length of life is associated with higher hospice enrollment.</td>
<td></td>
</tr>
<tr>
<td>Performance gap:</td>
<td></td>
</tr>
<tr>
<td>• Changes in concordance have been not studied longitudinally, potentially distorting the extent of concordance.</td>
<td></td>
</tr>
<tr>
<td>• Automating assessment of concordant end-of-life care, with a chart review to verify findings, could be a valuable screening tool to identify potential overtreatment and undertreatment.</td>
<td></td>
</tr>
<tr>
<td>Costs:</td>
<td></td>
</tr>
<tr>
<td>National:</td>
<td></td>
</tr>
<tr>
<td>• Chronic care management and intensive interventions during the last three years of life are tied to high aggregate medical spending, even though such interventions do not always align with patient preferences.</td>
<td></td>
</tr>
<tr>
<td>• Among Medicare beneficiaries, receiving end-of-life care that is not consistent with a patient’s preferences is associated with higher health care costs. About one-fourth of all Medicare spending goes to end-of-</td>
<td></td>
</tr>
<tr>
<td>Measure Construct</td>
<td>Receipt of Preference-Concordant End-of-Life Care</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>Global:</td>
<td>The cost of providing care that is not consistent with patient preferences has not been estimated on a global scale. However, estimates for end-of-life care are available from multiple countries. For example, a nine-country study showed that medical spending per capita in the last 12 months of life is high, ranging from about $20,000 in Taiwan to $80,000 in the United States. However, medical care at the end of life (last 12 months) accounted for only 8 to 11% of overall medical spending in most countries. A national Canadian study showed that the mean health care cost in the last year of life among children (aged 1 month–19 years) was $78,332 (Canadian dollars), with the majority incurred in acute care settings (67%). Use of advance planning directives programs in the United States and Canada, in which facilitators assist patients in documenting their preferences for end-of-life care, have been shown to reduce costs of care from $1,031 to $64,827 U.S. dollars per patient, depending on study period and cost measurement.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Importance: Related U.S. Department of Health and Human Services priority topics</th>
<th>Value-based care:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaps in quality of care are linked to pain and symptom management, communication, and care planning. Interventions to develop structured discussions about goals of care can improve patient and family satisfaction. Value-based care incentives promote an environment that relies on keeping patients comfortable and at home when possible. Ensuring that end-of-life care decisions are concordant with patient preferences reflects shared decisionmaking, which is a key component of value-based care.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Validity</th>
<th>Convergent, predictive, and face validity:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall, there is limited reliability and validity of measures used to determine concordance of preferred and received end-of-life care. Comparisons of different methods of communicating end-of-life preferences are limited by the inconsistent use of valid and reliable measures across studies. Specific issues identified that can affect validity include response shift (i.e., when people’s values and choices change over time), a lack of specificity and consistency of documentation, and biases related to chart reviews. An example measure with available psychometric properties is the DES-10 scale, which assesses patient engagement in decisionmaking for cancer care. It has been found to have convergent and predictive validity as well as good face validity.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feasibility for international comparisons</th>
<th>Definitions:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Although there is general consensus that this construct refers to the level of agreement between a patient’s preferred and received end-of-life care, there is currently no standard definition or measure for this construct.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data availability:</th>
<th>National:</th>
</tr>
</thead>
<tbody>
<tr>
<td>The National Health and Aging Trends Study (NHATS)—a prospective nationally representative survey of Medicare beneficiaries aged 65 and</td>
<td></td>
</tr>
</tbody>
</table>
### Measure Construct

<table>
<thead>
<tr>
<th>Receipt of Preference-Concordant End-of-Life Care</th>
</tr>
</thead>
</table>
| older—collects data from participants and, after death, from informants knowledgeable about the participants’ experiences in the month before death. To determine concordance between preferred versus received end-of-life care, informants are asked whether any decisions were made about treatments in the last month of the participant’s life that the participant would not have wanted (a “yes” would indicate discordance).  
  - The Physician Orders for Life-Sustaining Treatment (POLST) form is becoming increasingly standardized to assess patient preferences for care, with almost all 50 states implementing the program.  
  - The Commonwealth Fund International Health Policy Survey assesses preference for end-of-life care among older adults (i.e., patients discuss their care wishes with their doctor or family or have a written plan in place). These data need to be linked with other data sources on receipt of care to determine concordance with preferences.  
  - As countries move to incorporate advance care directives in which patients can document their preferences for care, these documents can be valuable for assessing concordance between preferred and received end-of-life care.  

### Data comparability:

- The lack of theoretical and conceptual clarity regarding concordance limits data comparability.
- There are also various methods to assess end-of-life care preferences and care received, which makes comparing results across studies difficult.

### Subgroup analysis and equity:

- Among U.S. older adults, fewer non-Latino Black patients (10.4%) reported that end-of-life care received was inconsistent with the patient’s preferences, compared with non-Latino White patients (13.7%), although this difference is not statistically significant.
- Black patients tend to receive more life-prolonging end-of-life care than White patients. Black patients do not experience the same benefits of end-of-life discussions, and their directives are not followed as frequently, even when directives are in place.

### Usability: Current evidence of use

Current evidence of the use of measures for receipt of preference-concordant end-of-life care (not exhaustive):

### National:

- The NHATS provides a single measure of concordance of a patient’s preferred and received end-of-life care; however, the measure is based on reports from informants who knew the patient’s experiences in the month before death. It is unclear how accurate the informant’s knowledge of preferred care was to the patient’s actual preferences.
- Historical chart reviews found approximately 99% concordance between care preferences and care received. In cases that were not concordant, the designated decisionmaker was not immediately available, and a clinician made his or her own determination on how best to proceed.
- The National Quality Forum encourages the documentation of treatment preferences and advanced care plans, recommending measuring the "percentage of patients aged 65 years and older who
Receipt of Preference-Concordant End-of-Life Care

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Receipt of Preference-Concordant End-of-Life Care</th>
</tr>
</thead>
<tbody>
<tr>
<td>have an advance care plan or surrogate decision maker documented in the medical record or documentation in the medical record that an advance care plan was discussed, but the patient did not wish or was not able to name a surrogate decision maker or provide an advance care plan.</td>
<td></td>
</tr>
<tr>
<td>• The Home Health Care Consumer Assessment of Healthcare Providers and Systems (CAHPS) Survey uses a multi-item scale with several measures for elements of and quality of communication between home health provider and patient.</td>
<td></td>
</tr>
<tr>
<td>• The hospice PEACE project measures the “[p]ercentage of seriously ill patients receiving specialty palliative care in an acute hospital setting for more than one day or patients enrolled in hospice for more than seven days with chart documentation of preferences for life-sustaining treatments.”</td>
<td></td>
</tr>
<tr>
<td>• The Family Evaluation of Hospice Care is an after-death survey administered to bereaved caregivers of individuals who died while in hospice. An indicator on end-of-life patient and family preferences could be added to this instrument.</td>
<td></td>
</tr>
<tr>
<td>Global:</td>
<td></td>
</tr>
<tr>
<td>• The Commonwealth Fund International Health Policy Survey assesses preference for end-of-life care among older adults (i.e., patients discuss their care wishes with their doctor or family or have a written plan in place).</td>
<td></td>
</tr>
</tbody>
</table>

Usability: Relationship to currently used Organisation for Economic Co-operation and Development (OECD) measures

Receipt of preference-concordant end-of-life care is not among the current OECD measures. However, OECD’s conceptual framework points to “coping with end-of-life” as a key feature of health care needs related to continuous and integrated health management. Closest related measures:

- OECD measures the number of people in long-term care but not their preferences on types or location of care.

References


### Table A.6. Care Continuity or Consistent Provider

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Care Continuity or Consistent Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>The World Health Organization defines <em>continuity of care</em> as &quot;the degree to which a series of discrete health care events is experienced by people as coherent and interconnected over time and consistent with their health needs and preferences.&quot;¹</td>
</tr>
<tr>
<td></td>
<td>The American Academy of Family Physicians defines <em>continuity of care</em> as &quot;the process by which the patient and his/her physician-led care team are cooperatively involved in ongoing health care management toward the goal of high-quality, cost-effective medical care.&quot;²</td>
</tr>
<tr>
<td></td>
<td>The emphasis on continuity of care can vary across health care domains. In primary care, the emphasis is on having a consistent single provider for each patient. For mental health care, the emphasis is on a team of providers delivering care through a coordinated plan, which can extend beyond health to social services.³</td>
</tr>
<tr>
<td></td>
<td>Continuity of care can include the quality of care over time, which includes the patient’s experience of a continuous caring relationship with an identified health care professional (i.e., a consistent provider) as well as the providers’ delivery of seamless service through integration, coordination, and the sharing of information between different providers.⁴</td>
</tr>
<tr>
<td></td>
<td>Three ways of providing care continuity are (1) personal continuity (via an ongoing patient-provider relationship), (2) information continuity (present care takes into account past information), and (3) management continuity (management of a health condition is consistent but also flexible to a patient’s changing needs).³</td>
</tr>
<tr>
<td><strong>Importance: Health or social impact</strong></td>
<td><strong>Population impact:</strong></td>
</tr>
<tr>
<td></td>
<td><em>National:</em></td>
</tr>
<tr>
<td></td>
<td>• In 2012, about 15% of Americans reported having no usual source of health care. From 1996 to 2012, the percentage of people reporting a provider as their usual source of care declined (about 35% reported this in 2012), while the percentage reporting a facility as their usual source of care increased (about 40% reported this in 2012).⁵</td>
</tr>
<tr>
<td></td>
<td><em>Global:</em></td>
</tr>
<tr>
<td></td>
<td>• Across 11 high-income countries, there is a high prevalence of patient-reported continuity of care in the primary care setting. Around 80% reported that their regular doctor often or always knows their medical history, spends enough time with them, involves them as much as they want, and explains things well.⁶</td>
</tr>
<tr>
<td></td>
<td>• In a three-country study, general practitioners (GPs) in England and Wales reported lower scores on their ability to provide informational and management continuity of care than their counterparts in the Netherlands and the United States. However, they reported similar (high) scores for personal continuity of care.⁷</td>
</tr>
<tr>
<td></td>
<td>• Across six Latin American countries, continuity of care in the primary care setting was poor, about half of users reported not having a regular doctor, and a third reported that their doctor did not know their medical history.⁸ The prevalence of reporting “not having a regular doctor” was</td>
</tr>
</tbody>
</table>

Care Continuity or Consistent Provider

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Care Continuity or Consistent Provider</th>
</tr>
</thead>
</table>

particularly high in Brazil, Colombia, and El Salvador (each over 60%). The prevalence of “PC [primary care] doctor not knowing medical history” was highest in Brazil (57%).

Meaningful clinical impact:
- Continuity of care allows for better support of individuals and a better understanding of their conditions, which increases patient satisfaction, patient-clinician trust, and quality of communication.⁹–¹³
- A review of 18 studies found several studies reporting positive associations of continuity of care with decreased health utilization (e.g., hospitalization and emergency department [ED] visits) and increased patient satisfaction.¹⁴
- Among Medicare beneficiaries (aged 65 years and older), patient-reported continuity of care is associated with lower ED use, preventable hospitalization, and mortality.¹⁵
- Among U.S. infants and children (up to three years of age), lower continuity of care (e.g., lower index score for having a usual provider of care) was related to more ambulatory care–sensitive hospitalizations, more ambulatory sick visits, and lower odds of lead screening, with stronger associations found in children with chronic conditions.¹⁶
- High levels of continuity of care are associated with fewer admissions to hospitals for conditions requiring ambulatory care, fewer ED visits, and lower probability of poor primary care coordination.¹,¹⁷,¹⁸
- Continuity of care with a primary care provider, as assessed by several indexes, is related to decreased number and costs of ED visits but higher costs of medical consultation visits among adults (aged 19 and older).¹⁹
- A review of 22 studies from nine countries showed that most studies (82%) reported reductions in mortality (all-cause or disease-specific) with increased continuity of care from generalist or specialist providers.²⁰
- A review of 18 studies found several studies reporting a positive association between continuity of care and improved social functioning, but there were mixed findings for other mental health outcomes (e.g., hospitalization, symptom severity, service satisfaction).²¹

Performance gap:
- A European study involving five countries with different health care systems showed that, overall, countries with national health systems are more effective at providing care continuity than countries with regulated-market systems.²²
- Among 11 high-income countries (including the United States), measures of high continuity of care are related to better primary care coordination.⁶

Costs:
- Increased care continuity has been associated with lower costs.¹⁴,²³
- Among Medicare beneficiaries with chronic conditions, for every 0.1-unit increase in the continuity of care index, episode costs of care were 4.7% lower for congestive heart failure (CHF), 6.3% lower for chronic obstructive pulmonary disease (COPD), and 5.1% lower for type 2 diabetes.²⁴
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Care Continuity or Consistent Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td>Importance: Related U.S. Department of Health and Human Services priority topics</td>
<td>Value-based care: Care continuity is a part of care coordination and a measure of quality of care. Higher provider continuity of care is related to a 14% reduction in health care costs and fewer hospitalizations. The value associated with this cost reduction is $1,000 per beneficiary per year. Therefore, continuity of care measures could be useful for assessing provider quality and resource use under value-based payment systems like Medicare’s Quality Payment Program.</td>
</tr>
</tbody>
</table>

| Validity | Construct validity: Claims-based measures of care continuity have several limitations, including an inability to capture information such as interpersonal continuity between clinicians and patients and coordination activities that can occur outside of visits between clinicians or between clinicians and patients. Also, some continuity of care analyses examine continuity during episodes of care (e.g., CHF, COPD) and could exclude a large proportion of patients and claims data; results might therefore not be generalizable. The National Health and Health Services Use Questionnaire (NHHSUQ) patient-reported continuity of care scale has been validated against insurance claims–based continuity. For example, the patient-reported continuity of care scale was significantly associated with claims-based longitudinal (duration of provider-patient relationship) continuity. However, this scale was significantly associated with lower mortality in older adults, while claims-based continuity of care measures were not. Patient-reported and claims-based measures might capture distinct dimensions of care continuity. That is, claims data capture care provided over time, but the patient-reported responses might better reflect the quality of the patient-provider relationship. |

<p>| Feasibility for international comparisons | Definitions: There are several definitions and measures of care continuity; however, this measure does not appear to be collected at a national or international level. Data availability: National: In 2004, the NHHSUQ had a 13-item scale that assessed continuity of care from a primary care provider among Medicare beneficiaries aged 65 years and older. Specifically, the questionnaire captures the length of care from a usual provider and personal continuity of care (e.g., provider’s competence, behaviors, and people skills). Global: Many countries use the Usual Provider Continuity index to assess continuity of care. It describes the proportion of visits to the patient’s regular physician out of all visits to any provider. However, this measure does not take into account the total number, frequency, or sequence of visits. |</p>
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Care Continuity or Consistent Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• The Commonwealth Fund International Health Policy Survey collects care continuity data from primary care patients from 11 high-income countries (e.g., personal and information continuity).&lt;sup&gt;6,31&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>• An adapted version of the Commonwealth Fund International Health Policy Survey collects self-reported continuity of care measures in the primary care setting among nationally representative samples of adults in six Latin American countries (Brazil, Colombia, El Salvador, Jamaica, Mexico, and Panama).&lt;sup&gt;8&lt;/sup&gt;</td>
</tr>
<tr>
<td>Data comparability:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Across studies, researchers have used different continuity of care measures (e.g., indexes, surveys of patients or providers, and insurance claims data) and time frames to define continuity (e.g., from a single weekend to several years), making comparisons difficult.&lt;sup&gt;20,21&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>• Several measures of care continuity have been shown to be highly correlated; studies using different measures might still be comparable.&lt;sup&gt;27&lt;/sup&gt;</td>
</tr>
<tr>
<td>Subgroup analysis and equity:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Among GPs in the United States, younger GPs have more-positive attitudes toward continuity of care than older GPs, but attitudes do not differ by sex, full- or part-time status, or practice list size. However, in England and Wales, female GPs have more-positive attitudes toward continuity of care than male GPs.&lt;sup&gt;7&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>• Among adults with chronic diseases, racial/ethnic minorities (non-Latino Black, Latino, and Asian patients) are less likely to have a usual source of care than non-Latino White patients. Black patients and Latino patients are also more likely than non-Latino White patients to report a facility rather than a person or a person at the facility as their usual source of care. Non-Latino Black patients are less likely than White patients to report that their usual provider listens to them.&lt;sup&gt;32&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>• The benefits of continuity of care are especially pronounced among older populations and those with chronic care conditions.&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>Usability: Current evidence of use</td>
<td>National:</td>
</tr>
<tr>
<td></td>
<td>• The NHHSUQ has a 13-item scale that assesses continuity of care from a primary care provider among Medicare beneficiaries.&lt;sup&gt;29&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>• Self-reported usual source of care (none, provider, or facility) is also assessed in the nationally representative Medical Expenditure Panel Survey.&lt;sup&gt;5&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Global:</td>
</tr>
<tr>
<td></td>
<td>The Commonwealth Fund International Health Policy Survey collects care continuity data in 11 high-income countries and six Latin American countries.&lt;sup&gt;6,8&lt;/sup&gt;</td>
</tr>
<tr>
<td>Usability: Relationship to currently used OECD measures</td>
<td>This measure is not currently included in the measures studied by the Organisation for Economic Co-operation and Development (OECD); however, a recent policy brief on primary care discussed the importance of studying continuity of care.&lt;sup&gt;33&lt;/sup&gt;</td>
</tr>
</tbody>
</table>
|                   | OECD Health at a Glance does collect information on patient experiences of ambulatory care, including<sup>34</sup>:
|                   | • doctor spending enough time with patient during consultation |
|                   | • doctor providing easy-to-understand explanations |
|                   | • doctor involving patient in decisions about care and treatment. |
References


   [https://www.commonwealthfund.org/series/international-health-policy-surveys](https://www.commonwealthfund.org/series/international-health-policy-surveys)


### Table A.7. Access to Mental Health Providers

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Access to Mental Health Providers</th>
</tr>
</thead>
</table>
| **Description**   | This construct refers to the availability of mental health providers to meet population behavioral health needs and is one dimension of overall access to mental health care.\(^1\) However, which providers are counted as mental health providers can vary.\(^2\) 

*Access* is a complex construct that generally encompasses a patient’s ability to enter a health care system (usually one that is covered by insurance), identify a health care service that is geographically accessible, and find a provider whom the patient trusts and can communicate with.\(^3\)

*Mental health provider* broadly includes psychiatrists, psychologists, licensed clinical social workers, counselors, marriage and family therapists, and mental health providers who treat alcohol and other drug abuse.\(^4\) The Health Resources and Services Administration (HRSA) only includes those with a graduate degree.\(^5\) |

<table>
<thead>
<tr>
<th>Importance: Health or social impact</th>
<th>Population impact:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>National:</strong></td>
<td></td>
</tr>
<tr>
<td>• There are about 5,565 designated mental health professional shortage areas in the United States (i.e., areas that have at least 30,000 persons to 1 provider).(^5)</td>
<td></td>
</tr>
<tr>
<td>• More than half of counties in the United States have no psychiatrists. The majority of providers are in the northeastern states and some counties on the West Coast.(^6)</td>
<td></td>
</tr>
<tr>
<td>• From 2003 to 2013, the median number of psychiatrists dropped by 10%, which may explain why patients report poor access to mental health care.(^7)</td>
<td></td>
</tr>
<tr>
<td>• Access can be limited by geography or cost.(^8) Demand for providers has increased because of new treatment options, expanded diagnostic categories, Medicaid expansion, and reduced stigma for psychiatric conditions.(^7)</td>
<td></td>
</tr>
<tr>
<td><strong>Global:</strong></td>
<td></td>
</tr>
<tr>
<td>• The global median of mental health workers is about 9 per 100,000 population, ranging from 1.6 per 100,000 in low-income countries to 71.7 per 100,000 in high-income countries.(^9)</td>
<td></td>
</tr>
<tr>
<td>• Among Organisation for Economic Co-operation and Development (OECD) countries, a significantly higher proportion of people consult general practitioners for mental health problems compared with psychiatrists or psychologists.(^10)</td>
<td></td>
</tr>
<tr>
<td>• Distance to providers could aggravate the disparities in utilization between patients with higher or lower socioeconomic position.(^11)</td>
<td></td>
</tr>
<tr>
<td><strong>Clinical impact:</strong></td>
<td></td>
</tr>
<tr>
<td>• Provider shortages result in lower access to care, long waits for necessary treatment, and provider burnout.(^12)</td>
<td></td>
</tr>
<tr>
<td>• Low access to mental health providers can contribute to unmet mental health needs. In 2019, about 18% of U.S. adults (about 44 million) had a mental illness, but 56% received no treatment, and 20.6% of those</td>
<td></td>
</tr>
<tr>
<td>Measure Construct</td>
<td>Access to Mental Health Providers</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>- who sought treatment reported not being able to receive treatment; in addition, 8.7% of youths (12–17 years of age) (more than 2 million) suffer from severe depression, but only 25% receive some consistent treatment (7–25+ visits in a year).&lt;sup&gt;13&lt;/sup&gt;</td>
<td></td>
</tr>
</tbody>
</table>

**Performance gap:**
- Low access to mental health providers can result in individuals relying on emergency departments (ED), which are poorly equipped to address psychiatric care needs.<sup>14,15</sup> ED length of stay is considered a surrogate marker for crowding, and, compared with non-psychiatric patients, psychiatric patients have a longer ED length of stay when they are admitted for observation, transferred to another facility, or discharged.<sup>16</sup>
- Fewer mental health providers participate in insurance plan networks than primary care providers, which contributes to narrow mental health networks and limits access to mental health care.<sup>17</sup>
- Following Affordable Care Act implementation, insurance plans sold on the individual marketplace are required to offer coverage for mental health services,<sup>18</sup> but treatment rates remain low (only 45% in 2014).<sup>19</sup>
- Some insurance plans still do not offer mental health coverage. In 2019, 7.8% of U.S. youths had a private insurance plan that did not cover mental or emotional difficulties.<sup>20</sup>
- Measuring access to mental health providers is complicated because a substantial amount of mental health care is administered by professionals outside the health sector—e.g., criminal justice, education, and social services.<sup>21</sup>

**Costs:**

**National:**
- Psychiatrists receive lower in-network reimbursement (13–20% lower depending on the severity of the problem) than non-psychiatrist medical professionals for the same services,<sup>22</sup> which may contribute to their lower participation rates in insurance networks<sup>23</sup> and higher patient costs for out-of-network services.<sup>24</sup>
- Absenteeism and lost productivity at work from untreated mental health issues costs the United States $23 billion annually.<sup>25</sup>

**Global:**
- The total cost of mental health illness for OECD countries is between 3.5% and 4% of GDP.<sup>26</sup>
- In developing countries, less than 1% of the development assistance for health goes to mental health.<sup>27</sup>
- It is estimated that mental health disorders will cost the global economy $16 trillion (in U.S. dollars) by 2030, in part due to loss of productivity across the life course.<sup>28</sup>

**Importance:**

**Related U.S. Department of Health and Human Services priority topics**

**Value-based care:**
- Value-based payment (VBP) models have been shown to improve quality and access to behavioral health care, but many structural and policy barriers to VBP adoption remain.<sup>29</sup>
- Quality measurement and value-based systems that do not take rurality and travel distance to a provider into account could unfairly penalize clinicians and lead to greater health disparities.<sup>30</sup>

**Validity**

**Construct validity:**
- Mental health quality measures, such as access to providers, have not been refined and validated. The evidence base for quality measures, including access, needs to be expanded and strengthened.<sup>21,31</sup>
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Access to Mental Health Providers</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Underreporting by ministries of health has been cited as an issue for accurately assessing access to mental health providers, which can compromise construct validity. The Mental Health America indicator for mental health workforce availability uses data from the National Provider Identification data file of the County Health Rankings &amp; Roadmaps data source. But limitations to the data include that the list of providers might not be up to date (some might no longer be practicing or accepting new clients) and that some small providers might not have an identification number.</td>
<td></td>
</tr>
</tbody>
</table>

**Feasibility for international comparisons**

**Definitions:** In the United States and internationally, access to mental health providers is often measured as a ratio of available providers to the population in a specified area.

**Data availability:**

*National:*
- The HRSA collects data to identify designated health professional shortage areas for primary, dental, and mental health care to determine eligibility for government programs.
- State-level data on mental health workforce availability (ratio of county population to number of mental health providers) are estimated by Mental Health America.
- Mental health workforce data are available from professional associations and Occupational Employment Statistics.
- U.S. and international research can be conducted with physician and hospital claims data, physician registries, and census data.

*Global:*
- The World Health Organization’s (WHO’s) Mental Health Atlas questionnaire collects data on availability of mental health professionals across countries.

**Data comparability:**
- WHO’s Mental Health Atlas mental health workforce questions include a broad set of professionals trained in mental health, making it possible to compare across countries which specialties are responsible for mental health care.
- However, measuring access can be challenging because it is based on the organization of services by country.

**Subgroup analysis:**
- Despite suffering from lower rates of mental and behavioral health illnesses, low-income and racial and ethnic minority populations are disproportionately affected by treatment gaps.
- Compared with non-Latino White populations, racial/ethnic minorities (those who are non-Latino Black, Latino, or Asian) have lower access to mental health services, are less likely to receive needed care, and are more likely to receive poor quality of care. In 2014, mental health care access improved for racial/ethnic minorities but remained significantly lower compared with that of non-Latino White populations.
- Rural areas have more designated mental health professional shortage areas (3,238) than nonrural areas (1,865).

**Usability: Current evidence of use**
- Several countries collect data on access to mental health providers. Evidence includes the following (not exhaustive):
## Access to Mental Health Providers

### Measure Construct

**Access to Mental Health Providers**

**National:**
- HRSA collects data to identify designated health professional shortage areas for mental health care.  
- States are no longer required to use time and distance standards to ensure provider network adequacy. Researchers can measure access to care by combining the Centers for Medicare & Medicaid Services Medicare Provider Utilization and Payment Data with census data.

**Global:**
- The WHO Mental Health Atlas collects data on the availability of a range of mental health professionals per population in member countries. These data cannot be used for longitudinal analysis.

### Usability: Relationship to currently used OECD measures

Access to mental health providers is not among the current OECD measures.

**Closest related measures:**
- For mental health care, OECD measures continuity of care, coordination of care, treatment, and outcomes. Continuity of care could be extrapolated to estimate provider access, but additional data would be necessary.
- OECD also collects data on type of provider consulted for mental health problems: general practitioner, psychiatrist, and psychologist.

### References


### Data Transfer and Interoperability

#### Table A.8. Data Transfer and Interoperability

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Data Transfer and Interoperability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>The transfer and interoperability of health data involves the hardware and software that deal with the storage, retrieval, sharing, and use of health care information.¹ Health information technology systems aim to facilitate the transfer and interoperability of health information.² Appropriately exchanging health information among health care professionals and patients allows the doctors, nurses, pharmacists, and other providers to obtain vital medical information.³ The objectives of health information technology and exchange are to improve the speed, quality, safety, and cost of patient care.³,⁴ - <strong>Health data</strong>: The data that are exchanged through health information technology consist of private and secure electronic health records (EHRs), personal health records, and electronic prescribing records.² - <strong>Transfer/Interoperability</strong>: Transfer of health information can be performed electronically or via manual processes. There are also potentially many different information systems that health care providers or organizations could adopt. The transfer and interoperability of health information refers to the ability of health information to be exchanged between systems and used across systems.⁵</td>
</tr>
<tr>
<td><strong>Importance: Health or social impact</strong></td>
<td><strong>Population impact:</strong> - Recent systematic reviews of the health effects of health information exchanges that serve to transfer and interoperate health information data demonstrate improvements to physical, psychological, and care continuity outcomes by reducing errors and improving compliance.⁶ - <strong>Clinical impact:</strong> - Health information exchanges can lead to higher-quality care, greater patient engagement,⁷ and fewer medical errors.⁶ - Health information technology is associated with greater utilization of preventive care, such as vaccinations.⁸ - Health information technology can support personalizing care to improve patient outcomes, and when these data are made available for large, representative patient populations, they can support selection of cohorts for clinical trials.⁹ - However, poorly designed or overly complex interfaces can pose a challenge to users, impede clinical efficacy, and generate other unintended consequences.¹⁰,¹¹</td>
</tr>
</tbody>
</table>
| **Performance gap:** | - The transfer and interoperability of health information can improve compliance with clinical guidelines.⁶ - Health information exchange systems have been shown to improve timely communication and increase efficient provider-patient interactions.¹² - Health insurance exchanges reduce overindicated and contraindicated prescribing.¹³ - "Reflecting ongoing challenges with the interoperability of health IT [information technology], just over half of U.S. physicians reported being able to exchange patient clinical summaries, laboratory and diagnostic test results, and patient medication lists with physicians
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Data Transfer and Interoperability</th>
</tr>
</thead>
<tbody>
<tr>
<td>outside the practice. . . In contrast, the vast majority of physicians (72–93 percent) in the Netherlands, New Zealand, Norway, and Sweden reported having these abilities, although Canadian and German physicians rarely reported having them (12–33 percent).&quot;14</td>
<td></td>
</tr>
<tr>
<td>Costs:</td>
<td></td>
</tr>
<tr>
<td>National: A national health information exchange system in the United States has been estimated to cost $97 million initially and $41 million annually in maintenance.15 On the other hand, a national health information exchange has been estimated to provide $78 billion in value annually.16</td>
<td></td>
</tr>
<tr>
<td>Global: Systematic reviews of studies across the globe indicate that health information exchange systems lead to total cost savings by reducing the costs of laboratory tests, radiology, consultations, and admissions.17</td>
<td></td>
</tr>
<tr>
<td>Importance: Related U.S. Department of Health and Human Services (HHS) priority topics</td>
<td>Value-based care:</td>
</tr>
<tr>
<td>The HHS priority to transform our health care system from one that pays for procedures and sickness to one that pays for outcomes and health relates to data transfer and interoperability because of the gains in efficiency that data transfer and interoperability produces. Promoting interoperability is one way to maximize the promise of health information technology.</td>
<td></td>
</tr>
<tr>
<td>Validity</td>
<td>Construct validity:</td>
</tr>
<tr>
<td>The Organisation for Economic Co-operation and Development (OECD) pilot study7 measures (1) the utilization of EHRs; (2) whether laboratory, radiology, or other results are shared with outside organizations; (3) e-prescription utilization; and (4) telehealth capacities.</td>
<td></td>
</tr>
<tr>
<td>Internal validity:</td>
<td></td>
</tr>
<tr>
<td>Evidence indicates that health information exchanges that serve to transfer and improve the interoperability of health information data can lead to higher-quality care, greater patient engagement,7 fewer medical errors,6 greater utilization of preventive care,8 and other positive effects.13</td>
<td></td>
</tr>
<tr>
<td>However, any measure of data transfer and interoperability describes processes for which there are substitutes. There is somewhat-limited evidence surrounding the benefits of health information exchange,18 and some argue that additional evidence is needed to justify the costs of implementation.19</td>
<td></td>
</tr>
<tr>
<td>Discriminant validity:</td>
<td></td>
</tr>
<tr>
<td>Health system characteristics play an important role in the adoption and utilization of health information technologies to facilitate data transfer and interoperability. For instance, nations like the United Kingdom that have single-payer health care systems might avoid (or have already avoided) many of the complicating barriers to adoption of health information technology.20 International comparisons would reflect the ease or difficulty of implementation of the health information technologies borne of the health system structure.</td>
<td></td>
</tr>
<tr>
<td>Feasibility for international comparisons</td>
<td>Definitions:</td>
</tr>
<tr>
<td>There are no standardized measures of (1) how seamlessly data can be transferred between providers and (2) how interoperable a nation's health information systems are.</td>
<td></td>
</tr>
</tbody>
</table>

Data Transfer and Interoperability 115
There are several types of interoperability to consider. Technical interoperability ensures basic data exchange capabilities between systems. Syntactic interoperability ensures that the same format and data structures are used across systems. Semantic interoperability ensures that the meaning of medical concepts can be shared across systems. Organizational interoperability enables seamless provision of health care across institutions. There are many domains and subdomains that can be measured as part of this concept, including availability of electronic health information, method of exchange, usability, comprehensibility, quality of data content, care coordination, accessibility, application (computable, human use), and patient and caregiver engagement.

Other factors include coverage of EHRs, national systems, patient data access, minimum data sets, use of structured data and clinical terminology standards, and unique IDs for patients and providers in the EHRs.

Survey data would be required to assess the data transfer and interoperability capacities of a nation’s health care system. A recent OECD pilot study demonstrates the type of data that could be used to inform international comparisons. Part of the survey would be fielded among health care professionals, and part of the survey would be fielded among chief information officers and administrators. The survey would assess the availability and use of EHRs and health information exchange.

The 2019 Commonwealth Fund International Health Policy Survey of Primary Care Physicians was administered to nationally representative samples of practicing primary care doctors in Australia, Canada, France, Germany, the Netherlands, New Zealand, Norway, Sweden, Switzerland, the United Kingdom, and the United States and obtains information about use of health information technology.

Data are not currently available. New data would need to be collected, similar to the pilot survey data recently collected by OECD. Countries could have different requirements for electronic capture of clinical data, minimum data set requirements, etc.

The U.S. National Quality Forum measures the adoption of health information technology that facilitates data transfer/interoperability. The following current evidence of use is not exhaustive:

The current evidence of the use of data transfer and interoperability focuses on the adoption of EHRs. EHR adoption is one component of health information technology that facilitates data transfer/interoperability.
Due in part to the structure of their health care systems, the United Kingdom, the Netherlands, Australia, and New Zealand had near-universal adoption of EHRs by 2009.\textsuperscript{23} More recently, survey evidence indicates that all but two of the OECD countries reported use of EHRs by at least half of primary care physicians; many had rates above 75%.\textsuperscript{7}

Data transfer and interoperability is not among the current OECD measures. However, OECD has piloted a survey to assess the adoption and utilization of health information technology in its member nations.\textsuperscript{7}

**Closest related measures:**
- CT/MRI units and exams: These measures describe the availability and utilization of medical technologies. U.S. data for this method come from a survey of hospitals and other medical facilities.\textsuperscript{24}

**References**

Table A.9. Self-Reported Pain

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Self-Reported Pain</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td><em>Pain</em> is physical suffering or discomfort caused by illness or injury. Because pain is subjective, the existence and intensity of pain is often measured by patient self-reports. Elements of self-report could include relevant pain and physical, psychological, emotional, and social domains of functioning, as well as patient-reported outcomes and defined goals. Pain assessment can also occur through physical examination or behavioral observation where self-report is not possible.</td>
</tr>
</tbody>
</table>

**Importance: Health or social impact**

- **Adults.** In 2016, 50 million adults (20.4%) reported chronic pain, defined as pain on most days or every day in the past 6 months, and 19.6 million (8%) reported high-impact chronic pain, defined as chronic pain limiting life or work activities on most days or every day in the past 6 months.
- **Youth.** There is no general self-reported pain assessment of the national youth population, but parent-reported data from 2012 suggest that 15.4 million youth (26.6%) had a pain condition (dental, abdominal, headache, other muscle or bone, etc.) in the last 12 months.
- **Global incidence:** Globally, approximately 20% of adults report chronic pain (that lasts at least 3 months). In low- and middle-income countries (LMICs), the prevalence of self-reported chronic pain is 33% among adults in the general population, 56% in the general elderly population, and 35% in workers. A lower prevalence of reporting any pain has been found in Asian countries—e.g., China (6.2%) and Japan (4.4%).
- In LMICs, there is insufficient access to pain management solutions. In Organisation for Economic Co-operation and Development (OECD) countries, analgesic opioid availability has grown by almost 60% since the 2000s.

**Clinical impact:**

- Although pain is associated with a negative impact on health and related to poor quality of life, patients often underreport pain because they consider it a normal part of the aging process. Consequently, pain is often underassessed and undertreated. Chronic pain is associated with reduced physical performance (e.g., physical activity and daily activities), lower quality of life, and poorer sleep. Persistent prescriptions of opioids to treat chronic pain have been linked to abuse and addiction.

**Performance gap:**

- Empirical studies and review articles provide evidence that self-reported pain is the optimal means to evaluate pain among adults without cognitive impairment. Patient self-reported pain has been used in the clinical setting to track progress (e.g., controlling pain), evaluate treatment, and change the course of care.
- Uneven pain management can be viewed as an equity issue. Existing self-report scales should scrutinize system-level factors that engender disparities.
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Self-Reported Pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Costs:</td>
<td><strong>National:</strong> The annual economic costs of pain can be divided into the incremental costs of treating pain and indirect costs due to lower economic activity associated with lost wages. The total loss in the United States annually ranges from $560 billion to $635 billion.(^\text{18})</td>
</tr>
<tr>
<td>Importance:</td>
<td><strong>Opioid crisis:</strong> Prescribing opioid doses based solely on pain intensity is problematic due to the subjectivity of pain intensity. There is no research linking a specific dose to pain relief. This subjectivity can lead to misuse and abuse of opioid analgesics.(^\text{19}) In 2017, 17.4% of the U.S. population received one or more opioid prescriptions, with the average person receiving 3.4 prescriptions.(^\text{20}) The prescribing rate increased annually by 3% from 2006 to 2010 but decreased thereafter, reaching a rate of 58.5 prescriptions per 100 persons in 2017. In 2016, 11.5 million Americans reported misusing prescription opioids in the past year.(^\text{21})</td>
</tr>
<tr>
<td>Validity</td>
<td><strong>Construct validity:</strong> Many pain measures for adults have been validated in different cultures and languages (e.g., the McGill Pain Questionnaire and the Brief Pain Inventory [BPI]).(^\text{27,28}) For example, the BPI has satisfactory to good construct validity and criterion validity.(^\text{29}) To date, no single measure of pain for youth has been found to be appropriate for use with all types of pain or across the developmental age span.(^\text{30})</td>
</tr>
<tr>
<td></td>
<td><strong>Internal validity:</strong> Psychometric studies have found self-reported pain measures to be internally consistent.(^\text{31})</td>
</tr>
<tr>
<td>Feasibility for international comparisons</td>
<td><strong>Definitions:</strong> Self-reported pain is a patient-reported outcome. Measures of self-reported pain are available for different populations (e.g., age- or disease-specific). Self-reported pain assessments are mainly appropriate for populations who can communicate their pain verbally, in writing, or by other means (e.g., finger span or blinking of their eyes to yes/no questions). Alternative tools (e.g., behavioral observation) may be more appropriate for noncommunicative populations, such as elderly people with advanced dementia, infants and preverbal toddlers, and intubated or unconscious patients.(^\text{3,4}) In two reviews of multicountry studies of chronic pain,(^\text{24,9}) there was great heterogeneity in pain measures used across studies. For example, although the International Association for the Study of Pain defines <em>chronic pain</em> as pain that lasts or recurs for longer than 3 months, some studies include intensity in the definition; others add complexity to the definition by stating that pain should be present for the whole day over a specific period (e.g., the preceding week or month).(^\text{24}) There are other chronic pain measures that identify pain as occurring any time during the last 6 months, which results in a significantly higher prevalence than studies using the definition of recurring pain that lasts 3 months.(^\text{24})</td>
</tr>
<tr>
<td></td>
<td><strong>Data availability:</strong> The U.S. National Health Interview Survey assesses pain in the adult population and pain-related conditions among youth.(^\text{5,6}) Although there is currently no global measure of pain, the National Health and Wellness Survey measures self-reported pain among adults in emerging and developed countries.(^\text{8})</td>
</tr>
</tbody>
</table>
Self-Reported Pain

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Self-Reported Pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Because self-reported pain is a complex and multidimensional experience, a single measure (e.g., intensity) could be insufficient to inform treatment.</td>
<td></td>
</tr>
<tr>
<td><strong>Data comparability:</strong></td>
<td>• Several self-reported pain scales have been validated in multiple languages and can be compared and aggregated in multicountry studies.</td>
</tr>
<tr>
<td><strong>Subgroup analysis and equity:</strong></td>
<td>• In the United States, chronic pain is higher in older adults. Furthermore, age-adjusted prevalence of chronic pain and high-impact chronic pain is higher in women, previously employed adults, adults living in or near poverty, and rural residents but lower in adults with at least a bachelor’s degree compared with all other education levels. Other countries also report higher pain in women and older adults.</td>
</tr>
<tr>
<td><strong>Usability:</strong></td>
<td><strong>Current evidence of use</strong> Several scales for self-reported pain are currently in use. Evidence includes the following (not exhaustive):</td>
</tr>
<tr>
<td>• The U.S. National Health Interview Survey assesses frequency of (chronic) pain in the past 6 months and “bothersomeness” of pain among adults aged 18 years and older.</td>
<td></td>
</tr>
<tr>
<td>• The Numerical Rating Scale (NRS) is one of the most widely used scales to assess self-reported pain in children and is considered the gold standard. It is an 11-, 21-, or 101-point scale in which the end points are the extremes of “no pain” and “pain as bad as it can be.” It emphasizes pain intensity and can be presented verbally or visually. Because it can be administered verbally, it can be used in telephone interviews.</td>
<td></td>
</tr>
<tr>
<td>• The Iowa Pain Thermometer (IPT) is a modified verbal numeric rating scale from 0–12 consisting of seven pain descriptors describing different levels of pain intensity plus response options between scores. It can be used with older adults, including those with cognitive impairment. Initially used in research, it can also be used in clinical settings; however, its 13-point scale does not align with other metrics commonly used in health care facilities.</td>
<td></td>
</tr>
<tr>
<td>• The Verbal Rating Scale (VRS) uses a series of adjectives to denote pain intensities. The most commonly used words are “no pain,” “mild pain,” “moderate pain,” and “severe or intense pain.” For ease of recording, each term can be assigned a number.</td>
<td></td>
</tr>
<tr>
<td>• The Visual Analogue Scale (VAS) is presented as a 10-cm line anchored by verbal pain descriptors, such as “no pain” and “worst pain imaginable.” It must be administered on paper or electronically. Users should orient it vertically or horizontally, depending on the language read by the audience.</td>
<td></td>
</tr>
<tr>
<td>• The Short-Form BPI was originally conceived to measure pain in cancer patients. It is also referred to as the 11-point NRS. It asks patients whether they are experiencing pain, its location, and its intensity. The BPI has been validated in 15 languages and fulfills many of the U.S. Food and Drug Administration recommendations for pain as a patient-reported outcome endpoint.</td>
<td></td>
</tr>
<tr>
<td><strong>Global:</strong></td>
<td>• The annual National Health and Wellness Survey assesses general pain among patients in the last 12 months; data are collected in Brazil,</td>
</tr>
</tbody>
</table>
Self-Reported Pain

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Self-Reported Pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>China, France, Germany, Italy, Japan, Russia, Spain, the United Kingdom, and the United States.8,39</td>
<td></td>
</tr>
<tr>
<td><strong>Usability: Relationship to currently used OECD measures</strong></td>
<td>Self-reported pain is not among the current OECD measures.</td>
</tr>
<tr>
<td><strong>Closest related measures:</strong></td>
<td></td>
</tr>
<tr>
<td>• OECD is planning to implement the first international Patient-Reported Indicators Survey (PaRIS) of Patients with Chronic Conditions,40 which will include Patient-Reported Outcome Measures (PROMS) such as pain, quality of life, physical functioning, and psychological well-being. Implementation of the survey in all participating countries will take place from 2021 to 2023.</td>
<td></td>
</tr>
<tr>
<td>• The World Health Organization collects data on self-reported pain in the context of other diseases, such as cancer.41</td>
<td></td>
</tr>
</tbody>
</table>

References

40. OECD. Putting People at the Centre of Health Care: PaRIS Survey of Patients with Chronic Conditions. 2019.
# Access to Primary Palliative Care

Table A.10. Access to Primary Palliative Care

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Access to Primary Palliative Care</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td><em>Primary palliative care</em> (PPC) is all of the palliative care competencies expected of the primary care physician, including but not limited to assessing a patient’s physical and nonphysical symptoms, assisting patients and their families to establish appropriate goals of care, communicating the patient’s illness and its trajectory, and affirming preference-concordant end-of-life care.¹ Primary care focuses on seeing the whole person and continuity of care, which is why the majority of palliative care is administered in that setting. In the United States, palliative care is provided within and outside hospice programs.² Providing PPC requires the following: involving an interdisciplinary team of providers to meet patient and family needs; creating systems for routine, reliable assessments of symptoms and care needs; creating pathways to triage patients who screen positive for symptoms or distress; and committing to continuing education and skills training for staff.³ The National Quality Forum developed a framework to assess performance and outcomes of palliative care at three levels: the patient/family-centered care level; the organizational/programmatic level; and the system level.⁴ Access to PPC could be operationalized in several ways.</td>
</tr>
<tr>
<td></td>
<td>• Provider surveys could assess whether health care providers have received training and continuing education on best practices for palliative care.</td>
</tr>
<tr>
<td></td>
<td>• Patient surveys could assess whether providers are following palliative care best practices and whether patients are experiencing serious health-related suffering (SHS).</td>
</tr>
<tr>
<td></td>
<td>• Use of morphine has been used as a crude measure of access to palliative care⁵ and could be accessed via administrative data.</td>
</tr>
<tr>
<td></td>
<td>• Insurance coverage of palliative care would provide insight into one aspect of access to palliative care.</td>
</tr>
<tr>
<td></td>
<td>• Surveys of national policies, clinical care guidelines, and training programs pertaining to palliative care could provide high-level insight.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Importance: Health or social impact</th>
<th>Population impact:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Unlike the Medicare hospice benefit, which can only be provided to people diagnosed with a terminal illness, palliative care can be delivered to a larger population living with chronic or advanced illnesses.⁶</td>
</tr>
<tr>
<td></td>
<td>• The World Health Organization (WHO) estimates that more than 20 million people require palliative care worldwide; most are adults over 60 years of age (69%) and 6% are children.⁷ Other researchers have estimated that 45% of all people who died in 2016 worldwide experienced SHS requiring palliative care (25.5 million deaths).⁸ This percentage is estimated to increase to 47% by the year 2060, accounting for more than 48 million deaths.⁸</td>
</tr>
<tr>
<td>Measure Construct</td>
<td>Access to Primary Palliative Care</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------------------</td>
</tr>
</tbody>
</table>
| **Clinical impact:** | • Palliative care can meaningfully improve quality of life, boost mood, lower symptom intensity, and decrease emotional distress. Some studies have demonstrated possible prolongation of life.\(^5,6,9\)  
• Many primary care providers might offer palliative care but not recognize or classify actions such as prescribing and titrating opioids for pain relief, maximizing quality of life, and symptom management as palliative care.\(^10\) |
| **Performance gap:** | • Demand for palliative care specialists exceeds supply, particularly in the outpatient supply.\(^11\)  
• PPC interventions often overlook the alignment of culturally based preferences with treatment.\(^12\)  
• An estimated 25.5 million (45%) of the 56.2 million deaths recorded worldwide experienced SHS in 2015. More than 80% of these people who died with SHS were from developing regions, and the vast majority lack access to palliative care and pain relief. Nearly 2.5 million children die with SHS every year; the vast majority (98%) of these children are from developing regions; additionally, at least 93% of child deaths associated with SHS in low-income countries are avoidable.\(^13\)  
• 91% of morphine consumed worldwide in 2013 was consumed in high-income countries, which have only 19% of the world’s population; this further illustrates the gap between wealthy and poor countries in access to palliative care.\(^14\) |
| **Costs:** | • Providing palliative care can lead to more-equitable resource utilization, reversing the trend of hospitalization and use of the intensive care unit at the end of life.\(^6,13,15\) |
| **Importance:** Related U.S. Department of Health and Human Services (HHS) priority topics | **Value-based care:**  
• The HHS priority to transform our health care system from one that pays for procedures and sickness to one that pays for outcomes also prioritizes quality of care. Gaps in quality of care are linked to pain and symptom management, communication, and care planning. Interventions to develop structured goals of care discussions can improve patient and family satisfaction.\(^16\) Value-based care incentives promote an environment that relies on keeping patients comfortable and, when possible, at home.\(^17\) |
| **Other related priorities:** Opioid crisis: The HHS priority to counter the crisis of opioid abuse, misuse, and overdose in the United States emphasizes the safe prescription and disposal of opioid pharmaceuticals.\(^18\) Palliative care specialists must strike a balance between patients who require pain management and potential harmful outcomes for patients who could require constant treatment for many years.\(^19\) |
| **Validity** | **Construct validity:**  
• Measures of attitudes toward PPC have shown strong content validity.\(^20\)  
• Depending on the measure selected, all aspects of palliative care might not be captured; for example, morphine consumption accounts only for care of physical symptoms and leaves out care of psychosocial and spiritual concerns. Insurance coverage of palliative care does not fully address access. |
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Access to Primary Palliative Care</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Internal validity:</strong></td>
<td>There is scientific evidence that palliative care improves patient symptoms; relieves physical, psychological, and spiritual suffering; and improves quality of life and patient and family satisfaction.⁵</td>
</tr>
<tr>
<td><strong>Discriminant validity:</strong></td>
<td>The International Classification of Diseases, Ninth Revision, code for PPC is v66.7. A study of veterans diagnosed with heart failure found that it was coded accurately 96% of the time.²¹</td>
</tr>
<tr>
<td><strong>Feasibility for international comparisons</strong></td>
<td><strong>Definitions:</strong></td>
</tr>
<tr>
<td></td>
<td>Currently, there is no commonly used definition for this measure.</td>
</tr>
<tr>
<td></td>
<td>The Palliative Care Registry measures inpatient palliative care rates by dividing service utilization by the inpatient population.²² A similar measure has not yet been developed for the outpatient population.</td>
</tr>
<tr>
<td><strong>Data availability:</strong></td>
<td><strong>National:</strong> No national data set measuring access to PPC exists. Chart reviews and surveys of the bereaved can be used to measure access, but validated universal measures are needed.</td>
</tr>
<tr>
<td></td>
<td><strong>International:</strong> WHO includes global access to palliative care in its package of essential services for universal health coverage.¹³ It used a Delphi study and WHO mortality and population estimates to estimate the need for palliative care.⁷</td>
</tr>
<tr>
<td></td>
<td>The Demographic and Health Surveys measure the provision of and guidelines for palliative care for people living with HIV/AIDS in countries around the world.²³</td>
</tr>
<tr>
<td></td>
<td>The UK Help the Hospices database includes the number of providers currently providing specialist palliative care, but this measure does not capture access to PPC.²⁴</td>
</tr>
<tr>
<td><strong>Data comparability:</strong></td>
<td>In the United States, palliative care and hospice care are distinguished by the former being based on need with no prognostic restriction and the latter based on a prognosis of living less than six months. In other countries, the terms are largely synonymous.²</td>
</tr>
<tr>
<td></td>
<td>The mechanisms of PPC need further study to identify the necessary components and core functions of effective care.¹²</td>
</tr>
<tr>
<td><strong>Subgroup analysis and equity:</strong></td>
<td>The lack of PPC for Black and Hispanic patients is partially offset by the increase in referrals to specialty palliative care. Nevertheless, the lack of PPC might lead to delayed access to any palliative care, poorer outcomes, and increased disparities in end-of-life care.²⁵</td>
</tr>
<tr>
<td><strong>Usability:</strong></td>
<td>Current evidence of the use of a measure of access to PPC (not exhaustive):</td>
</tr>
<tr>
<td><strong>Current evidence of use</strong></td>
<td>Historical chart reviews found high degrees of unmet need for PPC. One review found that 37% of patients had no evidence of any palliative care.²⁵</td>
</tr>
<tr>
<td></td>
<td>Estimating prognosis is challenging with complex patients and is key to prescribing palliative care. Online resources, such as <a href="http://www.eprognosis.org">www.eprognosis.org</a>, could guide primary care estimates.²⁶</td>
</tr>
</tbody>
</table>
The National Home and Hospice Care Survey studies palliative care as a process measure but does not differentiate between primary and specialized levels of care. The WHO affirmed that providing palliative care to patients and families is core to the role and identity of primary care clinicians. The Palliative Outcome Scale can be used to measure an unmet need for PPC in clinical trials and longitudinal studies. It has been translated into 12 languages and tested and validated in many countries and cultural contexts.

Access to PPC is not among the current OECD measures. It has also expressed interest and need for further study of palliative care provision. WHO has recommended a measure of suffering-adjusted life years to augment existing measures of the burden of ill health.

References
Prices for Brand-Name and Generic Drugs

Table A.11. Prices for Brand-Name and Generic Drugs

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Prices for Brand-Name and Generic Drugs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>Pharmaceuticals have enabled a multitude of improvements in health outcomes for patients. One in every five health care dollars is spent on pharmaceuticals in Organisation for Economic Co-ordination and Development (OECD) countries. Although overall spending has not changed in recent years, it has become increasingly skewed toward expensive specialty medicines. Measuring drug pricing at the international level could lead nations to make decisions that impact prices, making them available more broadly or to specific subpopulations of patients. • Brand-name and generic prescription drug pricing is dependent on quantity of services, mix of compounds, prices, and other factors. International comparisons are best done with sales of single-molecule products.</td>
</tr>
<tr>
<td><strong>Importance: Health or social impact</strong></td>
<td>Population impact: • <strong>National:</strong> In the United States, annual expenditures on pharmaceuticals exceed half a trillion dollars and account for 17% of national spending on health care. Americans have reported skipping prescriptions due to cost, and this phenomenon is more common for those who have lower income and more chronic conditions. Recent reports have described dramatic increases in prices for common medications, such as epinephrine and insulin, as well as high prices for new medications, such as biologic agents that treat cancer. In a 2017 poll, 61% of Americans deemed lowering the cost of prescription medications to be a top health care priority. • <strong>Global:</strong> In 2017, retail pharmaceuticals accounted for nearly one-fifth of all health care expenditure in OECD nations. Funding from governments or compulsory schemes (including Affordable Care Act–mandated insurance programs in the United States) covered 58% of spending on retail pharmaceuticals. Most of the remainder is financed from household out-of-pocket payments, which makes pricing an important factor in affordability and access. Compared with other high-income countries, the United States spends the most per capita on prescription drugs.</td>
</tr>
<tr>
<td><strong>Clinical impact:</strong></td>
<td>Patients rationing medications due to high prices (or out-of-pocket cost) may experience worse outcomes. For example, approximately 7.4 million Americans need insulin, and 25% of patients with diabetes ration insulin because of high cost. Recent reports have highlighted specific cases in which patients rationed insulin because of high prices and then died of diabetic ketoacidosis (uncontrolled diabetes).</td>
</tr>
<tr>
<td><strong>Performance gap:</strong></td>
<td>The World Health Organization (WHO) has recognized the importance of keeping pharmaceuticals accessible in all nations, including through the use of external reference pricing to identify nations with extraordinarily high prices.</td>
</tr>
<tr>
<td>Measure Construct</td>
<td>Prices for Brand-Name and Generic Drugs</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------------------</td>
</tr>
<tr>
<td>• Prescription drug prices were higher in the United States than in nine other high-income countries considered in a Commonwealth Fund report from 2017.⁵</td>
<td></td>
</tr>
</tbody>
</table>

**Costs:**

*National:* Generics account for 89% of prescriptions but only 26% of total drug costs in the United States.¹³

*Global:* Overall drug spending is estimated to increase 2–5% annually, exceeding $1 trillion by 2024.¹⁴

• High pharmaceutical prices may have equally high opportunity costs. With such a large portion of public and private spending dedicated to spending on pharmaceuticals, spending on other areas important to health, education, or national security could be decreased.¹⁵

• The long-term costs of morbidity due to skipping doses of key medications have been demonstrated to exceed the costs of providing medications for patients with certain chronic illnesses.¹⁶

<table>
<thead>
<tr>
<th>Importance: Related U.S. Department of Health and Human Services (HHS) priority topics</th>
<th>Drug pricing:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• The HHS priority to lower the price of prescription drugs without discouraging innovation relates to the control of prices relative to national budgets and research and development (R&amp;D) priorities.</td>
<td></td>
</tr>
</tbody>
</table>

**Other related priorities:**

*Value-based care:* If drug prices decrease, this may facilitate better adherence to medications, thus improving outcomes for patients and producing a better ratio of quality to cost.

<table>
<thead>
<tr>
<th>Validity</th>
<th>Construct validity:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Differences in how drug prices are measured can affect validity. For example, prices are up to 190% higher in the United States per capita, but as a share of total national expenditures, the United States is closer to the median.⁵,¹⁷</td>
<td></td>
</tr>
</tbody>
</table>

**Internal validity:**

• Prior research has demonstrated that more than 50% of the decline in health insurance coverage was attributable to increases in health insurance premiums,¹⁸ and currently 14% of the growth in health insurance premiums is due to growth in drug prices.¹⁹

**Discriminant validity:**

• A simple comparison of prices can overlook price differences in the types of drugs prescribed, the impact of using older products, and the impact of distribution costs.³ Measures of drug pricing might be more useful for comparing international trends over time rather than to produce rankings of countries in a single year, given the challenging comparability issues described above.

• Price variation between countries is not always considered a negative finding. Differences in pricing are considered an appropriate policy to allow companies to earn profits in high-income countries while still being able to conduct R&D and distribute products in lower-income markets.²

• Orphan drugs often target severe diseases and unmet medical needs for small populations, resulting in high prices. The high pressure on decisionmakers to fund them may exempt them from economic evaluation.² Orphan drugs have been found to be more clinically...
### Feasibility for international comparisons

**Definitions:**
*Price* is defined as the amount asked by a distributor for a product, while *cost* is the expense incurred to a patient, provider, or payer. Because of a complex web of payment arrangements, some of which are not transparent, high prices can, but do not always, lead to high prices for payers or onerous out-of-pocket costs for patients.

*Ex-factory prices* are prices at which manufacturers sell their products to wholesalers. In the United States, ex-factory prices correspond to the wholesaler acquisition cost. IQVIA cites a proprietary derived estimate of price called *net-price spending*, which is the amount received by pharmaceutical manufacturers after accounting for rebates, off-invoice discounts, and other price concessions made to distributors, health plans, and intermediaries.

*Rebates* are defined as payments from manufacturers to health plans in exchange for favorable formulary placement.

*Coupons* are defined as discounts to consumers that cover some or all of the consumer’s share of the retail price.

The United States distinguishes *retail sales* (directly to patients) from *institutional sales* (to hospitals, clinicians, nursing homes, infusion centers, home health agencies, and clinics). WHO distinguishes pricing, reimbursement, distribution, and availability of generics as separate but interrelated issues of importance to pharmaceutical policy.

**Data availability:**
- **National:** Medi-Span tracks wholesaler acquisition cost, which has recently been used as a proxy for price in a study comparing drug prices internationally (though prices were then adjusted downward by 17.8% to estimate the influence of rebates on price). Data describing net-price spending in the United States are also available from IQVIA (see definition above).
- **International:** Data are available through reference pricing authorities. Additionally, some OECD nations report data on pharmaceutical sales either as manufacturers’ prices or as expenditures at retail prices.

**Data comparability:**
- Comparing prices within and between countries is complicated by differences in product availability and formulations between countries, the proprietary nature of the rebates that manufacturers charge different payers, inclusion of different market sectors (retail versus inpatient versus outpatient), variability in definition of price, use of data describing price versus spending, distribution costs, value-added tax rates, and subpopulations of patients included in the evaluation (if inclusion varies based on insurance coverage). Actual prices paid by payers for expensive therapies are generally not known due to managed entry agreements that disconnect list prices from prices paid. A recent study found price variations between 28% and 388% for each drug surveyed.
- In the United States, rebates on branded drugs have been demonstrated to reduce price from as little as 16% (private insurance)}
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Prices for Brand-Name and Generic Drugs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>to as much as 61% (Medicaid). Limited availability of post-rebate prices in the United States thus limits interpretability of measures of pricing domestically and also internationally.26</td>
</tr>
<tr>
<td></td>
<td>• For international comparisons of generics and biosimilars, many countries only include data for the community pharmaceutical market or the reimbursed pharmaceutical market.7</td>
</tr>
<tr>
<td>Subgroup analysis and equity:</td>
<td>• Future analyses could evaluate prices for subgroups of patients (stratified by age, comorbidities, insurance status, or socioeconomic status).33</td>
</tr>
<tr>
<td>Usability: Current evidence of use</td>
<td>Current evidence of the use of drug pricing measures includes the following (not exhaustive):</td>
</tr>
<tr>
<td>National:</td>
<td>The Centers for Medicare &amp; Medicaid Services collects manufacturer-reported data of average sales price on a quarterly basis. These data inform the reimbursement of medications administered in physician offices and hospital outpatient departments under Medicare Part B.34</td>
</tr>
<tr>
<td>Global:</td>
<td>• In an academic study, a British research team used the Intercontinental Medical Statistics MIDAS database to examine variation in drug prices among selected OECD countries.35,36</td>
</tr>
<tr>
<td>Usability: Relationship to currently used OECD measures</td>
<td>Drug pricing is not represented among the current OECD measures.37</td>
</tr>
<tr>
<td>Closest related measures:</td>
<td>• Pharmaceutical spending: This measure describes expenditure on prescription and over-the-counter medications. It is measured as a share of total health spending in U.S. dollars per capita and as a share of gross domestic product. In the United States, the main data source for this measure is the National Health Interview Survey, a continuously updated survey of the noninstitutionalized civilian population.37,38</td>
</tr>
<tr>
<td></td>
<td>Although expenditure is a function of both price and utilization, simply dividing by utilization would not yield a valid measure of drug pricing. Instead, a measure of drug pricing would need to measure prices by analyzing sales of single-molecule products.3,7</td>
</tr>
</tbody>
</table>

References


Table A.12. Diffusion of and Access to New Prescription Drugs

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Diffusion of and Access to New Prescription Drugs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Pharmaceutical innovations have led to reductions in morbidity and mortality caused by such conditions as infectious diseases, cardiovascular diseases, and cancer. However, barriers to widespread diffusion of and access to new medications exist. For example, some health systems may be slower to adopt new medications despite good evidence of benefit because of poor communication about such benefits or because of policies such as price regulation. In the literature, the following definitions of diffusion of new drugs have been identified: 1. the share of daily doses and of expenditures in a therapeutic class that is captured by new medicines(^1) 2. the percentage of the population that has access to new products, by drug or class, reported by years since first approval (and or use) anywhere in the world.(^2)</td>
</tr>
<tr>
<td>Importance: Health or social impact</td>
<td>• <strong>Population impact:</strong>  o <em>National:</em> Many conditions are amenable to therapy with pharmaceuticals, including infectious diseases, cardiovascular diseases, and cancers. In the United States, roughly 20% of poll respondents reported that they had been unable to fill a prescription due to cost; a third of these individuals reported a serious health condition. One-third reported that prescription drug costs were a financial burden.(^3) In the United States in 2020, certain emerging pharmaceuticals are expensive, which has led to restriction in access for certain patients who could benefit (e.g., direct antivirals for hepatitis C).(^4) Compared with other Organisation for Economic Co-operation and Development (OECD) nations, market launches and adoption by providers could occur more rapidly in the United States.(^5) o <em>Global:</em> Multiple studies have demonstrated that approvals of new medications are slower in other nations than the United States.(^6,7) For example, lovastatin became available in Egypt 12 years after it gained FDA approval. According to a National Bureau of Economic Research (NBER) report, “long launch lags are common and nearly 40 percent of all new drugs are only launched in ten or fewer countries.”(^5) • <strong>Clinical impact:</strong> In the United States, age-adjusted heart disease mortality fell by more than 50% from 1950 to 1996, and then by another 22% from 1990 to 2013.(^8) Much of this improvement might be due to innovations in pharmaceuticals that became accessible to patients across the nation.(^9) Similar advances for other conditions could be made more rapidly. However, rapid adoption of new medications could prove costly or harmful if they have not undergone thorough testing of efficacy, safety, and cost-effectiveness prior to adoption. Questions have also been raised around the approval and diffusion of new medicines that might not be value-enhancing beyond what is currently available, as in the case of “me-too” drugs.(^10-12)</td>
</tr>
</tbody>
</table>
Diffusion of and Access to New Prescription Drugs

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Diffusion of and Access to New Prescription Drugs</th>
</tr>
</thead>
<tbody>
<tr>
<td>• <strong>Performance gap:</strong> Diffusion of and access to new drugs varies greatly from nation to nation.(^{13,14}) The public health benefits of a new drug depend on how quickly drugs are launched in markets and how widely they are adopted within a country once they are launched. Lower-income countries are less likely to have access to new medications because their markets are less profitable. This is an issue in the European Union in particular, because each country has its own policies regulating approval.(^5,15) Variation within countries after approval depends on multiple factors as well, including insurance coverage for subpopulations, physicians' scientific attitudes and prescribing patterns, communication among prescribers, and marketing.(^{16})</td>
<td></td>
</tr>
<tr>
<td>• <strong>Costs:</strong> The financial cost to society could be large if new medications experience diffusion and adoption more widely, since these medications tend to be costly. However, populations denied essential medications could experience poor outcomes leading to long-term disability, which is also costly.(^7)</td>
<td></td>
</tr>
</tbody>
</table>

| Importance: Related U.S. Department of Health and Human Services priority topics | Value-based care  
Access to cost-effective new drugs can improve health care value. |
|--------------------------------------------------------------------------------|------------------------------------------------------------------|
| **Other related priorities:**  
*Drug pricing:* Accelerating diffusion of or improving access to new drugs can be costly for new brand-name medications. Alternatively, regulating price for expensive medications could lead to more-rapid diffusion and equitable access to medications. |

| Validity | Construct validity: Multiple measures of diffusion of and access to new drugs exist, and it is challenging to comment on their validity without concrete examples. However, if nations do not agree on definitions of innovation, diffusion, and access, these measures could present threats to construct validity. In particular, definitions of innovation vary among nations. For example, there is some evidence that lower-value drugs are more likely to diffuse in the United States compared with other nations.\(^18,19\) Depending on which measure definition is used, this measure might be assessing aspects related to access, such as insurance coverage or cost, rather than drug diffusion. |
| Internal validity: Diffusion and access to new therapies might only be a valid measure of health system performance if it is measured for medications that are innovative and/or cost-effective. |
| Discriminant validity: International comparisons can be challenging due to difficulty accounting for differences in socioeconomic status and patient preferences for treatment options. Two factors influence whether a company will launch a new drug in a given country: demand, or the size of the potential market; and the nature of competition, impacted by a country’s policies on market regulation and intellectual property rights.\(^20\) |

| Feasibility for international comparisons | Definitions:  
According to OECD, a medicine is characterized as innovative if it (1) meets a previously unmet or inadequately met substantive health need or (2) offers enhanced effectiveness (e.g., greater efficacy, reduced toxicity, or both) or another incremental benefit (e.g., a substantive improvement in patient convenience) relative to existing therapeutic alternatives. |
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Diffusion of and Access to New Prescription Drugs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diffusion</td>
<td>Diffusion is commonly defined as “the process by which an innovation is communicated through certain channels over time among members of a particular social system.”</td>
</tr>
<tr>
<td>Access</td>
<td>Access to medications is defined by the United Nations Development Group as “having drugs continuously available and affordable at public or private health facilities or drug outlets.”</td>
</tr>
</tbody>
</table>
| Data availability:| • National: Sources of data could include commercial insurance databases and Medicare Part D data.  
• International: Multiple reports cite the MIDAS data set produced by IMS Health as a source of commercial market research data describing product launches in multiple therapeutic classes, including quarterly unit sales and revenues at the medication package level. |
| Data comparability:| • Measures comparing use of medications internationally generally use the World Health Organization’s (WHO’s) Anatomical Therapeutic Chemical classification system to identify active ingredients. This covers all classes, but not all countries.  
• Identifying drug launches over time can be challenging because brand names are often referenced rather than active ingredients, which makes one-to-one comparison difficult.  
• There may be a question of how to define what is considered new or innovative. The United Kingdom’s NICE system, for example, does a cost-benefit analysis and might not allow certain drugs to be covered if their benefit is deemed low. This lack of access to lower-value drugs could count against the United Kingdom in assessing health system performance, for example. |
| Subgroup analysis and equity:| • There is subgroup analysis, as demonstrated in the NBER report: differences in access to a drug of interest by subpopulation (e.g., patients at least 65 years old) in hazard analysis. This could be used for other subpopulations as well. |
| Usability:        | • National: Emphasis has largely been on mathematical modeling to describe the diffusion and access to new prescription drugs.  
• International: Europe uses the Patients W.A.I.T. indicator to measure access to innovative therapies. This indicator shows the rate of availability (the number of medicines available to patients in each European country) and the average time between marketing authorization and patient access (number of days elapsed from the date of European Union marketing to the day of completion of post-marketing authorization administrative processes). WHO recently measured access to direct antiviral therapies for hepatitis C by surveying producers of these medications about pricing (price per 28-day supply), licensing, and regulatory status. WHO also used pharmaceutical sales data to estimate the number of people who received direct-acting antivirals. |
## Measure Construct

**Usability: Relationship to currently used OECD measures**

Diffusion of and/or access to new prescription drugs is represented among current OECD measures:

- Biosimilar market share in treatment days for anti–tumor necrosis factor alphas and erythropoietin versus accessible market, 2017 (or nearest year), in European countries: Based on this measure, a nation’s access to innovative medications would be thought of as superior if there was greater availability of biosimilars (biochemically similar products that are more affordable than comparable reference medications).²⁸
- Annual approvals of new medicines per billion U.S. dollars of pharmaceutical business expenditure on research and development (R&D) in the United States, inflation-adjusted: This measure analyzes the ratio of number of new approvals of medications per unit of expenditure on drug R&D. In other words, this is a measure of return on investment in development and approval of drugs.²⁹

## References

2. Expert-provided definition.
16. Lublóy Á. Factors affecting the uptake of new medicines: A systematic literature review. *BMC Health Serv Res.* 2014;14:469.
Avoidable Emergency Department Use

Table A.13. Avoidable Emergency Department Use

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Avoidable Emergency Department Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Avoidable (or preventable) emergency department (ED) visits are those in which people use hospital EDs for nonurgent care and for conditions that could have been treated in a primary care setting. These are discharged ED visits not requiring any diagnostic or screening services, procedures, or medications. ED visits for ambulatory care–sensitive conditions (ACSCs) are considered avoidable. Avoidable ED visits can be experienced by all populations, regardless of sociodemographic or health insurance factors.</td>
</tr>
<tr>
<td>Importance: Health or social impact</td>
<td>Population impact:</td>
</tr>
</tbody>
</table>
| National: | • Nationally, about 139 million people (across all age groups) visited the ED in 2017, and 3.9% of these visits were classified as nonurgent (i.e., the patient could be seen in 2–24 hours). However, urgency and preventability are not necessarily synonymous. Using the above definition, from 2005 to 2011, about 3.3% of ED visits among adults (those 18–64 years of age) nationally were classified as avoidable, with most cases related to mental health and dental conditions.  

• Youth. About 13% of ED visits in 2010 among the national pediatric population (0–19 years of age) were deemed avoidable.  

• Older adults. Nationally, approximately 36% of avoidable ED visits (those that were related to an ACSC) between 2007 and 2009 occurred among adults aged 65 or older.  

Global:  

• Adults. There is currently no global study of avoidable ED visits. However, the prevalence of ED visits from 2011 to 2013 was 29.4% among adult patients (18 years or older) in the EU Seventh Framework project Quality and Costs in Primary Care (includes 31 European countries, Australia, New Zealand, and Canada); of these visits, it is unclear how many were avoidable. Some estimates are available for individual countries. For example, a national analysis from Taiwan suggests that 20% of ED visits are emergent but preventable by primary care.  

• Youth or older adults. Currently, global estimates of avoidable ED visits in youth or older adults are not available.  

Meaningful clinical impact:  

• Avoidable ED visits could contribute to unnecessary testing and treatment and weaker patient–primary care provider relationships.  

• Reducing unnecessary admissions is important because of their cost to commissioners, service providers, and patients and their families.  

Performance gap:  

• Avoidable ED visits can indicate poor care management or inadequate access to care.  

Costs: In the United States, ED visits account for 12.5% of national health expenditures ($328.1 billion U.S. dollars in 2010). Furthermore,
### Avoidable Emergency Department Use

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Avoidable Emergency Department Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Importance: Related U.S. Department of Health and Human Services (HHS) priority topics</td>
<td>Avoidable ED visits accounted for $64.4 billion, or 2.4% of national health expenditures. In the pediatric population (0–19 years), the cost of avoidable ED visits due to ACSCs was $3 billion in 2010.</td>
</tr>
<tr>
<td>Health insurance reform:</td>
<td>Policies or strategies to reduce avoidable ED visits could help reduce health care costs and, therefore, have important relevance to the HHS priority to reform health insurance. Evidence from the Affordable Care Act (ACA) shows that extending eligibility for health insurance to young adults (aged 19–25 years) resulted in significant decreases in ED visits in this age group. However, one in ten ED visits are still made by uninsured populations in the United States, suggesting the need for additional reforms. Furthermore, with the ACA, a growing prevalence of high-deductible health plans has been observed in both the group and individual markets, and this can have important implications for ED use. For example, high-deductible health plans incur higher costs for an ED visit for an individual compared with traditional health plans and, as a result, may contribute to reduced ED visits. An unintended consequence of such plans is that the high costs may pose a barrier to using the ED for a true emergency. Since the Deficit Reduction Act of 2005, states may impose mandatory cost-sharing for nonurgent ED visits for Medicaid patients. Most recently, Indiana set copayments for nonemergency care delivered in the ED as part of its Medicaid expansion waiver in 2015. Health plans, too, have emphasized the importance of avoiding the ED to reduce costs.</td>
</tr>
<tr>
<td>Validity</td>
<td>There is wide variability in the data sources used across studies to define avoidable ED visits (e.g., retrospective diagnoses, hospital admissions, triage scores, patient self-reported data), and validity of these measures is not always reported.</td>
</tr>
<tr>
<td>Predictive validity:</td>
<td>One widely used algorithm for classifying avoidable ED visits developed by researchers at New York University (known as the NYU-ED algorithm) based on diagnosis data has demonstrated acceptable predictive validity in the United States. Specifically, the algorithm was able to differentiate between individual ED visits that were and were not likely to result in hospitalization or death. However, a limitation of this algorithm is that it depends on diagnosis codes and physician coding patterns, which can change over time, thereby altering the validity of the algorithm.</td>
</tr>
<tr>
<td>Feasibility for international comparisons</td>
<td>Definitions: There is currently no internationally agreed-upon definition of avoidable ED visits. Important differences across health care policies and systems could make it difficult to develop a standard measure that can work on a global scale.</td>
</tr>
<tr>
<td>Data availability:</td>
<td>National: In the United States, the National Hospital Ambulatory Medical Care Survey obtains nationally representative data on all ED visits with information on level of urgency (but not on whether they were avoidable or preventable).</td>
</tr>
<tr>
<td>Measure Construct</td>
<td>Avoidable Emergency Department Use</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------------------</td>
</tr>
</tbody>
</table>
| **Global:**       | - The multicountry Quality and Costs of Primary Care in Europe (QUALICOPC) survey collects ED visit data from hospitals, but it does not distinguish between avoidable and true emergency visits.\(^6\)  
|                   | - A related variable is OECD’s *avoidable hospital admissions indicator*, defined as the number of hospital admissions with a primary diagnosis of asthma, chronic obstructive pulmonary disease (COPD), or congestive heart failure (CHF) among people aged 15 years and over. These conditions can be effectively treated in the primary care setting.\(^19\) However, this indicator focuses on a limited set of avoidable admissions and is not specific to the ED. |
|                   | **Data comparability:** As demonstrated by the QUALICOPC data, ED visit data can be pooled across countries, but to distinguish avoidable visits from those that are not avoidable will require adding information about the patient’s reason for the visit (e.g., self-report health condition or symptoms or using hospital International Classification of Diseases codes). Furthermore, the QUALICOPC is primarily focused on higher-income countries; thus, the feasibility of obtaining ED data from low- and middle-income countries is unknown. |
|                   | **Subgroup analysis and equity:** In the United States, avoidable ED visits due to ACSCs among adults are higher among those who are non-Latino Black and those who are Latino than among those who are non-Latino White, higher among Medicare recipients than among those with private insurance, higher among older adults (50–64 and 65 or older) than among younger adults (18–29 years), and higher among women.\(^3\) Data from the multicountry QUALICOPC survey suggest that women and older adults are less likely to visit the ED, while those with less education are more likely to visit it, but this is not specific to avoidable visits.\(^6\) |
| **Usability:**    | **Current evidence of use**  
|                   | National: No national study specifically collects avoidable ED visits. However, general ED visit information is obtained by the National Hospital Ambulatory Medical Care Survey\(^18\) and the Agency for Healthcare Research and Quality Healthcare Cost and Utilization Project.\(^20\) These data sources may need to be linked to other data (e.g., medical records) to distinguish avoidable ED visits. |
|                   | **Global:** No global study collects avoidable ED visits. However, general ED visit data are obtained by the QUALICOPC survey,\(^21\) and avoidable hospitalization data are collected by the Organisation for Economic Co-operation and Development (OECD).\(^19\) |
| **Usability:**    | **Relationship to currently used OECD measures**  
|                   | A measure for avoidable ED visits is not among the current OECD measures. |
|                   | **Closest related measures:**  
|                   | - The avoidable hospital admissions indicator measures the number of hospital admissions with a primary diagnosis of asthma, COPD, or CHF among people aged 15 years and over per 100,000 population. Rates are age-sex standardized to the 2010 OECD population aged 15 and over. Admissions resulting from a transfer from another hospital and where the patient dies during admission are excluded from the calculation, as these are considered unlikely to be avoidable.\(^19\) |
References

The Percentage of Patients with an Opioid Use Disorder Who Were Referred to or Prescribed Medication-Assisted Treatment

Table A.14. The Percentage of Patients with an Opioid Use Disorder Who Were Referred to or Prescribed Medication-Assisted Treatment

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>The Percentage of Patients with an Opioid Use Disorder Who Were Referred to or Prescribed Medication-Assisted Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>Medication-assisted treatment (MAT) is the most effective medical treatment for opioid use disorder (OUD).(^1) Despite the prescribing options available, estimates suggest that only 20–40% of people with OUD receive MAT.(^2) The referral treatment process consists of assisting a patient with accessing MAT; selecting treatment facilities; and helping navigate any barriers, such as treatment cost or lack of transportation, that could hinder accessing treatment in a specialty setting.(^3) Measuring the number of people who were or were not able to access MAT requires measuring the scope of the problem, treatment providers available, and education of those providers.</td>
</tr>
<tr>
<td></td>
<td>• Sometimes referred to as “opioid abuse or dependence” or “opioid addiction,” OUD is a problematic pattern of opioid use that causes significant impairment or distress.(^4) • MAT is the use of medications, in combination with counseling and behavioral therapies, to provide a “whole-patient” approach to the treatment of substance use disorders (SUDs). MAT is primarily used to treat addiction to opioids, but it is also used for alcohol dependence.(^5) Historically, the primary treatment for OUD was methadone, which was only administered by accredited opioid treatment programs.(^6) Since the Drug Addiction Treatment Act of 2000, buprenorphine can be prescribed in any location by qualified physicians who have applied for and received a Substance Abuse and Mental Health Services Administration waiver.(^7)</td>
</tr>
<tr>
<td><strong>Importance: Health or social impact</strong></td>
<td><strong>Population impact:</strong></td>
</tr>
<tr>
<td></td>
<td>• <strong>National:</strong> The misuse of and addiction to opioids is a major public health crisis in the United States, causing over 1.6 million years of life lost annually—more than the burden attributable to homicides, hypertension, HIV/AIDS, and pneumonia.(^8) As of 2018, 42.3% of opioid treatment centers offered MAT.(^9) • <strong>Global:</strong> Although the opioid epidemic is most pronounced in the United States, more and more countries are seeing growing numbers of OUDs and overdose-related deaths.(^10) The United Nations 2017 report on drug abuse stated that more than 200 million people, or 4.2% of the world population, were addicted to illegal drugs.(^11)</td>
</tr>
<tr>
<td></td>
<td><strong>Clinical impact:</strong></td>
</tr>
<tr>
<td></td>
<td>• <strong>Consequences:</strong> Being diagnosed with an OUD increases the risk of early death by a factor of 20.(^12) MAT is the most effective treatment for OUD.(^1) Outpatient treatment facilities are particularly important for the treatment of SUDs; only 13% of individuals who receive treatment do so in a private physician’s office.(^13)</td>
</tr>
<tr>
<td></td>
<td><strong>Performance gap:</strong></td>
</tr>
<tr>
<td></td>
<td>• In 2018, 2 million people in the United States had an OUD.(^14) However, the capacity to treat OUD remains at less than 50% of the total number</td>
</tr>
</tbody>
</table>
The Percentage of Patients with an Opioid Use Disorder Who Were Referred to or Prescribed Medication-Assisted Treatment

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>The Percentage of Patients with an Opioid Use Disorder Who Were Referred to or Prescribed Medication-Assisted Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The National Institute on Drug Abuse estimates that one-third of patients with opioid dependence at treatment programs receive MAT.(^\text{16})</td>
</tr>
<tr>
<td></td>
<td>• Rural populations are particularly underserved—60% of small, nonmetropolitan county residents live in an opioid treatment shortage area.(^\text{17})</td>
</tr>
<tr>
<td></td>
<td>• Only 28% of opioid overdose survivors are referred to MAT, despite evidence of its effectiveness.(^\text{18})</td>
</tr>
<tr>
<td></td>
<td><strong>Costs:</strong></td>
</tr>
<tr>
<td></td>
<td>• <strong>National:</strong> In the United States, recent estimates suggest that the costs of the opioid epidemic exceed $500 billion or 2.8% of GDP.</td>
</tr>
<tr>
<td></td>
<td>• <strong>Global:</strong> Because the global opioid epidemic is in its nascent stage, the global costs have not yet been realized or projected. But there is evidence that drug companies are applying tactics used in the United States in order to increase opioid sales in Asia and Europe.(^\text{19})</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Importance: Related U.S. Department of Health and Human Services (HHS) priority topics</th>
<th>Opioid crisis:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• HHS is committed to ending the crisis of opioid addiction and overdose in the United States. One point of HHS’s 5-point strategy to combat the opioid crisis is “better addiction prevention, treatment, and recovery services.” The clinician workforce is critical to providing better treatment and recovery services.</td>
</tr>
<tr>
<td></td>
<td><strong>Other related priorities:</strong></td>
</tr>
<tr>
<td></td>
<td>• <strong>Health insurance reform:</strong> The HHS priority to improve the availability and affordability of health insurance seeks to help all Americans access health care that meets individual needs and budgets.(^\text{20}) HHS has a vision of “a system [of] affordable, personalized care.”(^\text{21}) In states that expanded access to Medicaid, MAT prescriptions have increased, driven by office-based buprenorphine.(^\text{22,23})</td>
</tr>
<tr>
<td></td>
<td>• <strong>Value-based care:</strong> Integrating MAT into general medical practices could increase access, helping to transform our health care system from one that pays for procedures and sickness to one that pays for outcomes and health, but low reimbursement levels disincentivize providers from offering it.(^\text{24})</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Validity</th>
<th>Construct validity:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Because MAT is the most effective medical treatment for OUD,(^\text{1}) the percentage of patients with an OUD who were referred to or prescribed MAT is an important determinant of health.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Internal validity:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• The rate at which a health care system prescribes the more proven, effective treatment for a disorder, including OUD, is a critical component of the quality of health care available to the consumer.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Discriminant validity:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Greater prevalence of diagnosed OUD and other SUDs may reflect diagnostic differences rather than actual use and abuse of substances.</td>
</tr>
<tr>
<td></td>
<td>• Additionally, greater prevalence of OUD and other SUDs may reflect elements of the public health system (e.g., laws) rather than the health care system.</td>
</tr>
</tbody>
</table>
### Feasibility for international comparisons

**Definitions:**
- The World Health Organization’s official recommendation for the treatment of OUD is “the use of a range of treatment options for opioid dependence which include psychosocial support, opioid maintenance treatments such as methadone and buprenorphine, supported detoxification and treatment with opioid antagonists such as naltrexone.”
- Each country has different regulations on who can prescribe MAT. For example, in Canada and Europe, buprenorphine and methadone can be prescribed by a pharmacist, but in the United States the majority of prescriptions are by specialist physicians.
- There may be international differences in the definition of OUD that could influence the needs and capacities of different countries.

**Data availability:**
- **National:**
  - The National Survey on Drug Use and Health (NSDUH) documents illicit drug use, including OUD, and describes the use of MAT.
  - Administrative data (including claims data from Medicare, Medicaid, state all-payer databases, and IQVIA) contain diagnostic codes that would indicate OUD and describe prescription claims (for methadone, buprenorphine, and naltrexone).
- **Global:**
  - MAT is available in 80 countries around the world. Many European countries and other OECD countries regularly field surveys like the NSDUH that attempt to describe licit and illicit drug use in their populations.
  - At least 25 countries routinely collect administrative data on diagnoses and prescribed and dispensed medications.

**Data comparability:**
- Differences in how certain data are collected across countries may threaten comparability. For instance, some countries have national prescription registries, while others collect separate data describing reimbursement and prescription. Other countries have good prescription data for specific subpopulations (e.g., those covered by certain insurance programs) but lack data for other subpopulations.
- Governmental and public health system characteristics may play a role in the prevalence of OUD of each country. For instance, in some countries where drugs have been decriminalized (e.g., Portugal), overdoses and infections have been reduced.
- International definitional differences (which would arise if there were not commonly agreed-upon standards for OUD diagnoses) would also threaten the comparability of the data.

### Usability: Current evidence of use

**National:**
- The National Quality Forum measures the initiation and engagement of alcohol and other drug abuse or dependence treatment. This is not limited to OUD or to MAT but would include OUD as a type of drug abuse and MAT as a treatment.
- The Joint Commission uses the percentage of those with acute opioid disorder (AOD) diagnosis who receive or refuse a MAT or referral at a hospital, which is comparable to the National Quality Forum’s AOD.
The Percentage of Patients with an Opioid Use Disorder Who Were Referred to or Prescribed Medication-Assisted Treatment

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>The Percentage of Patients with an Opioid Use Disorder Who Were Referred to or Prescribed Medication-Assisted Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>treatment provided or offered at discharge from inpatient hospitalization.</td>
<td></td>
</tr>
<tr>
<td>• The National Survey of Substance Abuse Treatment Services collects data on gross measures of organizational structure, such as the percentage of referrals, but does not measure a program’s ability to prescribe medications. Nor is implementation measured, which may lead to inflated inferences about how therapies are used.</td>
<td></td>
</tr>
<tr>
<td>Global:</td>
<td></td>
</tr>
<tr>
<td>• The Canadian Institute for Health Information measures the hospital stays for harm caused by substance use. However, this measure does not involve the workforce size or capacity.</td>
<td></td>
</tr>
<tr>
<td>• The Australian Institute of Health and Welfare tracks referral to treatment by treatment source, such as family, law enforcement, or a health professional. It also tracks the different types of treatments to which patients are referred, such as whether MAT is combined with counseling.</td>
<td></td>
</tr>
<tr>
<td>Usability: Relationship to currently used OECD measures</td>
<td></td>
</tr>
<tr>
<td>The percentage of patients with an OUD who were referred to or prescribed MAT is not among the current OECD measures.</td>
<td></td>
</tr>
<tr>
<td>Closest related measures:</td>
<td></td>
</tr>
<tr>
<td>• Opioid use: These measures describe the availability of analgesic opioids and the rate of opioid-related deaths. Therefore, these measures describe the population (or a portion thereof) that could be prescribed MAT.</td>
<td></td>
</tr>
<tr>
<td>• Safe primary care prescribing: OECD tracks the volume of opioids prescribed and the proportion of chronic opioid users. Again, these measures describe the population (or a portion thereof) that could be prescribed MAT.</td>
<td></td>
</tr>
<tr>
<td>• Opioid-related deaths: The European Monitoring Centre for Drugs and Drug Addiction collects these data and shares them with OECD.</td>
<td></td>
</tr>
</tbody>
</table>

References

The Percentage of Patients with an Opioid Use Disorder Who Were Referred to or Prescribed Medication-Assisted Treatment


Wallack SS, Thomas CP, Martin TC, Chilingerian J, Reif S. Substance abuse treatment organizations as mediators of social policy: Slowing the adoption of a congressionally approved medication. J Behav Health Serv Res. 2010;37(1):64-78.


effective-treatments-opioid-addiction


Estimates of Administrative Complexity and Cost

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Estimates of Administrative Complexity and Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>In the clinical setting, administration refers to the nonclinical activities that regulate or control the behavior of individuals in a health care organization, enabling them to implement policy decisions and achieve their goals. Administrative tasks must be carried out at all levels of the health care system to ensure high-quality health care. Although more resources invested in administration can be productive, especially high administrative complexity and cost can be a sign of wasteful spending that could be curtailed without worsening patient health outcomes. Specifically, administrative complexity is the complicated workflows and fragmented financing procedures that can waste clinicians’ time and interfere with their caring for patients. Estimates can include, but are not limited to, time spent preparing paperwork (e.g., processing and billing of claims) and contacting payers. Administrative costs are the costs to carry out administrative services that help ensure against illness and deliver medical care. Estimates can include, but are not limited to, transaction-related costs, benefits management, selling and marketing, and regulatory/compliance costs.</td>
</tr>
<tr>
<td>Importance: Health or social impact</td>
<td>Population impact:</td>
</tr>
<tr>
<td>National: In 2018, the United States spent 17.7% of the gross domestic product on health care. This amount of spending is higher than that of any other nation. Recent estimates of the percentage of health spending dedicated to administration in the United States range from 8.3 to 25.3%, which is higher than administrative spending in other Organisation for Economic Co-operation and Development (OECD) nations. (Note that the wide range is due to different definitions of administration.) Furthermore, high administrative complexity can have negative impacts on the nation’s health professionals who are involved in administration. For example, physicians who spend more time on complex administrative duties (e.g., prior authorizations, clinical documentation, and medication reconciliation) report lower levels of career satisfaction and higher levels of burnout and are more likely to consider cutting back on seeing patients in the future.</td>
<td></td>
</tr>
<tr>
<td>Global: Rising health care costs have proved to be problematic for many nations for a myriad of reasons, including the advancement of expensive innovations in care. However, high administrative costs could lead to unnecessary health-related expenditures in some countries more than others. Reducing wasteful spending on administrative spending could allow nations to divert resources toward other more-productive activities in health care and other domains. A recent study using OECD data found that governance and financing-related administrative spending at the macro level has remained stable over the last decade at slightly more than 3% of total health spending. In the sample, Iceland had the lowest administrative spending (1.3% of total health spending) and the United States had the highest (8.3%). Nations having multiple payers tended to have higher administrative spending than those with single payers.</td>
<td></td>
</tr>
</tbody>
</table>

<p>| Importance: Health or social impact | Population impact: |
| National: In 2018, the United States spent 17.7% of the gross domestic product on health care. This amount of spending is higher than that of any other nation. Recent estimates of the percentage of health spending dedicated to administration in the United States range from 8.3 to 25.3%, which is higher than administrative spending in other Organisation for Economic Co-operation and Development (OECD) nations. (Note that the wide range is due to different definitions of administration.) Furthermore, high administrative complexity can have negative impacts on the nation’s health professionals who are involved in administration. For example, physicians who spend more time on complex administrative duties (e.g., prior authorizations, clinical documentation, and medication reconciliation) report lower levels of career satisfaction and higher levels of burnout and are more likely to consider cutting back on seeing patients in the future. |
| Global: Rising health care costs have proved to be problematic for many nations for a myriad of reasons, including the advancement of expensive innovations in care. However, high administrative costs could lead to unnecessary health-related expenditures in some countries more than others. Reducing wasteful spending on administrative spending could allow nations to divert resources toward other more-productive activities in health care and other domains. A recent study using OECD data found that governance and financing-related administrative spending at the macro level has remained stable over the last decade at slightly more than 3% of total health spending. In the sample, Iceland had the lowest administrative spending (1.3% of total health spending) and the United States had the highest (8.3%). Nations having multiple payers tended to have higher administrative spending than those with single payers. |</p>
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Estimates of Administrative Complexity and Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Meaningful clinical impact:</strong> Administrative complexity can result in frequent interruptions in the work of physicians and their staff, which are likely to interfere with patient care.(^1) There is some evidence from the United States that health care organizations spending more resources on administration deliver lower-quality care.(^2) Some stakeholders perceive money spent on administration to be money wasted because it may be less likely to produce direct benefits to patients compared with spending on health care delivery itself.(^3)</td>
<td></td>
</tr>
</tbody>
</table>

**Performance gap:** Recent comparisons have demonstrated that administrative spending in the United States is double that of Canada.\(^4\) A 2019 article by Shrank et al. estimated current levels of waste in the U.S. health care system in six previously developed domains, including spending on administration complexity.\(^5,6\) Countries with multipayer systems tend to exhibit higher administrative costs than countries with a single payer, and private insurers’ administrative costs are generally higher than those of public schemes. In many OECD nations, key areas where waste can be reduced include making better use of information and communications technology between payer and provider, simplifying administrative procedures, and finding the right size of administrative bodies.\(^2\)

The performance of a health care system is also measured by its efficiency, among other domains, and efficiency includes estimates of administrative complexity and costs. Across five performance health care system dimensions, the United States scores lowest in efficiency.\(^6\) More time spent on administrative tasks could reflect lower efficiency. An analysis of 11 high-income countries found a high administrative burden in the United States, with 54% of U.S. physicians reporting time spent on administrative issues related to insurance or claims as a major problem, compared with the mean of 41% across all 11 countries.\(^6\)

**Costs:**
- **National:** In the United States, administrative costs account for 8–25% of health care spending, which is higher than in other high-income countries (1–3%).\(^7\) An analysis comparing administrative costs in the United States with those in Canada in 2010 found that the United States generated $15 billion to $32 billion in excess costs per year.\(^4\) A physician-level study of administrative waste found that the costs of billing and insurance-related expenses varied from $20 for a primary care visit to $215 for an inpatient surgical procedure.\(^5,7\) The estimated total annual cost of waste from administrative complexity in the United States is $265.6 billion (2019 U.S. dollars).\(^5\)
- **Global:** Among OECD countries, health system administrative costs represent an average 3% of health spending, with higher spending among multipayer systems with free choice of insurer (as in the case of the United States) than among those with automatic affiliation or single-payer systems (whether the payer is a social health insurance fund or a government entity).\(^8,10\) System-specific investigations have uncovered administrative waste in other countries, including Australia (AUD 106 million was regained by removing duplication of administrative services) and Germany (EUR 4.3 billion of administrative costs were related to documentation and reporting).\(^2\)
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Estimates of Administrative Complexity and Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Importance:</td>
<td>Value-based care: Estimating administrative complexity and costs aligns with the HHS priority to deliver value-based care because reducing administrative burden in health care can empower health professionals to spend more time caring for patients while reducing overall health care costs. Furthermore, paying for value rather than volume may encourage health care organizations and payers to decrease spending on inefficient administration.</td>
</tr>
<tr>
<td>Related U.S.</td>
<td><strong>Validity</strong></td>
</tr>
<tr>
<td>Department of Health and Human Services (HHS) priority topics</td>
<td>Construct validity: Measures of administrative spending and/or waste have been used to identify low-hanging fruit where cuts can be made to save resources without sacrificing quality of care. However, one threat to construct validity is that nations may vary substantially in how they classify different types of spending. Analyzing one nation’s administrative spending trend over time could be more meaningful than comparing absolute numbers across nations. Furthermore, administrative complexity has been measured by collecting physician-reported time spent on administrative tasks, but such questions have not been validated. How administrative tasks interfere with patient care has been measured by a 16-item validated instrument, the Maslach Burnout Inventory—General Survey.</td>
</tr>
<tr>
<td></td>
<td>Internal validity: There is some evidence from the United States that health care organizations spending more resources on administration deliver lower-quality care. However, floor and ceiling effects likely characterize administrative spending, since some administrative spending is necessary to coordinate high-quality care, but too much spending is likely to be wasteful.</td>
</tr>
<tr>
<td></td>
<td>Discriminant validity: If an organization, health system, or nation categorizes “backroom” support personnel as administrative waste, cutting such personnel could lead to lower physician productivity and thus higher costs of providing care to patients.</td>
</tr>
<tr>
<td>Feasibility for international comparisons</td>
<td>Definitions: Per OECD, administrative costs refer to the costs associated with the governance and administration of the health system and the collection and pooling of financial resources by different health financing schemes. Estimates of administrative complexity in the United States and internationally are often based on physician-reported time spent on various administrative tasks (i.e., perceived burden). In the United States, administrative costs have previously been subdivided into costs attributable to private insurers (e.g., claims processing, profits), the government (e.g., eligibility determination, revenue collection), providers (e.g., billing, collections, collection and reporting of quality data), and employers (e.g., choosing plans, designing benefits). The administrative loss ratio has been defined as a health insurance plan’s total administrative costs divided by its total premium revenues. The medical loss ratio is the proportion of insurance premium revenues spent on clinical services and quality improvement. A higher medical loss ratio means that the insurer is spending less money on overhead. The System of Health Accounts described by OECD provides an international framework for the definition, demarcation, and categorization of administrative costs.</td>
</tr>
</tbody>
</table>

Estimates of Administrative Complexity and Cost
Estimates of Administrative Complexity and Cost

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Estimates of Administrative Complexity and Cost</th>
</tr>
</thead>
</table>

of health expenditure. Under this system, transactions are classified around who pays (financing), who provides (provision), and what the purpose is (function). For non-OECD nations, the World Health Organization coordinates a reporting system that compiles health care expenditures as *National Health Accounts.*

**Data availability:**
- **National:** A recent study used data from the U.S. National Health Expenditure Accounts (NHEA) to estimate costs of administering government programs and health insurance. The study included an estimate of administrative costs at the physician level (which are not accounted for in NHEA) by using physician effort dedicated to interaction with payers.
- **Global:** Himmelstein et al. performed an international comparison of eight nations using publicly available data to estimate administrative costs using prespecified categories of cost. Another study used OECD System of Health Accounts data to estimate macro-level financing-related expenditures in health care (but did not estimate administrative expenditure by health care providers). Furthermore, the Commonwealth Fund International Health Policy Survey's Physician Supplement collects prevalence of administrative efficiency (e.g., complexity) as measured by physician-reported time spent on administrative issues (e.g., insurance claims, reporting clinical or quality data, and paperwork or disputes related to medical bills).

**Data comparability:** Estimates of administrative costs vary even within the United States. Much of this variation is due to different definitions of administration (e.g., whether insurer profits and spending on population health are administrative costs). The debate around administrative costs of Medicare, for example, is an example of such challenges. Moreover, approaches in the estimation and valuation of administrative spending vary across countries, which may affect comparability of data. For example, (1) lack of data for certain types of administration (e.g., physician practice overhead) might lead to underestimation of administrative spending; (2) spending on preventive services, such as vaccination, could be misclassified as administrative spending in some nations; (3) differences in depreciation methods could lead to variance in administrative spending estimates across nations; and (4) differences in accounting for profits and brokerage fees of private health insurance could lead to biases in estimates of administrative spending. A scoping review of administrative costs in health care published by a German author in 2018 determined that heterogeneous reporting of costs might affect comparability across nations. International comparisons can also be complicated by differences across countries in how budgets are set, how reimbursement rates with providers are negotiated, differences in types of staff or technology, and differences in prices to carry out the same administrative functions, for example. To collect these additional pieces of information would entail enormous expense, time, and logistical difficulties.

Moreover, electronic health record (EHR) systems are being adopted in some countries, such as the United States, which contributes to the administrative complexities and costs; however, including EHR administrative costs could make it difficult to compare overall estimates to countries that do not have EHR systems. Furthermore, given that past
Estimates of Administrative Complexity and Cost

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Estimates of Administrative Complexity and Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>international measures of administrative complexity and costs have been limited to higher-income countries, their feasibility in developing countries is unclear.</td>
</tr>
</tbody>
</table>

**Subgroup analysis and equity:**
- Subgroup analyses have looked at administrative spending by payer type, generally for publicly versus privately funded payers. For example, a study of U.S. billing and insurance-related administrative costs determined that 3.1% of Medicare costs were directed toward administration, compared with 18% for private insurers.\(^{30}\)

**Usability:**

**Current evidence of use**
The United States and OECD collect data on administrative complexity and costs to track wasteful spending of health care dollars, a sign of poor health systems management, and thus identify areas for potential savings.\(^{15,18}\) Current evidence of the use of measures of administrative complexity includes the following (not exhaustive):

**National:**
- The percentage of national health expenditures dedicated to administrative spending: based on NHEA data\(^ {26}\)
- Medical loss ratio: The Affordable Care Act requires health insurers to submit data on the proportion of premium revenues spent on clinical services and quality improvement. A higher medical loss ratio means that the insurer is spending less money on overhead.\(^ {23}\)
- Dollars spent on reporting of quality measures.\(^ {31}\)

**Global:**
- The percentage of national health expenditures dedicated to administrative spending: based on NHEA data (using the Medicare hospital costs reporting format).\(^ {2,7}\)

**Usability:**

**Relationship to currently used OECD measures**
Administrative costs are represented among the current OECD measures. However, no standard measure to estimate administrative complexity (e.g., burden or time spent on administrative tasks) is currently collected by OECD.

**Closest related measures:**
- Government health administration expenditure related to share of total government expenditure financed by social health insurance or other compulsory schemes, 2014 (or nearest year)\(^ {35}\)
- Health administration expenditure as a share of financing schemes’ total health spending, 2014.\(^ {32}\) However, this measure leaves out administrative costs associated with health services providers (e.g., hospitals or practices of physicians).\(^ {33,34}\)

OECD could consider adopting other example measures of administrative complexity and costs (e.g., medical loss ratio; dollars spent on reporting of quality measures).

**References**


### Table A.16. Disadoption of Ineffective Medical Services

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Disadoption of Ineffective Medical Services</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>This measure construct, depending on its precise specification, could describe the rate at which the health system identifies ineffective treatments and reduces their availability or utilization in the health system.</td>
</tr>
<tr>
<td></td>
<td>• <em>Treatment</em> is defined as medical care assigned to a patient to address an illness or injury. Treatment includes medications and interventions (e.g., procedures, surgeries, behavioral interventions).</td>
</tr>
<tr>
<td></td>
<td>• A disproven treatment is a treatment that has fallen out of favor due to evidence overwhelmingly demonstrating that it is more harmful than beneficial. In the literature, this is often referred to as a <em>medical reversal</em>. Medical reversals occur when new evidence (better-designed, -controlled, and/or -powered than predecessors) contradicts the current standard of care. Examples of disproven treatments include vitamin E for cardiovascular benefits, beta carotene to prevent cancer, estrogen to prevent Alzheimer’s disease, arthroscopic partial meniscectomy for meniscal tears, stenting for stable heart disease, Provenge to slow the progression of prostate cancer, and gentamicin-collagen sponge for infection prophylaxis in colorectal surgery.</td>
</tr>
<tr>
<td></td>
<td>• <em>Abandonment</em> could be defined as discontinuing to prescribe or administer. Alternatively, abandonment could be defined as patients no longer receiving the treatment. The process for actually abandoning standard treatments when new evidence contradicts them also involves societal values, cultural tensions, and politics.</td>
</tr>
<tr>
<td><strong>Importance: Health or social impact</strong></td>
<td><strong>Population impact:</strong></td>
</tr>
<tr>
<td></td>
<td>• <em>National:</em> Failure to abandon disproven therapies has not been measured. A similar concept is low-value care, which costs an estimated $75 billion to $100 billion per year in the United States.</td>
</tr>
<tr>
<td></td>
<td>• <em>Global:</em> Review of evidence between 2003 and 2017 identified 2,017 studies of medical practice, 396 of which were medical reversals. 92% of the studies were conducted in high-income countries, and the remaining 8% in mid- or low-income countries. Other related estimates involve low-value care, which has been measured in Canada, Australia, Europe, Israel, India, and Korea. The World Health Organization estimates that 20–40% of health spending is wasted.</td>
</tr>
<tr>
<td></td>
<td><strong>Clinical impact:</strong></td>
</tr>
<tr>
<td></td>
<td>• <em>Consequences:</em> Patients can incur harm due to side effects of disproven or ineffective medications they are taking. The continued use of medical practices that do not work wastes resources, jeopardizes patient health, and undermines trust in medicine. For example, some women may still take estrogen in an effort to prevent Alzheimer’s disease, but estrogen is known to increase risk of stroke.</td>
</tr>
<tr>
<td></td>
<td><strong>Performance gap:</strong></td>
</tr>
<tr>
<td></td>
<td>• The American Board of Internal Medicine launched the Choosing Wisely campaign to identify treatments that should not be offered.</td>
</tr>
<tr>
<td></td>
<td>• The UK health system has endeavored to disinvest in more than 800 disproven or low-value treatments in recent years.</td>
</tr>
<tr>
<td>Measure Construct</td>
<td>Disadoption of Ineffective Medical Services</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td><strong>Costs:</strong></td>
<td></td>
</tr>
<tr>
<td>National/global:</td>
<td>It is difficult to quantify the cost of ineffective treatment, but costs for low-value care may provide a proxy. In the United States, the costs for the top five overused clinical activities across three primary care specialties (pediatrics, internal medicine, and family medicine) accounted for $6.76 billion annually. 86% of these costs were due to prescribing brand-name drugs instead of generics.</td>
</tr>
</tbody>
</table>

| Importance:       | Value-based care:                         |
| Related U.S.      | Low-value medical services are not cost-effective and carry the opportunity costs of underutilized but more-effective treatments. Institutions could provide higher-value care if they stopped administering costly or harmful therapies soon after they were disproved. |
| Department of Health and Human Services priority topics | Other indirectly related priorities: |
|                   | Health insurance reform: The continued utilization of disproven treatments is a part of health system waste that curtails the availability and affordability of health insurance. |
|                   | Drug pricing: The prices of disproven drugs represent waste and an area in which to improve the cost-effectiveness of the health care system. |

| Validity          | Construct validity:                       |
|                   | Although there is a common definition for medical reversal, there is disagreement over the wisdom of abandoning these treatments. Medical reversals are based on randomized controlled trials (RCTs), the “gold standard” of evidence; however, RCTs also carry concerns over external validity and generalizability, individual patient response not being characterized by the average reported effect, and a number of other design issues. |
|                   | Discriminant validity:                    |
|                   | There could be variation in which treatments are identified as disproven across countries. |
|                   | The utilization of medical treatments is, in part, a function of the characteristics of the population. Consequently, international comparisons of measures of the utilization of any medical treatment, disproven or not, could be confounded by population characteristics, access to care overall, and subpopulations. Additional threats to validity include nation-to-nation variation in baseline treatment practices and variations in standards of care for different countries. |

<p>| Feasibility for international comparisons | Definitions:                               |
|                                          | There is no commonly agreed-upon set of treatments for which the evidence represents a medical reversal. |
|                                          | There is no commonly agreed-upon definition for the concept of abandonment of disproven treatments. |
|                                          | However, there is international interest in the more-general concept of low-value care, as well as how to measure low-value care (which sometimes includes therapies). |</p>
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Disadoption of Ineffective Medical Services</th>
</tr>
</thead>
</table>
| **Data availability:** | • Data describing which medical treatments are disproven would need to be generated via regular systematic reviews of the existing literature surrounding treatments.  
• Data regarding prescription medications and procedures are already collected and reported by many nations. This process could likely be adapted to quantify rates of abandonment of medications and procedures over time.  
  ○ *National*: The United States collects data describing use of medications as part of the Medical Expenditure Panel Survey, in state-level databases through the Healthcare Cost and Utilization Project, through private companies like IQVIA, and via Medicare claims. For procedures, the United States collects and submits data via the National Hospital Discharge Survey (inpatient procedures) and the National Survey of Ambulatory Surgery (outpatient procedures).  
  ○ *Worldwide*: At least 25 countries routinely collect data on prescribed and dispensed medications. At least 17 countries provide data on procedures, including coronary angioplasty and stenting, in the inpatient or outpatient setting.  
| **Data comparability:** | • Differences in how certain data are collected across countries could threaten comparability. For instance, some countries have national prescription registries, while others collect separate data describing reimbursement and prescription. Other countries have good prescription data for specific subpopulations (i.e., those covered by certain insurance programs) but lack data for other subpopulations.  
Regarding procedures, some nations count the number of main procedures only, so there could be undercounting if patients undergo multiple procedures during the same hospitalization or visit, while in other countries overcounting of procedures could occur because procedures are counted without any restriction.  
• Health system characteristics could play a role in the utilization of disproven treatments. Specifically, if the treatments are costly, there could be incentives to utilize them or to avoid disinvesting in them. An example of this is the drug bevacizumab, which received accelerated U.S. Food and Drug Administration approval for metastatic breast cancer. The cost to each patient was $88,000 per year, but the drug was ultimately shown not to improve overall survival.  
• International definitional differences (which would arise if there are not commonly agreed-upon treatments that represent medical reversals or what constitutes an abandonment of the treatment) would also threaten the comparability of the data. |

| **Usability:** | **National:**  
| **Current evidence of use** | • There are existing measures of low-value care that count the incidence of therapies that are thought to be low yield, including ones promulgated through Choosing Wisely in the United States.  
• Additionally, medical reversal evidence is used by insurers or within health care systems to determine which procedures to cover or allow. For instance, high-dose chemotherapy with autologous bone marrow transplantation for advanced breast cancer was thought to be helpful. Many women with advanced breast cancer demanded that their health insurance plans cover the procedure and sued the companies in court. |
Disadoption of Ineffective Medical Services

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Disadoption of Ineffective Medical Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>when they did not. Later, a definitive RCT found no benefit, which lead to insurers halting coverage for the procedure.22</td>
<td></td>
</tr>
</tbody>
</table>

Global:

- There is no currently accepted measure for abandonment of disproven therapies. However, there are existing measures of low-value care that have been promulgated in the United Kingdom23 and other countries.24 For example, the United Kingdom National Institute for Health and Clinical Excellence identifies low-value activities that could be stopped—for example, because they are not clinically effective (and therefore not cost-effective), have a poor risk-benefit profile, or are not supported by adequate evidence. This work often leads to disinvestment in the disproven treatments.10

Usability:

Relationship to currently used Organisation for Economic Co-operation and Development (OECD) measures

The abandonment of disproven treatments is not among the current OECD measures. However, OECD expressed interest in low-value care in a recent publication.28

Closest related measures:

- **Caesarean section rates:** This measure describes the percentage of live births that occur via caesarean section. Caesarean sections can be lifesaving and necessary, but the procedure is known to be overutilized in some contexts, and the overutilization is related to financial incentives, liability concerns, and other factors. Consequently, the international comparison of caesarean section rates reflects, in part, the inappropriate utilization of the procedure.26 Because this measure only partly reflects inappropriate use, the national vital statistics data used for this measure would not be relevant in describing the abandonment of low-value treatments.26

- **Total volume of antibiotics prescribed:** The harms of antibiotics may outweigh the benefits for patients with conditions that are not likely attributable to bacterial infection. OECD tracks the volume of antibiotics so that nations with high prescription rates can consider taking action to achieve lower prescribing rates.

- **Benzodiazepine prescriptions for older adults:** Benzodiazepines can cause falls for older adults, meaning that potential harms tend to outweigh benefits. The OECD measure is designed to influence nations to take action to reduce inappropriate prescribing of potentially harmful therapies.27

References

17. OECD. **OECD Health Statistics 2019: Definitions, Sources and Methods, Surgical Procedures (shortlist).** 2019.

Disadoption of Ineffective Medical Services 161
Healthy Days at Home

Table A.17. Healthy Days at Home

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Healthy Days at Home</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Healthy days at home (HDAH) is a population-based measure that attempts to quantify how well health care organizations keep people alive, healthy, and not utilizing health care services. This measure is one of a select few population-based measures of the health care system that attempts to quantify health, as opposed to the lack thereof.</td>
</tr>
<tr>
<td></td>
<td>- The measure is calculated by subtracting the total number of days out of the year that the individual was not alive or was utilizing acute or postacute health care services from 365 days and then calculating the mean number of HDAH for the population of interest.</td>
</tr>
<tr>
<td></td>
<td>- Acute or postacute health care services include inpatient days, observation days, skilled nursing facility days, home health visits, outpatient emergency department visits, inpatient psychiatry days, inpatient rehabilitation days, and days of long-term hospital stays.</td>
</tr>
<tr>
<td></td>
<td>- Each time the individual utilizes acute or postacute health services counts as 1 day, though some of these services will generally not take the whole day.</td>
</tr>
<tr>
<td></td>
<td>- Mortality days are counted as the number of days remaining in the calendar year after the date of death. If an individual dies on December 28, mortality days = 3. If an individual is alive for the entire year, mortality days = 0.</td>
</tr>
<tr>
<td></td>
<td>- Alternative definitions or specifications of the measure consider that home health visits are neither as undesirable nor as costly as the other health services in the measure and reduce the weight of each home health visit.</td>
</tr>
<tr>
<td></td>
<td>- Population: The measure was developed for the full U.S. Medicare population; alternative formulations consider subsets of the Medicare population (e.g., those with a cancer diagnosis, those with multiple chronic conditions).</td>
</tr>
<tr>
<td>Importance: Health or social impact</td>
<td>Population impact:</td>
</tr>
<tr>
<td></td>
<td>HDAH is a positive measure, describing the lack of health care service utilization; it is intended to assess overall population health as opposed to health outcomes following an acute event. Healthy time at home is an outcome that patients value and reflects a very broad understanding of health and quality of care.</td>
</tr>
<tr>
<td></td>
<td>Clinical impact:</td>
</tr>
<tr>
<td></td>
<td>- Measures such as HDAH seek to maximize patient experiences and reorient the focus away from procedures and toward the prevention of needs that require a health care response.</td>
</tr>
<tr>
<td></td>
<td>- Current measures of health care quality have been criticized for “looking under the streetlight” (i.e., focusing attention on conditions that are easily measured or health care interventions that are considered widely appropriate, to the detriment of other aspects of care that are not measured). Additionally, focusing only on acute care events that require hospitalization does not credit health care systems that prevented that hospitalization in the first place.</td>
</tr>
<tr>
<td>Measure Construct</td>
<td>Healthy Days at Home</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>• Most individuals at the end of life would prefer to spend their last days at home; however, most do not.  This measure incorporates more nuance than mortality alone.</td>
<td></td>
</tr>
<tr>
<td>Performance gap:</td>
<td>• In a study of a sample of Medicare beneficiaries aged 65 and older, individuals in the worst-performing market had 5.8 fewer adjusted HDAH on average relative to the national mean; those in the best-performing market had 5.0 more HDAH on average relative to the national mean. HDAH is also strongly associated with beneficiary age, gender, and chronic conditions.</td>
</tr>
<tr>
<td>Costs:</td>
<td>• Health care spending in the United States is projected to grow 5.5% annually, reaching 19.4% of gross domestic product by 2027. Maximizing HDAH would presumably lower costs via reduction in unnecessary health care utilization.</td>
</tr>
</tbody>
</table>

**Importance:**

<table>
<thead>
<tr>
<th>Related U.S. Department of Health and Human Services (HHS) priority topics</th>
<th>Value-based care:</th>
</tr>
</thead>
<tbody>
<tr>
<td>The HHS priority to transform our health care system from one that pays for procedures and sickness to one that pays for health is reflected in this measure. The measure emphasizes efforts to improve health and prevent the utilization of health care services.</td>
<td></td>
</tr>
</tbody>
</table>

**Validity**

<p>| Construct validity: | • The measure is built on the intuitive premise that if consumers have access to care, their lack of utilization of care is indicative of their health. The measure was developed for use among a Medicare population, which presumably has access to care. It could be difficult to interpret this measure across countries with different health care payment systems and possible differences in access to care. |
| • The measure is new and therefore has not been extensively studied to show that its variation accurately measures health. |
| • One concern with the measure is the level of variation and whether the variation captures health variation across all age groups. For the full population, the utilization of health services will be close to 0, and among that population the measure may not reflect actual health. |
| Internal validity: | • The appeal of the measure is its simplicity and intuition. However, it is built on the premise of equal and ready access to health care. This implies that if access is not equal or ready, the measure may not accurately reflect the quality of health care. |
| • The measure is new and therefore has not been extensively studied to show that its variation accurately measures the quality of health care. |
| Discriminant validity: | • Variation in the measure might reflect use of health care services but not HDAH if a population has less access to affordable health care (e.g., an individual who does not seek treatment for a health concern because of costs would appear to be healthy using this measure). |
| • The utilization of health services is, in part, a function of the characteristics of the population. Consequently, international comparisons of the measure could be confounded by population characteristics overall and by subpopulations. |</p>
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Healthy Days at Home</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feasibility for international comparisons</td>
<td>Definitions:</td>
</tr>
<tr>
<td></td>
<td>• The measure is new (2020); the original publication provided a precise definition.²</td>
</tr>
<tr>
<td></td>
<td>• Although the measure has been applied to a specific population (Medicare, aged 65+)² and has been further applied to more-narrow populations,³ it could, in theory, be applied to broader populations.</td>
</tr>
<tr>
<td></td>
<td>• There are alternative definitions or specifications of the measure that could be explored (e.g., weighting of home health visits).²</td>
</tr>
<tr>
<td></td>
<td>Data availability:</td>
</tr>
<tr>
<td></td>
<td>• The data used in generating the measure are Medicare administrative claims data, the use of which has been shown to be technically feasible.² Gathering comparable data elements for people under 65 in the United States would likely be more difficult.</td>
</tr>
<tr>
<td></td>
<td>• Data describing utilization of health services that are used to generate the measure are common to all administrative claims data sources.</td>
</tr>
<tr>
<td></td>
<td>o National: Medicare, Medicaid, and state-level all-payer claims databases.</td>
</tr>
<tr>
<td></td>
<td>o Worldwide: As evidenced from hospital admissions data currently employed by the Organisation for Economic Co-operation and Development (OECD), nearly all OECD countries have administrative data describing the utilization of health services;¹⁰ however, it is not clear whether all of the elements required to generate comparable measures would be readily available (e.g., date of service, utilization types).</td>
</tr>
<tr>
<td></td>
<td>Data comparability:</td>
</tr>
<tr>
<td></td>
<td>• The specific variables used in constructing the measure are simple, which facilitates its construction and comparability.</td>
</tr>
<tr>
<td></td>
<td>• Health system characteristics play a role in the utilization of health services. Cost-sharing influences the utilization of health services¹¹ so that utilization might not reflect health. Consequently, the utilization of health services will not be comparable between countries or populations that do and do not require cost-sharing.</td>
</tr>
<tr>
<td></td>
<td>Subgroup analysis and equity:</td>
</tr>
<tr>
<td></td>
<td>• In some countries, data describing health services utilization might not be available (or uniformly available) for the entire population. For instance, some countries have national administrative databases describing health services utilization, while others have data for specific subpopulations (i.e., those covered by certain insurance programs) but lack data for other subpopulations.¹²</td>
</tr>
<tr>
<td></td>
<td>• Additionally, if the cost-sharing requirements for certain subgroups differ, the measure might not be comparable.</td>
</tr>
<tr>
<td></td>
<td>• This measure could be stratified by the presence of chronic conditions, which could necessitate more health care utilization and fewer HDAH.²</td>
</tr>
<tr>
<td>Usability: Current evidence of use</td>
<td>National:</td>
</tr>
<tr>
<td></td>
<td>• The Medicare Payment Advisory Commission commissioned the development of the measure.²</td>
</tr>
<tr>
<td></td>
<td>• Similar measures, such as health-related quality of life (HRQoL), have been developed by the Centers for Disease Control and Prevention (CDC) and other organizations and are currently used.¹³</td>
</tr>
</tbody>
</table>
Measure Construct | Healthy Days at Home
--- | ---
**Global:**
- This measure has not been used on a global population; however, there have been multiple studies examining similar measures across the globe. For example, a New Zealand study measured days alive and out of the hospital and showed improvement over time.\(^\text{14}\)
- Measures based on the CDC's HRQoL measure are currently used internationally. For example, the EuroQol EQ-5D measure has been validated and is used across Europe to measure generic health status.\(^\text{15}\)

**Usability:**
**Relationship to currently used OECD measures**
HDAH is not among the current OECD measures. In terms of measures that positively describe health, self-rated health is the closest. There are many measures describing the prevalence of negative health outcomes that, if aggregated and reversed, could be comparable with HDAH. OECD also captures measures related to patient satisfaction with doctors involving that patient in decisions about care, which is tangentially related to HDAH.

**Closest related measures:**
- Self-rated health: Respondents are asked to self-report their health.\(^\text{16}\) This measure is intuitively similar to HDAH, though HDAH is not self-reported. Consequently, HDAH is likely more objective but might also be less patient centered.
- Avoidable hospital admissions: Measures the number of hospital admissions with a primary diagnosis of asthma, chronic obstructive pulmonary disease, and congestive heart failure with the idea that admissions for these diagnoses could be avoidable with better primary care.\(^\text{16}\)
- “Patients reporting having been involved in decisions about care or treatment by their regular doctor” and “Patients reporting having been involved in decisions about care or treatment by any doctor.”\(^\text{10,17}\)

**References**
Availability of Emergency Medical Services to Prevent Opioid Death

Table A.18. Availability of Emergency Medical Services to Prevent Opioid Death

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Availability of Emergency Medical Services to Prevent Opioid Death</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Emergency medical services (EMS) are a potential mechanism for provision of naloxone, an opioid antagonist that reverses the effects of opioid overdose. Naloxone is increasingly being used by police officers, emergency medical technicians, and nonemergency first responders to reverse opioid overdoses.¹</td>
</tr>
<tr>
<td>• EMS: Denotes the pre-hospital component of emergency services and includes out-of-hospital EMS practitioners, such as emergency medical technicians and paramedics (i.e., first responders); initial out-of-hospital treatment; and transport, by air or by ground, to a hospital.²</td>
<td></td>
</tr>
<tr>
<td>• The availability of EMS could reduce opioid-related deaths because emergency medical technicians might be able to administer naloxone, which rapidly reverses an opioid overdose; EMS could also be able to coordinate with other treatment services to prevent future overdoses.³</td>
<td></td>
</tr>
<tr>
<td>• To measure inpatient emergency department (ED) treatment for opioid overdose, cases are defined as having an International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) code of one or more discharge codes: 965 (.00, .01, .02, .09), E8580 (.0, .1, .2).⁴</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Importance: Health or social impact</td>
<td></td>
</tr>
<tr>
<td>Population: In 2018, 46,802 Americans died as a result of an opioid overdose, and an estimated 1.7 million people in the United States suffered from substance use disorders related to prescription opioid pain relievers; 526,000 suffered from a heroin use disorder (not mutually exclusive).⁵ Emergency medical service administration of naloxone can be used as a proxy for estimating opioid overdoses.³ From July 2016 to September 2017, there were 142,557 ED visits for suspected opioid overdose.⁶</td>
<td></td>
</tr>
<tr>
<td>Worldwide, roughly 34 million people used opioids (synthetic narcotic analgesic) and about 19 million used opiates (natural narcotic analgesic) in 2016; roughly 27 million suffered from opioid use disorders (OUDs) in 2016. The majority of people dependent on opioids used illicitly cultivated and manufactured heroin, but an increasing proportion used prescription opioids. Approximately 118,000 people died of OUDs in 2015.⁷ EMS vary across OECD countries. Decisions on which providers can provide emergency aid, such as naloxone, are determined at a national or subnational level.⁸</td>
<td></td>
</tr>
<tr>
<td>Meaningful clinical impact: The longer the delay between recognition of an overdose and calling emergency services, the greater the risk of</td>
<td></td>
</tr>
</tbody>
</table>
severe damage or death. This problem is exacerbated in areas with poor access to EMS. EMS responses for suspected opioid overdose reduce the likelihood that the overdose will be fatal and also increase the likelihood that an opioid user will be referred to treatment. Despite data suggesting that heroin and opioid misuse is stabilizing, ED use continues to increase.

**Performance gap:** EDs do not typically have consistent protocols or guidance for what to do after an opioid overdose is resolved.

**Costs:** The Society of Actuaries estimated that the opioid epidemic cost the United States $179 billion in 2018; $72.6 billion of that cost was attributable to mortality costs (lost lifetime earnings). Despite data suggesting that heroin and opioid misuse is stabilizing, ED use continues to increase.

- National: Caring for all opioid overdoses in the ED costs more than $632 million annually.
- Global: Estimating costs is dependent on whether the measure captures the costs of medication, associated health services, or both.

**Importance:** Related U.S. Department of Health and Human Services (HHS) priority topics

The opioid crisis is an HHS priority topic, and this measure directly addresses the goal of reducing deaths from opioid overdoses, as well as improving access to naloxone and increasing referrals to additional treatment programs.

**Validity**

Measurement of the time of response to opioid overdose EMS calls or the availability of EMS could be biased by individuals’ unwillingness to call EMS during an overdose. Examples include an ability to administer naloxone and supervise recovery without EMS support, fear of arrest or incarceration (even in the presence of Good Samaritan laws), distrust of police and the legal system, avoidance of costly EMS interventions, and an unwillingness to draw unwanted attention with an EMS call and response. Measures such as the percentage of EMS calls for opioid overdose that do not result in overdose death can vary based on cultural differences by country in the factors described above. Greater attention to quality control of EMS records will enhance the internal validity of these data. Linking data to toxicology results and patient outcomes will further validate the data.

**Feasibility for international comparisons**

**Definitions:** There is no common definition, nor is there a well-defined measurement, for this concept. A measure could be constructed using a subset of existing data on EMS availability, limited to responses for opioid overdoses (e.g., “time of response” data for EMS runs to opioid overdose calls, percentage of EMS runs for opioid overdoses that result in death). However, even within the wider field of prehospital emergency care, there are very few performance metrics that are generalizable or widely used. Although there is interest in measurement of this aspect of the health care system, development of such measures is in the early phases.

**Data availability:** The United States lacks strong opioid abuse and overdose surveillance systems. The National Survey of Substance Abuse Treatment Services (N-SSATS) collects data on the location, characteristics, and use of treatment facilities at the state level. It tracks the number and location of facilities that provide emergency medical treatment. In the United States, there have been several national and state-based studies of the distribution of EMS and performance;
however, these metrics are varied and generally are limited to availability of EMS in general rather than specifically for opioid overdoses.

Some countries collect more-robust data from EMS providers and systems; however, others have less-developed data systems.21–26 Additionally, some countries do not have EMS systems in place.27

EMS departments could provide administrative data for this measurement concept, but it is unclear how many are collecting all of the elements that would be needed to fulfill this measurement concept.

Data comparability: Without defined measurements, it is difficult to assess data comparability across countries. Additionally, the organizational structure and functions of EMS vary from country to country, which will also have an impact on data comparability.

Subgroup analysis: There is evidence of disparities in naloxone administration by EMS providers in rural communities.28 Generally, EMS response times and availability vary by urbanicity; stratifying results by rural, urban, and suburban regions would provide additional insights for this topic.29

Usability: Current evidence of use

This measure is not currently in use on a global or national level.16

- The Centers for Disease Control and Prevention tracks ED visits for suspected opioid overdoses.30
- The Australian Institute of Health and Welfare tracks the rate of emergency admissions for opioid dependence by age, socioeconomic group, urbanicity, and degree of urgency.31

Usability: Relationship to currently used OECD measures

OECD currently collects the following data for distantly related concepts, but none approach the information that is needed for this concept. These measures would be helpful in establishing the extent of opioid abuse and the risk for overdose deaths but not this specific measurement concept.

- availability of analgesic opioids
- opioid-related deaths; overall volume of opioids prescribed
- proportion of population who are chronic opioid users

References

### Table A.19. Clinician Workforce Who Can Prescribe Medication-Assisted Treatment or Naloxone

<table>
<thead>
<tr>
<th>Measure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>Medication-assisted treatment (MAT) is the most effective medical treatment for opioid use disorder (OUD) and other substance use disorders (SUDs), including alcohol. Three medications have been approved by the U.S. Food and Drug Administration (FDA) to treat OUD: methadone, buprenorphine, and naltrexone. Additionally, the FDA has approved naloxone, an injectable drug used to prevent opioid overdose. Historically, the primary treatment for OUD was methadone, which was only administered by accredited opioid treatment programs (OTPs). However, barriers to methadone treatment remain, including locations of OTPs, provider shortages, and treatment waitlists. Naltrexone can be prescribed by any licensed physician, but patients must be abstinent for 7 days prior to treatment. Since the Drug Addiction Treatment Act of 2000, buprenorphine can be prescribed in any location by qualified physicians who have applied and received a Substance Abuse and Mental Health Services Administration (SAMHSA) waiver. MAT is the use of medications, in combination with counseling and behavioral therapies, to provide a “whole-patient” approach to the treatment of SUDs. MAT is primarily used in the treatment of addiction to opioids but is also used for alcohol dependence. Sometimes referred to as “opioid abuse or dependence” or “opioid addiction,” OUD is a problematic pattern of opioid use that causes significant impairment or distress. SUDs occur when the recurrent use of alcohol or drugs causes clinically significant impairment, including health problems; disability; and failure to meet major responsibilities at work, school, or home.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Importance: Health or social impact</th>
<th>Population impact:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>National:</strong> The misuse of and addiction to opioids is a major public health crisis in the United States, causing over 1.6 million years of life lost annually, more than the burden attributable to homicides, hypertension, HIV/AIDS, and pneumonia. As of 2018, 42.3% of opioid treatment centers offered MAT. Despite the prescribing options available, estimates suggest that only 20–40% of people with OUD receive MAT. Much of this treatment gap is attributed to the workforce, specifically (1) insufficient prescriber training, (2) lack of institutional support, (3) poor care coordination, (4) provider stigma, (5) inadequate reimbursement, and (6) burdensome regulatory procedures. Consequently, the size of the clinician workforce who can prescribe MAT is a critical variable in effectively addressing the ongoing opioid crisis. <strong>Global:</strong> While the opioid epidemic is most pronounced in the United States, more and more countries are seeing growing numbers of OUDs and overdose-related deaths. The United Nations 2017 report on drug abuse reports that more than 200 million people, or 4.2% of the world population, are addicted to illegal drugs.</td>
<td></td>
</tr>
</tbody>
</table>
### Measure: Clinician Workforce Who Can Prescribe Medication Assisted Treatment or Naloxone

#### Clinical impact:
- Being diagnosed with an OUD increases the risk of early death by a factor of 20.\(^1\) MAT is the most effective treatment for OUD.\(^1\)

#### Performance gap:
- In 2018, 2 million people in the United States were diagnosed with an OUD.\(^16\) However, the capacity to treat OUD remains at less than 50% of the total number of people diagnosed with OUD.\(^4\) Only 28% of opioid overdose survivors are linked to MAT despite evidence of its effectiveness.\(^17\)
- Rural populations are particularly underserved—60% of small, nonmetropolitan county residents live in an opioid treatment shortage area.\(^18\)

#### Costs:
- **National:** In the United States, recent estimates suggest that the costs of the opioid epidemic exceed $500 billion or 2.8% of gross domestic product. Wide disparities in Medicaid coverage of MAT across states have implications for access because OUD predominantly affects low-income populations.
- **Global:** Because the global opioid epidemic is in its nascent stage, the global costs have not yet been realized or projected. But there is evidence that drug companies are applying tactics used in the United States to increase opioid sales in Asia and Europe.\(^19\)

### Importance: Related U.S. Department of Health and Human Services (HHS) priority topics

#### Opioid crisis:
- HHS is committed to ending the crisis of opioid addiction and overdose in the United States. One point of HHS’s 5-point strategy to combat the opioid crisis is “better addiction prevention, treatment, and recovery services.” The clinician workforce is critical to providing better treatment and recovery services.

#### Other related priorities:
- **Value-based care:** Integrating MAT into general medical practices could increase access, helping to transform our health care system from one that pays for procedures and sickness to one that pays for outcomes and health, but low reimbursement levels disincentivize providers from offering MAT.\(^20\)

### Validity

#### Construct validity:
- The measure is built on the premise that a greater number of authorized MAT prescribers will increase the number that can access, and access quickly, effective treatments for OUD.\(^12\)
- However, given the effectiveness (or potential effectiveness) of other public health approaches to reduce the prevalence of SUD in the population,\(^21\) increasing the number of authorized MAT prescribers might not be a necessary condition for the improvement of health as it relates to OUD.

#### Internal validity:
- The capacity of the health care system to deliver proven, effective treatments is a critical component of the quality of health care as available to the consumer.
### Measure Construct

**Clinician Workforce Who Can Prescribe Medication Assisted Treatment or Naloxone**

#### Discriminant validity:
- Greater MAT prescribing capacity might reflect the greater population need rather than the quality of the health care system. Consequently, the measure could be biased against countries where OUD and other SUDs are less prevalent. Bias in the measure would be prevented by accounting for the prevalence of OUD in the population.
- Greater prevalence of diagnosed OUD and other SUDs might reflect diagnostic differences rather than actual use and abuse of substances.
- Additionally, greater prevalence of OUD and other SUDs could reflect elements of the public health system (e.g., laws) rather than the health care system.

#### Feasibility for international comparisons

**Definitions:**
- The World Health Organization’s official recommendation for the treatment of OUD is “the use of a range of treatment options for opioid dependence which include psychosocial support, opioid maintenance treatments such as methadone and buprenorphine, supported detoxification and treatment with opioid antagonists such as naltrexone.”
- Each country has different regulations on who can prescribe MAT. For example, in Canada and Europe, buprenorphine and methadone can be prescribed by a pharmacist, but in the United States the majority of prescriptions are by specialist physicians. Any international differences in the definition of OUD and in who is authorized to prescribe MAT could influence the needs and capacities of different countries.

**Data availability:**
- **National:**
  - Capacity needs could be assessed with regularly updated surveys on drug use, such as the National Survey on Drug Use and Health (NSDUH).
  - The National Survey of Substance Abuse Treatment Services (NSSATS) is an annual survey conducted by SAMHSA that captures detailed information on all known substance abuse treatment facilities throughout the United States, including OTPs.
  - SAMHSA maintains a database describing all physicians that have received waivers to prescribe buprenorphine.
- **Global:**
  - MAT is available in 80 countries around the world.
  - Many European countries and other OECD countries regularly field surveys like the NSDUH that attempt to describe licit and illicit drug use in their populations.
  - In many European and other OECD countries, methadone treatment is available via dedicated facilities (like OTPs) or via primary care providers or others. Consequently, in many locations the capacity of authorized prescribers could be assessed with administrative data.

**Data comparability:**
- Governmental and public health system characteristics could play a role in the capacity needs of each country. For instance, in some countries where drugs have been decriminalized (e.g., Portugal), overdoses and infections have been reduced, reducing the needs for physicians authorized to prescribe MAT.
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Clinician Workforce Who Can Prescribe Medication Assisted Treatment or Naloxone</th>
</tr>
</thead>
</table>
| • Health system characteristics might play a role in the capacity needs of each country. For instance, in many European and other OECD countries, OTPs are not necessary because primary care physicians and others are authorized prescribers.  
• International definitional differences (which would arise if there were not commonly agreed-upon standards for OUD diagnoses) would also threaten the comparability of the data. | |
| Usability: Current evidence of use | National:  
• There is no currently accepted measure for the clinician workforce who can prescribe MAT.  
• The National Quality Forum measures the initiation and engagement of alcohol and other drug abuse or dependence treatment, but this measure does not involve the workforce size or capacity.  
• The Joint Commission uses the percentage of those with an acute opioid disorder (AOD) diagnosis who receive or refuse a MAT or referral at hospital, which is comparable to the National Quality Forum’s AOD treatment provided or offered at discharge from inpatient hospitalization.  
• The National Survey of Substance Abuse Treatment Services collects data on gross measures of organizational structure, such as the percentage of referrals, but does not measure a program’s ability to prescribe medications. Nor is implementation measured, which could lead to inflated inferences about how therapies are used.  
Global:  
• There is no currently accepted measure for the clinician workforce who can prescribe MAT.  
• The Canadian Institute for Health Information does measure the hospital stays for harm caused by substance use, but this measure does not capture the workforce size or capacity.  
• The Australian Institute of Health and Welfare tracks referral to treatment by treatment source, such as family, law enforcement, or a health professional. It also tracks the different types of treatment to which patients are referred, such as whether MAT is combined with counseling. |
| Usability: Relationship to currently used OECD measures | A measure of the clinician workforce who can prescribe MAT is not among the current OECD measures.  
Closest related measures:  
• Opioid use: These measures describe the availability of analgesic opioids and the rate of opioid-related deaths. Therefore, these measures describe the population need for a clinician workforce who can prescribe MAT.  
• Opioid-related deaths: The European Monitoring Centre for Drugs and Drug Addiction collects these data and shares them with OECD.  
• Safe primary care prescribing: OECD tracks the volume of opioids prescribed and the proportion of chronic opioid users. Again, these measures describe the population need for a clinician workforce who can prescribe MAT. |


24. U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality. *National Survey on Drug Use and


Access to Opioid Treatment Centers

Table A.20. Access to Opioid Treatment Centers

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Access to Opioid Treatment Centers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Access to treatment facilities (for substance use disorders [SUDs], specifically opioid use disorders [OUDs]) is defined as the ability to offer same-day or walk-in appointments in the outpatient setting and the ability to optimize resources, refer patients, and support them if the necessary level of care is not immediately available.¹</td>
</tr>
</tbody>
</table>
| Importance: Health or social impact | Population impact:  
- National incidence: Of the 19.9 million adults who needed treatment for an SUD in 2016 (inclusive of but not limited to opioid use), only 2.1 million (10.8%) received addiction treatment.² Structural barriers and stigma inhibit access.³ Client outreach varies widely between states, ranging from 47% in Wisconsin to 85% in Montana and Oklahoma.⁴  
- Although approximately 33,000 physicians have obtained the waiver needed to prescribe medications on U.S. Food and Drug Administration–approved (FDA-approved) Schedules II, IV, or B for treating OUDs, fewer than half offer this treatment to patients.⁵  
- Nearly all U.S. states do not have sufficient treatment capacity to provide medication-assisted treatment (MAT) to all patients with an OUD.⁶ |
|                     | Clinical impact:  
- Increasing access to medications for OUD treatment is widely acknowledged to be a key strategy for addressing the opioid epidemic.⁷⁻⁸ The use of OUD medications is associated with reductions in opioid use, withdrawal and craving, infectious disease transmission, and treatment dropout.¹⁰⁻¹⁵  
- Outpatient treatment facilities are particularly important for the treatment of SUDs because only 13% of individuals who receive treatment do so in a private physician’s office.¹⁶ |
|                     | Performance gap:  
- Almost all states and the District of Columbia (96%) had opioid abuse or dependence rates higher than their buprenorphine treatment capacity rates; 19 states had a gap of at least 5 per 1,000 people.⁶  
- A review of national survey data found that less than 20% of persons with OUD have used any SUD service in a given year.¹⁷ A separate analysis found that OUD medication treatment capacity only has the potential to cover about 61% of those in need.⁶ |
|                     | Costs:  
- National: High out-of-pocket costs and access to a provider are both barriers to accessing treatment.¹⁸ Individuals accessing opioid treatment tend to have comorbidities that increase health care utilization and costs. The costs of treatment alone do not account for the higher costs.¹⁹ |
<p>| Importance: Related U.S. Department of Health and Human Services (HHS) priority topics | Opioid crisis: The availability of opioid treatment centers is directly related to this HHS priority and contributes to a multifaceted response to addressing the opioid crisis. |</p>
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Access to Opioid Treatment Centers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Other related priorities:</strong></td>
<td><strong>Construct validity:</strong></td>
</tr>
<tr>
<td><em>Value-based care:</em> The HHS priority to transform our health care system from one that pays for procedures and sickness to one that pays for outcomes and health relates to access to opioid treatment centers because of the increased access to long-acting MAT and assignment of case managers to addicted patients. These additional steps have lessened the need for emergency medical treatment of opioid overuse.*</td>
<td></td>
</tr>
<tr>
<td><em>Value-based care:</em> The HHS priority to transform our health care system from one that pays for procedures and sickness to one that pays for outcomes and health relates to access to opioid treatment centers because of the increased access to long-acting MAT and assignment of case managers to addicted patients. These additional steps have lessened the need for emergency medical treatment of opioid overuse.*</td>
<td></td>
</tr>
<tr>
<td><strong>Validity</strong></td>
<td><strong>Construct validity:</strong></td>
</tr>
<tr>
<td>• The availability of opioid treatment centers has high construct validity because it captures the presence or absence of treatment facilities. Access to opioid treatment centers is sometimes measured by utilization of services, which leaves room for more interpretation and somewhat lowered construct validity.</td>
<td></td>
</tr>
<tr>
<td><strong>Internal validity:</strong></td>
<td></td>
</tr>
<tr>
<td>• Because access is a component of quality of care, a measure of access to opioid treatment centers would have high internal validity as a measure of care quality.</td>
<td></td>
</tr>
<tr>
<td><strong>Feasibility for international comparisons</strong></td>
<td><strong>Definitions:</strong></td>
</tr>
<tr>
<td>• It is challenging when estimating drug treatment access to isolate drug-specific data (e.g., treatment for OUD) from broader data sets. Additional modeling could be required if data from international sources are going to be used to estimate drug-specific expenditures.</td>
<td></td>
</tr>
<tr>
<td><strong>Data availability:</strong></td>
<td></td>
</tr>
<tr>
<td>• National accountability and reporting systems might record public but not private expenditure data on drug treatment, which can be linked to access. Additional research would need to be done using claims data.</td>
<td></td>
</tr>
<tr>
<td><strong>Data comparability:</strong></td>
<td></td>
</tr>
<tr>
<td>International definitional differences and approaches to isolating data specific to opioid treatment center access could threaten the comparability of the data.</td>
<td></td>
</tr>
<tr>
<td><strong>Usability:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Current evidence of use</strong></td>
<td></td>
</tr>
<tr>
<td>Current evidence of the use of measures related to accessing opioid treatment centers includes the following (not exhaustive):</td>
<td></td>
</tr>
<tr>
<td>• Access to opioid treatment is measured by the National Survey on Drug Use and Health (NSDUH) by the number of people aged 12 or older who needed substance abuse treatment in the past year (inclusive of opioid use) and the number who received any or specialty treatment.</td>
<td></td>
</tr>
<tr>
<td>• There is no currently accepted measure for access to opioid treatment. The NSDUH measures the need for and receipt of specialty substance abuse treatment in the past year but does not specify whether specialty treatment includes MAT.</td>
<td></td>
</tr>
<tr>
<td>• A validated instrument based on World Health Organization policy guidelines identified barriers to treatment in 11 European countries.</td>
<td></td>
</tr>
<tr>
<td><strong>Usability:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Relationship to currently used Organisation for Economic Co-operation and Development (OECD) measures</strong></td>
<td></td>
</tr>
<tr>
<td>Access to opioid treatment is not among the current OECD measures.</td>
<td></td>
</tr>
<tr>
<td><strong>Closest related measures:</strong></td>
<td></td>
</tr>
<tr>
<td>• OECD collects data on opioid access and related deaths but not access to treatment. In its reports, OECD urges that MAT should be part of any long-term comprehensive treatment and rehabilitation program. MAT is available in all OECD countries except Japan and Korea.</td>
<td></td>
</tr>
</tbody>
</table>
References

22. U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration. *Key Substance Use and Mental Health Indicators in the United States: Results from the*


The Percentage of Patients with a Follow-Up Visit Within Four Weeks of Starting an Opioid for Chronic Pain

Table A.21. The Percentage of Patients with a Follow-Up Visit Within Four Weeks of Starting an Opioid for Chronic Pain

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>The Percentage of Patients with a Follow-Up Visit Within Four Weeks of Starting an Opioid for Chronic Pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td><em>Chronic pain</em> is typically defined as pain that lasts more than 3–6 months and is present on most days or every day. Chronic pain affects a significant proportion of the population in the United States and is associated with poor physical and mental health outcomes and risk of opioid use disorder (OUD), as opioids are frequently prescribed as a primary treatment for chronic pain.(^1) Increasing rates of opioid prescribing over the past 15 years have contributed significantly to increasing rates of OUD and overdose.(^2) In 2016, the Centers for Disease Control and Prevention (CDC) released its guidelines for prescribing opioids for chronic pain in the United States in order to encourage careful and selective use of opioid therapy for pain management.(^1) These guidelines included three groups of recommendations pertaining to how to determine when to prescribe opioids for chronic pain, how to determine dose and duration and when to follow up and discontinue opioids, and how to address the risks and harms of opioid use. In the second group of recommendations, the CDC advised (Recommendation 7) that &quot;clinicians should evaluate benefits and harms within 1 to 4 weeks of starting opioid therapy for chronic pain or of dose escalation.&quot;(^1) The CDC also developed quality improvement measures to incentivize providers to follow the new guidelines, including an assessment of the recommendation for follow-up within four weeks of starting opioid therapy for chronic pain in order to evaluate the benefits and harms of continued opioid therapy.(^3) <em>Opioid medication:</em> The CDC annually updates a list of all opioids, with associated National Drug Code, generic name, and product name, which may be used to match with electronic health records (EHRs).(^3) This measure is limited to patients 18 years and older.(^3) <em>Numerator:</em> The number of patients with a new opioid prescription (within the previous 45 days) for chronic pain with an in-person follow-up visit with the prescribing clinician within four weeks.(^3) <em>Denominator:</em> The number of patients in an outpatient panel of patients 18 years of age or older prescribed an opioid for chronic pain who had no opioid prescription in the previous 45 days.(^3) <em>Exclusions:</em> Exclusion criteria include cancer patients, palliative care, and end-of-life care. Additional exclusion criteria are (1) drugs not typically used in outpatient settings and (2) buprenorphine products indicated for medication-assisted treatment (MAT) to treat OUD.(^3)</td>
</tr>
<tr>
<td>Measure Construct</td>
<td>The Percentage of Patients with a Follow-Up Visit Within Four Weeks of Starting an Opioid for Chronic Pain</td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Importance: Health or social impact</td>
<td>Population impact:</td>
</tr>
</tbody>
</table>
|                                       | National: In the United States, research has shown that 11–40% of the adult population experiences chronic pain, for which opioids are commonly prescribed long term (3–4% of the adult population receives long-term opioid therapy).  

Global: Chronic pain is estimated to impact 20% of adults worldwide, and low-back pain is the leading cause of age-standardized years of lifetime disability worldwide. Prescriptions for opioids to treat chronic noncancer pain have increased over time, which has coincided with an increase in opioid misuse, opioid abuse, and opioid-related deaths. |
<p>| Meanings clinical impact:             | The recommendation for follow-up after starting opioid medications is supported by type 4 evidence: clinical experience and observations, observational studies with important limitations, or randomized clinical trials with several major limitations. If the visits are effective, they could result in reductions in adverse events, such as overdose and OUDs, as well as other improved outcomes, such as proper dosing. |
| Performance gap:                      | The United States and a few other developed countries consume the vast majority of the world’s opioids. Within the United States, there is variation in the frequency of opioid prescriptions between states. Currently, 15 OECD countries have clinical opioid guidelines in place, most of which are focused on treating patients with chronic pain. However, adherence to these guidelines, as well as specific information (such as the recommendation of a follow-up visit within 4 weeks of beginning a new opioid prescription), is not uniformly captured, and thus it is difficult to assess whether a performance gap exists. |
| Costs:                                | The cost of chronic pain in the United States is estimated at $560 billion per year in direct medical costs, lost productivity, and disability programs. The Society of Actuaries estimated that the opioid epidemic cost the United States $179 billion in 2018. |
| Importance: Related U.S. Department of Health and Human Services (HHS) priority topics | Opioid crisis:                                                                                         |
|                                       | This measure addresses the HHS priority of addressing the opioid crisis, specifically the goal of advancing the practice of pain management by encouraging selective and appropriate opioid prescriptions to individuals experiencing chronic pain. |
| Validity                              | Construct validity: EHRs can differ in their ability to capture opioid prescriptions; countries with EHRs that do not capture all of the available opioid medications could have biased results. If survey data are being used, a minimum standard for a follow-up visit (e.g., duration, content) should be established. |
|                                       | Internal validity: This measure was determined to have an evidence base rating of 4, indicating that it is based on clinical experience and observations, observational studies with important limitations, or randomized clinical trials with several major limitations. There are currently no studies evaluating the effectiveness of monitoring patients at this interval (e.g., within 4 weeks). The recommendation was based on evidence that continuing opioid therapy for more than 3 months significantly increases the risk of developing OUD and that patients who |</p>
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>The Percentage of Patients with a Follow-Up Visit Within Four Weeks of Starting an Opioid for Chronic Pain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>are not experiencing pain relief within 1 month are unlikely to experience pain relief by 6 months. Therefore, following up with patients prior to 3 months and discontinuing opioids if a patient is not experiencing effective pain relief should reduce OUD risk.¹</td>
</tr>
<tr>
<td></td>
<td><strong>Discriminant validity:</strong> This measure does not account for the quality of the follow-up visit, questions asked, or effectiveness in preventing adverse outcomes from use of opioids for chronic pain. The CDC guidelines include recommendations for what questions clinicians should ask patients during the follow-up visit and considerations for determining whether to continue with opioid therapy. These recommendations should be compared with other countries' follow-up practices to determine what counts as a follow-up visit, as well as availability of EHR and pharmacy record data.</td>
</tr>
<tr>
<td>Feasibility for international comparisons</td>
<td><strong>Definitions:</strong> This measurement definition is derived from CDC guidelines (detailed above).¹ Currently, this measure is only proposed for use in the United States; other countries are not using this measure. There are no specific data available for international comparisons. However, there are other relevant data on rates of chronic pain and rates of opioid prescription and overdose, and it is possible that this measure could be obtained for other countries from existing data.</td>
</tr>
<tr>
<td></td>
<td><strong>Data availability:</strong> This measure could be obtained using prescription data from the practice EHR, as well as from patient follow-up visit data from the EHR or the medical chart. The Organisation for Economic Co-operation and Development (OECD) has identified ten countries with EHR systems that could adequately support this type of research and reporting, as well as other countries that are currently developing EHR systems to be used for national measurement of quality in the future.¹⁶ Survey data could also be used.</td>
</tr>
<tr>
<td></td>
<td><strong>Data comparability:</strong> The same OECD report found that there is variation in the EHRs that would complicate comparisons across countries, including lacking or inadequate terminology standards for the same terms, incomplete records or records not kept up to date, variation in quality of providers' recordkeeping, lack of provider-level quality checks, low rates and low quality of diagnostic and/or procedural coding, incomplete coverage, complex legacy systems, lack of standardization, inability to identify and match patients across providers, and difficulty with transitioning from paper to electronic records resulting in duplication.¹⁶ Additionally, there may be differences in the ways that opioid prescriptions are captured in the EHR, with some opioid prescriptions not identified in some systems.¹</td>
</tr>
<tr>
<td></td>
<td><strong>Subgroup analysis and equity:</strong> In the United States, there are significant disparities in the treatment of pain and in rates of opioid prescribing and rates of OUD by race/ethnicity,¹⁷,¹⁸ gender, and socioeconomic status.⁵ These disparities must also be evaluated in other countries and must be considered when defining this measure. The ability to conduct subgroup analyses to assess disparities will be limited by the same data quality issues that impact data comparability.</td>
</tr>
<tr>
<td>Usability: Current evidence of use</td>
<td>The CDC metric is relatively new; there are no available studies that report on this measure on a national or global level.</td>
</tr>
</tbody>
</table>
The Percentage of Patients with a Follow-Up Visit Within Four Weeks of Starting an Opioid for Chronic Pain

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>The Percentage of Patients with a Follow-Up Visit Within Four Weeks of Starting an Opioid for Chronic Pain</th>
</tr>
</thead>
</table>
| Usability: Relationship to currently used OECD measures | This measure is new, proposed by the CDC in 2016, and is not among the current OECD measures. OECD currently collects the following data related to opioid prescribing practices that are closely related to this measurement concept but do not address the same parameters:  
  - availability of analgesic opioids  
  - opioid-related deaths; overall volume of opioids prescribed  
  - proportion of population who are chronic opioid users.  
  Overall volume of opioids prescribed is the closest measure, as it helps to inform the denominator. |

References

Time to Regulatory Approval for New Prescription Drugs

Table A.22. Time to Regulatory Approval for New Prescription Drugs

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Time to Regulatory Approval for New Prescription Drugs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>When new prescription medications are discovered, it may take different amounts of time for the same biologic agent to be approved for sale in different nations. An international measure of time to regulatory approval could encourage nations with slower approval processes to consider streamlining them, particularly for medications that have demonstrable benefits, such as improved health outcomes or cost-effectiveness.</td>
</tr>
</tbody>
</table>
| **Importance: Health or social impact** | **Population impact:**

*National:* Many conditions are amenable to therapy with pharmaceuticals, including infectious diseases, cardiovascular diseases, and cancers. Compared with many other Organisation for Economic Co-operation and Development (OECD) nations, approvals of new prescription medications occur more rapidly in the United States (the time from new drug application to approval is typically 6–10 months for the U.S. Food and Drug Administration [FDA]). However, some studies suggest that the European Medicines Agency approves medications even faster, leading critics of the United States to argue that approvals should be accelerated. 

*Global:* Multiple studies have demonstrated that approvals of new medications are slower in nations other than the United States. For example, lovastatin became available in Egypt 12 years after it gained FDA approval. According to a National Bureau of Economic Research report, "long launch lags are common and nearly 40 percent of all new drugs are only launched in ten or fewer countries," suggesting that patients in some nations could be denied beneficial medications due to regulatory barriers.

**Clinical impact:** Approval of a new drug for a life-threatening disease can result in significantly decreased mortality rates and improved quality of life for individuals with that disease. In the United States, age-adjusted heart disease mortality fell by more than 50% from 1950 to 1996, and then by another 22% from 1990 to 2013. Much of this improvement could be due to innovations in pharmaceuticals that became accessible to patients across the nation. Improvements in outcomes for other conditions could be made more rapidly if new prescription drugs were approved rapidly. Alternatively, more-rapid adoption of new medications could prove costly or harmful if new products have not undergone thorough testing of efficacy, safety, and cost-effectiveness.

**Performance gap:** Time to regulatory approval for prescription drugs varies greatly from nation to nation. The United States and Europe tend to have faster approval times compared with other OECD nations. Lower-income countries are less likely to approve medications rapidly for various reasons, including that public or private payers might not have the means to pay for expensive new medications, making their markets less profitable.
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Time to Regulatory Approval for New Prescription Drugs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Costs:</strong></td>
<td>The financial cost to society could be large if new medications are approved more rapidly, since these medications tend to be costly. However, populations denied essential medications could experience poor outcomes leading to long-term disability, which are also costly. 12</td>
</tr>
<tr>
<td><strong>Importance:</strong></td>
<td><strong>Value-based care:</strong> Bringing new and transformative pharmaceutical products to the market faster could lead to improved health outcomes for patients. Alternatively, it could hinder the provision of cost-effective care if new pharmaceuticals are expensive.</td>
</tr>
</tbody>
</table>
| **Other related priorities:** | *Drug pricing:* Accelerating approval of new drugs may be costly due to costs related to accelerating the review process. Additionally, accelerating approval could lead to more demand for new and expensive medications, which could cause the average price of medications to increase.  

*Opioid crisis:* Decreasing time to regulatory approval of new prescription medications could aid in addressing the opioid crisis. For example, policy options to reduce abuse of opioids include (1) approving new analgesic medications that are not addictive and (2) increasing availability of new and innovative pharmacologic treatments for opioid addiction. Alternatively, increasing spending on new pharmaceuticals could divert resources away from addressing the opioid crisis. |
| **Validity:**     | **Construct validity:** Time to regulatory approval for new prescription medications is a relatively straightforward measurement concept, though differences in approval processes could affect interpretability of international comparisons (see the section on data comparability).  

**Internal validity:** Time to regulatory approval is likely a good measure of the ability of a health system to innovate quickly. The association between good performance on measures of innovation and improved health outcomes is less clear. For example, if most candidate medications are orphan drugs, then outcomes are likely to improve for eligible patients. However, if low-value drugs are rapidly gaining approval, outcomes on health could be minimal or negative. 13,14  

The time to regulatory approval represents only the final step in a much longer, risk-prone process of drug development that begins with initial discovery and basic research. 12 There are many regulatory policies governing the entire pathway from discovery to market that differ widely by country in cost, duration, and complexity. 12 The unique contribution of each step of this process, including but not limited to time to regulatory approval, quality of health care, and patient outcomes, needs to be determined.  

**Discriminant validity:** A nation with exceptionally slow approvals of medications would likely benefit from streamlining approval processes. However, a nation with exceptionally rapid approvals might experience poor patient outcomes and high costs if newly approved drugs are expensive and lead to avoidable harms that could have been averted with more rigorous preapproval testing. 10 |
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Time to Regulatory Approval for New Prescription Drugs</th>
</tr>
</thead>
</table>
| **Feasibility for international comparisons** | **Definitions:**  
*Time from concept to market:* The time between initiating product design and bringing it to the commercial market.  
*Time of regulatory review:* The time required for a regulatory agency such as the FDA to review and approve a product.  
*Benchmarks for regulatory review could be different if a drug is considered for accelerated approval.*  
*Launch lag:* The time between the product’s first launch date until its launch date in another country.  
*First to market:* Whether a medication was approved first in one country versus any other.  
**Data availability:**  
- *National:* The FDA website lists approval dates for new medications.  
- *International:* Multiple reports cite the MIDAS data set produced by IMS Health as a source of commercial market research data describing product launches in multiple therapeutic classes. The European Medicine Agency also posts information describing new drugs and approval dates.  
**Data comparability:**  
- Measures comparing use of medications internationally generally use the World Health Organization’s Anatomical Therapeutic Chemical classification system to identify active ingredients. This covers all classes but not all countries.  
- Identifying drug launches over time can be challenging because brand names are often referenced rather than active ingredients, which makes one-to-one comparison difficult.  
- Country-specific differences in approval processes can make comparability challenging. For example, some countries prioritize safety monitoring prior to approval, while others approve medications earlier and then perform post-marketing surveillance of safety.  
**Usability:**  
- *National:* The FDA monitors time of regulatory review for all products.  
- *International:* Multiple OECD nations track time to regulatory approval, including the European Union, Canada, and Japan. Using publicly available data, academic investigators have compared approval times in the United States with those of other nations. For example, one study published in 2012 collected data on 510 novel drug applications in the United States, Europe, and Canada between 2001 and 2010. The investigators tracked medical length of time for completion of first review, median total review time, and number of drugs first to market in each geographic area.  
**Usability: Relationship to currently used OECD measures** | Time to regulatory approval for new prescription drugs is not currently represented among current OECD measures; however, OECD has expressed interest in the topic.  
The closest available measures include the following:  
- Annual approvals of new medicines per billion U.S. dollars of pharmaceutical business expenditure on research and development (R&D) in the United States, inflation-adjusted: This measure analyzes the ratio of the number of new approvals of medications per unit of expenditure on drug R&D. In other words, this is a measure of return on
Multiple options exist for future development of OECD measures of time to regulatory approval for new prescription drugs.

- Median time to approval for new drugs each year would require data on all approved drugs, including data describing when applications were submitted and approved.
- Alternatively, OECD could identify lists of specific drugs that are likely to greatly improve population health outcomes or value and monitor dates they came to market in each member nation (or first-to-market statuses).

### References

Travel Time to Provider Office

Table A.23. Travel Time to Provider Office

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Travel Time to Provider Office</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>The burden of travel from a patient’s residence to his or her health care provider can be an important measure of access to preventive and treatment services. Travel time can be estimated directly (e.g., using geographic information systems [GIS], online routing websites, or self-report measures), but when data are unavailable, travel distance is often used as a proxy measure of travel time. However, although time and distance are highly correlated,(^1) they can differ depending on several factors (travel mode, time of day, travel speed, etc.).(^2) Travel time to a provider is sometimes measured under the umbrella term network adequacy. Network adequacy is a health plan’s ability to deliver the benefits promised by giving reasonable access to in-network providers; the maximum travel time or distance to providers should be 30 miles or 30 minutes from the home.(^3,4)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Importance: Health or social impact</th>
<th>Population impact:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Adverse outcomes associated with greater distance or travel time may include treatment delays and discontinuity of care.(^5)</td>
<td></td>
</tr>
<tr>
<td>• Low-income populations bear a relatively high burden of travel because they often reside in remote areas and/or have limited means of transportation.(^6) Having to travel greater distances for mental health care or clinical interventions negatively affects utilization in rural and urban areas.(^7,8)</td>
<td></td>
</tr>
<tr>
<td>• National prevalence: Studies have found transportation barriers in as little as 3% or as much as 67% of the population sampled, making it difficult to assess the impact on utilization or outcomes.(^8) Adults report traveling 38 minutes, on average, to their health care provider.(^9) Travel to a provider can also affect provider selection, such as electing to visit an emergency department (ED) instead of a primary care provider for nonurgent care.(^10)</td>
<td></td>
</tr>
<tr>
<td>• Global prevalence: Despite lower user fees in Organisation for Economic Co-operation and Development (OECD) countries, distance to providers may aggravate the disparities in utilization between patients with higher or lower socioeconomic position.(^11)</td>
<td></td>
</tr>
<tr>
<td>• Among 21 low- and middle-income countries (LMICs), the majority of respondents reported traveling 10 minutes or less (about 21%) or 11–20 minutes (about 22%) to get to a health facility; however, there were variations across countries, with Burkina Faso having the lowest proportion of respondents reporting a health facility within 10 minutes (about 4%) and Jordan having the highest (about 90%).(^12)</td>
<td></td>
</tr>
</tbody>
</table>

Clinical impact: |
• Living farther away from care has clinical and cost implications.\(^13\) |
• Rural counties report worse Medicare quality measures than urban counties.\(^14\) |
• A review of 108 studies from the global north countries (the United States, Canada, the countries of Western Europe, Australia, New Zealand) found that the majority of studies (77%) identified a “distance decay” association, where shorter travel time to health care services |
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Travel Time to Provider Office</th>
</tr>
</thead>
<tbody>
<tr>
<td>was related to better health outcomes among adults (18 years or older; excluding maternity and emergency medical travel); however, six studies showed the reverse (distance bias association), and 19 showed no association.¹⁵</td>
<td></td>
</tr>
<tr>
<td>• Among 21 LMICs, longer time to travel (i.e., more than 10 minutes) to a health care facility was associated with higher odds of neonatal mortality and lower odds of receipt of antenatal care during pregnancy and in-facility delivery.¹²</td>
<td></td>
</tr>
</tbody>
</table>

**Performance gap:**
- Transportation barriers like excessive travel time are an important barrier to health care access, especially for those with lower incomes, the underinsured, and the uninsured.⁸ Closing of rural hospitals limits local access to specialty care and can delay screenings, diagnoses, and treatment, particularly for older populations.¹⁶
- Most studies of neighborhood environments and their relationship to health and health equity have been cross-sectional and are limited in establishing causal relationships.¹⁷

**Costs:**
- Costs to seek care are usually measured as opportunity costs, or the value of patient time based on the value of foregone activities.¹⁸
- A study of U.S. adults found a mean of $42 in opportunity costs per visit (based on travel plus clinic time) to seek ambulatory medical care for oneself, which is higher than the mean out-of-pocket costs of $34 per visit.⁹
- A multistate study in the United States showed that longer travel times (e.g., more than an hour) for cancer care among older adults (aged 65 years or older) were associated with higher Medicare spending (e.g., $669/month more for those traveling more than 1 hour than those with commutes of 30 minutes or less in the initial phase of care) and higher patient cost responsibility (e.g., $67/month more for those traveling more than 1 hour versus those with commutes of 30 minutes or less in the initial phase of care).¹³ The higher Medicare spending could be due to higher hospitalization rates.
- Health maintenance organization members consume more health care and experience higher disease burden as their distance from primary care providers increases.¹⁹

**Importance:**
- Related U.S. Department of Health and Human Services priority topics
  - Value-based care: Quality measurement and value-based systems that do not take rurality and travel distance to a provider into account might unfairly penalize clinicians and lead to greater health disparities.¹⁴
  - Other related priorities:
    - Health insurance reform: Travel time is a potential indicator of access to personalized health care that meets individual needs and budgets, a key component of HHS’s priority for health insurance reform.
    - Opioid crisis: Longer drive times to the nearest opioid treatment program could affect adherence to the federal requirement that individuals must have six visits per week to receive medication.²⁰

**Validity**
- Construct validity: Analysis of measures of distance or travel time to a provider shows that these measures have strong construct validity.²¹,²² A measure of self-
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Travel Time to Provider Office</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>reported travel times to a hospital showed a moderately strong association with travel times estimated by GIS ($r = 0.86$).&lt;sup&gt;23&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

**Feasibility for international comparisons**

**Definitions:**
- Travel to a provider can be expressed in distance or time. Time is often more representative in urban areas where public transportation is utilized.<sup>24</sup>

**Data availability:**
- Researchers in the United States have combined disease registries with the Centers for Medicare & Medicaid Services (CMS) Physician Compare data, which list the name and location of clinicians who have billed Medicare in the past year.<sup>24</sup>
- The American Time Use Survey collects self-reported time to travel for care among a nationally representative sample of adults (18 years or older).<sup>25</sup> This measure is not specific to a particular site of care (e.g., ED or primary provider’s office).
- U.S. and international research can be conducted with physician and hospital claims data, physician registries, GIS mapping tools, and census data.<sup>11,26,27</sup>
- The cross-sectional Demographic & Health Survey, which collects data from nationally representative samples on a range of topics and across multiple countries, has a service availability module that identifies distance to the nearest health facility of interest, the most common transportation mode used by people to go to the facility, and the time it takes to travel to the facility using the most common transportation mode.<sup>28</sup> GIS and online routing services (e.g., Google Maps) are also increasingly being used to estimate travel time in various countries.<sup>15</sup>

**Data comparability:**
- Estimating travel time is complicated by multiple factors, such as transportation mode and availability, environmental conditions, transportation costs, and any other barriers to travel.<sup>29</sup> Other studies have used objective measures, such as GIS or online routing websites, but various geospatial methods exist to estimate travel time, and different assumptions are made (e.g., researcher-selected route to the provider), which may make comparisons across studies difficult.<sup>15</sup>

**Subgroup analysis and equity:**
- In the United States, American Time Use Survey data show that self-reported travel time to access care was higher for those who are non-Hispanic Black and those who are Hispanic (on average, about 45 minutes in each group) compared with those who are non-Latino White (36 minutes); higher for men (40 minutes) than for women (37 minutes); higher for those aged 25–44 years (38 minutes) and 45–64 years (41 minutes) compared with younger adults aged 18–24 years (31 minutes); higher for those who were unemployed or not in the labor force (41 minutes) compared with employed adults (about 35 minutes); and higher when the visit occurred on a weekday during work hours or after hours (38 minutes) compared with a weekend (33 minutes); no differences were found between urban (38 minutes) and rural (40 minutes) residence.<sup>30</sup> However, another national survey of adults shows that, taking distance to the nearest hospital and traffic patterns into account, rural residents are estimated to have higher travel times (17
Travel Time to Provider Office

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Travel Time to Provider Office</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>minutes) than suburban (12 minutes) and urban (10 minutes) residents.31</td>
</tr>
</tbody>
</table>

**Usability: Current evidence of use**

Several countries currently assess travel distance or time to health providers. Evidence includes the following (not exhaustive):

- **National:** Self-report measures (e.g., the American Time Use Survey) and geospatial techniques are being used to estimate travel time to providers, especially to identify disparities among population sociodemographic groups and geographic settings to identify where policies and interventions are needed to improve access to care.15,30,32
- States are no longer required to use time and distance standards to ensure provider network adequacy.33 Researchers can measure access to care by combining the CMS Medicare Provider Utilization and Payment Data with census data.34,35
- **International:** For countries with national health systems, health registries, mapping software, and census data can be combined to measure distance to health centers and the types of care provided. In the United Kingdom, distance to an ED similarly affected utilization.36
- The Demographic and Health Surveys collect multiple measures of travel burden (including time) to assess health service availability across countries.28

**Usability: Relationship to currently used OECD measures**

Travel time to a provider office is not among the current OECD measures.

**Closest related measures:**

- OECD measures the number and concentration of health providers and hospitals but not their geographic distribution or travel distances.37
- OECD collects data on unmet care needs (due to costs, travel distance, or waiting time) as a dimension of health systems performance. In 2013, 3% of the population among the OECD European Union countries reported unmet needs for medical care for those reasons, but there were variations across countries, with a higher prevalence in Latvia, Greece, Poland, and Estonia, and less than 1% prevalence in the Netherlands, Austria, Spain, Luxembourg, and the Czech Republic.38

**References**

Travel Time to Provider Office


Health Care Spending in the Last Year of Life

Table A.24. Health Care Spending in the Last Year of Life

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Health Care Spending in the Last Year of Life</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Health care spending in the last year of life is thought to be an indicator of wasteful spending on care for terminally ill patients with poor prognoses. After a patient’s death, the amount spent on that patient’s care can be calculated retrospectively. If the same calculation is performed for all patients who died in a one-year period, the total amount spent can be averaged over the number of persons who died.</td>
</tr>
</tbody>
</table>
| Importance: Health or social impact | • Population impact:  
  o National: The United States spent 17.1% of its gross domestic product (GDP) on health care in 2017.1 Estimates of expenditures for patients in the last year of life range from 10% of health spending for the general population to as much as 25% of Medicare spending for Medicare beneficiaries in the last year of life.2,3 One estimate of mean per capita medical spending in the last 12 months of life was $80,000 in the United States.3 Maximizing value in the care of high-cost, seriously ill individuals is a major public health challenge domestically, especially as the population ages.4  
  o Global: Annual health care spending is a major component of national budgets each year, accounting for between 7 and 12% of GDP for most wealthy Organisation for Economic Co-operation and Development (OECD) nations in 2017. Efforts to reduce unnecessary spending are widespread among member nations, including efforts to reduce spending at the end of life by redirecting patients toward palliative care and away from high-cost, low-value services.5,6  
• Meaningful clinical impact: Optimizing national performance on a measure of health care spending at the end of life would mean that patients received appropriate care when necessary and did not receive low-value (e.g., futile) services.4 For example, a well-performing nation would likely (1) maximize use of palliative care, which has been demonstrated to improve patient quality of life at the end of life;7 (2) minimize nonbeneficial treatments at the end of life;8,9 and (3) encourage patients to live their last days at home rather than in the acute care setting.10 It is important to note that some critics of end-of-life spending measures report that only 11% of the highest-cost individuals are in the last year of life, while even fewer have predictable prognoses, making spending reductions challenging.11  
• Performance gap: An academic study from 2017 found similar performance among nine countries on health care spending at the end of life. Each country with complete data spent between 8 and 11% of total health expenditures on care for patients in the last 12 months of life.3 There is some evidence that high-cost, low-value care could be more prevalent in the United States in the last three months of life.12 However, there is also some evidence that end-of-life care costs are decreasing in the United States.13  
• Costs: According to one estimate, the United States spent $200 billion on health care services for patients in their last year of life.11 Any reduction in such spending could be redirected toward potentially higher-value services. |
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Health Care Spending in the Last Year of Life</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Importance:</strong></td>
<td><strong>Value-based care:</strong> Reducing unnecessary spending at the end of life would improve the value of care delivered.</td>
</tr>
<tr>
<td>Related U.S.</td>
<td><strong>Construct validity:</strong> The challenge in constructing a valid international measure of health care spending at the end of life would be in calculating decedent-level expenditures. Although OECD has performed international comparisons of health care expenditures for many years, calculating individual expenditures would involve various data sources and methods across countries. This could be challenging for the United States, in particular, since estimates for patients not covered by Medicare would be hard to obtain.¹</td>
</tr>
<tr>
<td>Department of Health</td>
<td><strong>Internal validity:</strong> One threat to internal validity is that cultural differences might account for significant international variation in health care spending in the last year of life.</td>
</tr>
<tr>
<td>and Human Services</td>
<td><strong>Discriminant validity:</strong> Nations with very high expenditures would likely have room to improve performance. However, concerns have been raised that reducing high levels of spending could be challenging because predicting which severely ill patients will recover and which will not is often infeasible.¹¹,¹⁴–¹⁶</td>
</tr>
<tr>
<td>priority topics</td>
<td><strong>Feasibility for international comparisons</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Definitions:</th>
<th>There is no accepted standard measure of health care spending at the end of life, and academic studies use various time intervals to calculate end-of-life spending (range of 1 month to 5 years).¹⁷–²⁰ Expenditures have been expressed as the percentage of total national health expenditures or the amount spent per decedent.³,²¹</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data availability:</strong></td>
<td></td>
</tr>
<tr>
<td>• <strong>National:</strong> Studies characterizing end-of-life health care spending most often focus on patients covered by Medicare and use fee-for-service claims data to estimate cost.¹⁸–²⁰,²² One study used the Medical Expenditure Panel Survey and the National Health Expenditure Survey.¹¹ Another potential data source is the Medicare Current Beneficiary Survey, though this survey may be limited by recall bias and could be subject to systematic biases in times proximate to beneficiary death.²³</td>
<td></td>
</tr>
<tr>
<td>• <strong>Global:</strong> Studies of end-of-life spending in Canada and Europe have used insurance claims, estimates of resource use, and administrative databases to estimate spending per decedent.¹⁸,²⁴,²⁵</td>
<td></td>
</tr>
<tr>
<td><strong>Data comparability:</strong> Comparability of hospital spending data across nations might be additionally limited because of different methods of calculating estimates of spending using different data sources (such as claims, resources utilization estimates, and administrative databases). Estimates of nonhospital care are limited because data sources are disparate and more difficult to acquire than those describing inpatient hospitalizations.¹⁸ In nations that have multiple payers, there also may be limited ability to capture expenses based on payer-specific availability of data.¹</td>
<td></td>
</tr>
<tr>
<td>Measure Construct</td>
<td>Health Care Spending in the Last Year of Life</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td><strong>Subgroup analysis and equity:</strong></td>
<td></td>
</tr>
<tr>
<td>• There is some evidence that spending at the end of life varies based on race, with higher spending rates for Black decedents than for White decedents.¹⁷,²⁶</td>
<td></td>
</tr>
<tr>
<td><strong>Usability:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Current evidence of use</strong></td>
<td></td>
</tr>
<tr>
<td>• National: Outside of the academic studies cited above, the United States does not appear to be routinely using this measure or any similar measure.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• International: Outside of the academic studies cited above, other countries do not appear to be routinely using this measure or any similar measure.</td>
</tr>
<tr>
<td><strong>Usability:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Relationship to currently used OECD measures</strong></td>
<td></td>
</tr>
<tr>
<td>Health care spending in the last year of life is not currently represented among OECD measures. Relevant measures include the following:</td>
<td></td>
</tr>
<tr>
<td>• <em>Health expenditure per capita:</em> This measure calculates the amount spent on health care per person in a nation. It is simpler to calculate than expenditures per decedent because it does not require individual-level calculations.¹</td>
<td></td>
</tr>
<tr>
<td>• <em>Current health spending by age under the System of Health Accounts Framework:</em> This measure is only available for three OECD member nations, but it shows spending for specific age groups of patients. Similar methods could be used to stratify expenses for decedents versus nondecedents.²⁷</td>
<td></td>
</tr>
<tr>
<td>• <em>Long-term care expenditure by government and compulsory schemes as a share of GDP:</em> Persons receiving long-term care are more likely to expire than persons who are not receiving this care, which makes spending on their care conceptually similar to spending on decedent care.²⁸</td>
<td></td>
</tr>
</tbody>
</table>

References


Spending on Mental Health (Percentage of Total Health Spending)

<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Spending on Mental Health (Percentage of Total Health Spending)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>The World Health Organization (WHO) defines <em>mental health</em> as “a state of well-being in which every individual realizes his or her own potential, can cope with the normal stresses of life, can work productively and is able to make a contribution to her or his community.” Mental health problems refer to a set of medical conditions that affect a person’s thinking, feeling, mood, ability to relate to others, and daily functioning. National health spending describes both individual needs and population health as a whole. Health spending includes the consumption of health goods and services and includes all types of financing arrangements. Percentage of total health spending is the ratio of health spending allocated to mental health services compared with total national health spending.</td>
</tr>
</tbody>
</table>
| Importance: Health or social impact | Population impact: National:  
  - Mental health is an indispensable component of health, and mental disorders represent one of the most significant and increasing burdens to global public health. In the United States, an estimated 50% of the population will be diagnosed with a mental disorder during their lifetime. Mental health disorders have the highest prevalence and disability burden among all noncommunicable diseases. Furthermore, mental health spending tends to be low and inefficiently allocated, with the least-effective interventions receiving the largest funding share.  
  - Nearly one in five U.S. adults lives with a mental illness (46.6 million in 2017). More than 10 million adults in the United States have unmet need for mental health treatment. The United States had the highest rate of death from mental health disorders and substance use disorders (SUDs) among comparable countries (France, Germany, Sweden, Canada, Austria, Switzerland, the United Kingdom, Belgium, Netherlands, Japan) in the 2015 Global Burden of Disease Study. In 2016, mental health spending accounted for $180.7 billion out of an estimated $3.1 trillion of total health care spending, or 6%. Spending on mental health is expected to total $280.5 billion in 2020. Other studies estimate that mental health spending accounts for 7% of total health care spending and has increased by 62% since 2006. However, as a percentage of total health spending, spending on mental health is expected to account for 6.5% in 2020, down from 7.4% in 2009. |
|                   | Global:  
  - Mental illness represents nearly 10% of the global burden of disease (measured in disability-adjusted life-years lost), up from 6.6% in 1991. Mental health disorders and SUDs affected more than 1 billion people globally in 2016. Globally, the allocations for mental health care are disproportionately low relative to the burden. WHO estimated that in 2015, mental health spending accounted for less than 2% of the global median health expenditure. The global median mental health |
expenditure per capita was $2.50 out of a total global median per capita health expenditure of $141. The increased international focus on mental health is expected to increase the resources devoted to mental health. By 2030, the economic burden of poor mental health is projected to cost the global economy $16 trillion.

Clinical impact:
• Consequences: Disproportionately low spending on mental health translates to limited access to care. Currently, as many as 80% of those with a common mental disorder and up to 50% of those with a severe mental disorder do not seek or receive treatment.

Performance gap:
• In 2008, the Mental Health Parity and Addiction Equity Act required that health insurance plans include the same level of benefits for medical care and for mental health and substance use care. Although spending on and utilization of mental health services in employer-sponsored group health plans increased after the passage of this law, other data suggest that there was no significant effect overall when all types of plans were considered, and there are still significant disparities between physical and mental health care and unmet need for mental health care.
• The 2011 Grand Challenges in Global Mental Health initiative, led by the U.S. National Institute of Mental Health, prioritized implementation research questions to reduce the treatment gap for mental health disorders.
• The 2017 Global Alliance for Chronic Diseases consortium and the UK Research Councils developed an agenda to improve implementation and reduce the treatment gaps for mental health.

Costs:
• National: The estimated financial costs of mental disorders are at least $467 billion annually and rising.
• Global: The economic burden of poor mental health is projected to cost the global economy $16 trillion by 2030. The Organisation for Economic Co-operation and Development (OECD) reported that the “direct and indirect costs of mental ill-health can exceed 4% of GDP [gross domestic product]” and were estimated at $2,493 billion globally. Indirect costs include lost productivity, absenteeism, and premature mortality.

Importance:
Related U.S. Department of Health and Human Services (HHS) priority topics

Value-based care:
Improvements in mental health care could significantly increase value because mental health disorders account for a significant proportion of total disease burden. Furthermore, there is a strong association between mental and physical health, such that untreated mental illness is associated with increased physical health costs. For example, mental disorders such as depression and anxiety are negatively related to engagement in preventive, health-promoting behaviors.

Other related priorities:
Opioid crisis: One point of HHS’s 5-point strategy to combat the opioid crisis is “better addiction prevention, treatment, and recovery services.” Improved access and utilization of mental health services will aid in prevention, treatment, and recovery.
<table>
<thead>
<tr>
<th>Measure Construct</th>
<th>Spending on Mental Health (Percentage of Total Health Spending)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Health insurance reform:</strong> Spending on mental health is related to the HHS priority for health insurance reform because there are significant disparities in spending on and coverage for mental health care relative to the burden of disease associated with mental disorders.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Validity</th>
<th>Construct validity:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Mental health spending is disproportionately low and inefficiently allocated.⁴</td>
<td></td>
</tr>
<tr>
<td>• Evidence indicates that health care spending is positively related to actual improvements in health.²³</td>
<td></td>
</tr>
<tr>
<td>• There is general agreement that the level of spending should reflect the burden of disease attributable to mental disorders.²¹</td>
<td></td>
</tr>
</tbody>
</table>

**Internal validity:**
• Mental health disorders are common and costly, yet the quality of care for these disorders has not improved over recent decades.²⁴
• A key reason for the lag in mental health performance is the lack of monetary resources.²⁵
• However, the ideal proportion of spending on mental health as a percentage of total health spending that will result in the best health outcomes is likely to vary by country, and a higher percentage of spending on mental health might not necessarily be an indicator of mental health care quality due to differences in the populations.

**Discriminant validity:**
• The utilization of mental health services is, in part, a function of the characteristics of the population. Consequently, international comparisons of measures of mental health spending may be confounded by population characteristics, by access to care overall, and by subpopulations.
Additional threats to validity include international variation in the definitions of mental illness, as well as mental health diagnostic and treatment guidelines.

<table>
<thead>
<tr>
<th>Feasibility for international comparisons</th>
<th>Definitions:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• The WHO definition of mental health is widely used;¹ however, there are other definitions, and there is some disagreement among experts on the best definition.²⁰</td>
<td></td>
</tr>
<tr>
<td>• Related to mental health spending, OECD notes: “Governments cannot fully quantify the cost of mental illness, as few countries systematically measure the resources they devote to mental health. The lack of data on costs, quality and outcomes inhibits a complete assessment of mental health system performance. The result is poor policy, and an inability to direct scarce resources to areas of need.”²¹</td>
<td></td>
</tr>
</tbody>
</table>

**Data availability:**
• Although some data on mental health spending and overall health spending are currently available and in use, OECD has suggested that available data are not sufficient for a full picture of mental health system performance.²¹
  ○ Global: The World Health Organization Assessment Instrument for Mental Health Systems (WHOAIMS) is a tool for collecting essential information on the mental health system of a country or region. National governments submit this information, which includes mental health spending and overall health spending, to WHO and other coalitions.²⁷
Measure Construct | Spending on Mental Health (Percentage of Total Health Spending)
---|---
**Data comparability:**
- Health system characteristics could play a role in diagnosis, treatment, and spending associated with mental illness. The level of cost sharing in the health system associated with health care generally, and mental health care specifically, influences the patient’s utilization of care.\(^{28}\)
- International differences in diagnoses and treatment guidelines, as well as cross-cultural differences in the presentation of mental health conditions and related help-seeking behaviors, would likely also threaten the comparability of the data.\(^{29}\)

**Subgroup analysis and equity:**
- Currently, available data on mental health spending are at the national level and do not describe spending among subgroups.
- Disparities in rates of mental health disorders and receipt of mental health treatment by gender\(^{30}\) and race/ethnicity\(^{31}\) will affect variation in spending within a given country.

**Usability:**

<table>
<thead>
<tr>
<th>Current evidence of use</th>
<th>National:</th>
</tr>
</thead>
<tbody>
<tr>
<td>- The United States, specifically the Substance Abuse and Mental Health Services Administration, measures the amount of spending on mental health services and the percentage of total health spending devoted to mental health services.(^{11})</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Global:</th>
</tr>
</thead>
<tbody>
<tr>
<td>- WHOAIMS collects information from national governments regarding mental health spending and overall health spending.(^{27})</td>
</tr>
</tbody>
</table>

**Usability:**

| Relationship to currently used OECD measures | Mental health spending (or mental health spending as a percentage of total health spending) is not among the current OECD measures. However, OECD has begun working to benchmark the performance of nations’ mental health care systems.\(^{32}\) |

**Closest related measures:**

- **Health expenditure by type of service:** This set of measures describes the percentage of national health spending devoted to primary care, inpatient care, outpatient care, long-term care, medical goods, collective services, and others.\(^3\) The measures demonstrate that the structure of spending across various types of care varies considerably by country.

### References