

1 APPENDIX B

2 Commentary

3

4 INTRODUCTION

5 The hospital care additional priorities project 2007 was formally launched in August 2006, based
6 on an agreement between National Quality Forum (NQF) and the Agency for Healthcare
7 Research and Quality (AHRQ) to pursue endorsement of a set or group of voluntary consensus
8 standards that can be used for public reporting, that specifically address gaps in availability of
9 measures related to inpatient quality (including patient safety and pediatrics).

10 A “Call for Measures” in the areas of morbidity and mortality, anesthesia and surgery
11 (including surgical volume and mortality), utilization rates for high-risk (or often unnecessary)
12 procedures, readmission rates and length of stay, and pediatric pain assessment was issued in
13 August 2006 as part of the NQF’s standard operating procedure. Following the Call for
14 Measures, Blue Cross Blue Shield Association and America’s Health Insurance Plans provided
15 funding to have methodologies for length of stay and readmission rates considered under the
16 auspices of the project. A “Call for Measures” for this portion of the project was subsequently
17 issued.

18 As with other NQF consensus projects, a Steering Committee representing key healthcare
19 constituencies and Technical Advisory Panels (TAP) with expertise in the areas to be addressed
20 were convened. (appendix E)

21 This appendix summarizes the deliberations of the Steering Committee related to the
22 proposed voluntary consensus standards.

23

24 APPROACH

25 The Steering Committee began its work by affirming its use of the framework delineated in
26 *A Comprehensive Framework for Hospital Care Performance Evaluation*¹ and the principles therein
27 that address promoting standardization, driving measure set improvement, and supporting
28 implementation.

¹ National Quality Forum (NQF). *A Comprehensive Framework for Hospital Care Performance Evaluation: A Consensus Report*. Washington, DC: NQF; 2003.

29 Additionally, the Steering Committee and TAPs espoused the principles of support for
30 evidence-based practice; and evidence that candidate measures are reliable and valid
31 throughout the process of evaluating the measures.

32

33 Identification and Evaluation of Candidate Measures

34 Potential candidate consensus standards were identified through:

- 35 • “Call for Measures” in the areas of interest;
- 36 • recommendations from the TAPs and the Steering Committee; and
- 37 • review of NQF-endorsed™ consensus standards applicable to the areas of interest.

38 NQF staff prepared detailed measure evaluations using the NQF-endorsed™ standard
39 criteria of important, scientifically acceptable, useable and feasible established in *A*
40 *Comprehensive Framework for Hospital Care Performance Evaluation*. Information for the measure
41 evaluations was obtained from the measure developer and literature review. The Steering
42 Committee had provided detailed guidance to help staff and the TAPs to focus on aspects of the
43 measures of particular interest. The TAPs then reviewed the measure evaluations prepared by
44 NQF staff and, after hearing presentations from representatives of the measure developers,
45 clarified points and concerns with those representatives. Following deliberations of the
46 perceived strengths and weaknesses of each of the measures and technical reasons why the
47 measures should or should not be recommended, the TAP made recommendations to the
48 Steering Committee.

49 The Steering Committee developed guidance for the work to ensure a common direction
50 and approach to be undertaken by the multiple TAPs. That guidance included a statement of
51 purpose, priorities and other things to be considered in evaluating the individual measures and
52 considering them within the overall context of hospital measures. That guidance is summarized
53 below.

54

55 Purpose

56 The purpose of the hospital care additional priorities project is to improve the quality of
57 healthcare by recommending:

- 58 • hospital care measures, including composite measures, for NQF endorsement that fill
59 gaps or voids in the current measure set, especially in the areas of pediatrics, surgery
60 and anesthesia, morbidity and mortality, and patient safety;

- 61 • methodologies for length of stay and readmission rates; and
62 • guidance for public reporting of measures.

63

64 **Considerations**

65 In considering which of the candidate consensus standards should move forward for
66 endorsement, the Steering Committee charged the TAPs to consider a number of factors in
67 recommending measures for endorsement, including that they:

- 68 • are fully open source;
69 • are fully developed and specified;
70 • apply to acute care hospitals, as relevant based on scope of service;
71 • are useful to and useable by the public, including stakeholder groups; and
72 • reflect those aspects of care over which hospitals have control or can substantially
73 influence.

74 And as a group, that the recommended measures encompass those that:

- 75 • address to the extent possible reducing disparities in the quality and safety of care for
76 minority populations regardless of the primary focus; and
77 • include elements, including education and awareness to improve the public's ability to
78 understand and use performance data.

79

80 **Priority Areas Within the Candidate Consensus Standards**

81 An overarching priority set by the Steering Committee was that measures proposed for
82 endorsement should continue the NQF effort to advance a set of measures representative of the
83 six IOM aims. Additional priorities approved by the Steering Committee for this project focus
84 on candidate consensus standards:

- 85 • that fill gaps or voids in the NQF-endorsed hospital care consensus standards;
86 • that can be applied to multiple units or services within the acute hospital setting (i.e.,
87 crosscutting);
88 • that are in common, widespread use and/or are required for other purposes (e.g.,
89 meeting accreditation requirements, addressing national goals or initiatives);
90 • suitable for accountability;
91 • that address misuse or overuse;
92 • that are directly applicable to specific at-risk populations (e.g., neonates, frail elderly);

A-3

- 93 • with high-level evidence;
- 94 • that address one or more of the six NQF-endorsed healthcare system quality “aims”² ;
- 95 and
- 96 • that minimize burden through use of electronically available data.

97

98 **Criteria for Selection of Measures**

99 To evaluate the measures, the TAPs and Steering Committee used the NQF-endorsed criteria
100 from *A Comprehensive Framework for Hospital Care Performance Evaluation* – that is that the
101 measures should be important, scientifically acceptable, useable, and feasible. Additionally, an
102 evidence-grading tool being piloted in NQF projects was used, since feedback from the project
103 Steering Committees that had previously used the tool indicates that the standardized grading
104 system provides more uniform recommendations from the various TAPs. The
105 recommendations of the TAPs were conveyed to the Steering Committee by their Chairs and
106 representatives of the developers were present to respond to questions regarding the measures.

107

108 **THE PROPOSED VOLUNTARY CONSENSUS STANDARDS FOR HOSPITAL CARE: ADDITIIONAL** 109 **PRIORITY AREAS – 2007**

110 The strengths and weaknesses of each measure, as assessed by the relevant TAP(s) and outlined
111 in each measure evaluation, were fully considered by the Steering Committee in its
112 deliberations. At least one representative of each TAP, usually the Chair, was in attendance to
113 present recommendations related to the measures the TAP had evaluated. Additionally,
114 representatives of the measure developers were present to respond to questions from the
115 Steering Committee.

116 Of particular note, when the Steering Committee took a position that differed from a TAP
117 with respect to advancing a measure, it did so with TAP representative input and only after
118 reviewing whether the measure in question met measure selection criteria and whether the
119 concerns could be mitigated through measure adjustments by the developer or were not
120 directly related to the construct or application of the measures; e.g., how reported. In the event
121 recommendations were made specific to individual measures, those recommendations have
122 been conveyed to the developers for appropriate action. In most cases, the recommendations

² Identified in the NQF document *A Comprehensive Framework for Hospital Care Performance Evaluation: A Consensus Report*. NQF, 2003.

123 were acted upon prior to final action by the Steering Committee and are reflected in the
124 discussion which follows.

125 Of the 43 measures recommended for endorsement in the National Voluntary Consensus
126 Standards for Hospital Care: Additional Priorities – 2007, Part 2 and this addendum, eight use
127 3M APR-DRG in their risk adjustment. Those measures are:

- 128 • Acute Stroke Mortality Rate
- 129 • Bilateral Cardiac Catheterization Rate
- 130 • Congestive Heart Failure Mortality
- 131 • Hip Fracture Mortality Rate
- 132 • Abdominal Aortic Aneurysm Repair Mortality Rate
- 133 • Esophageal Resection Mortality Rate
- 134 • Incidental Appendectomy in the Elderly Rate
- 135 • Pancreatic Resection Mortality Rate

136 The measure steward of these eight measures, AHRQ, holds a limited license 3M APR-DRG
137 grouper which is included with the AHRQ QI Software to which users have free access.

138 Additionally, 3M has provided a written commitment to the measure steward that it will
139 provide access to the APR DRG methodology through a website during the NQF consensus
140 development process and for the measures that are subsequently endorsed, “3M will provide
141 the necessary information regarding the APR DRG methodology so that it can be included in
142 the measure specifications for the final public consensus report.”

143 The summary which follows does not include all topics discussed; it is intended to capture
144 key points and areas of concern, recommendations for measure improvement and actions taken
145 by developers to address Steering Committee recommendations.

146

147 Length of Stay/Readmission

148

149 Risk-adjusted average length of inpatient hospital stay. In its initial deliberations, the Steering
150 Committee was divided in its support of this measure. Its concerns mirrored those of the TAP;
151 i.e., insufficient, information provided by the developer to allow a thorough assessment of the
152 CACR and concern about the inclusion of socioeconomic variables in the model. With respect
153 to the latter, the developer advised the Steering Committee that these factors are included in an
154 effort to avoid penalizing a hospital for providing care to populations whose socioeconomic

- Deleted: discussion
- Deleted: of this measure,
- Formatted: Font: Book Antiqua, 11 pt
- Deleted: e
- Deleted: lack of
- Deleted: on
- Deleted: risk adjustment model, including the proprietary comorbidity-adjusted complication risk (
- Deleted:)
- Deleted: ,
- Deleted: With respect to the former, the CACR is described as a probability model when the measure is approved, the details of the methodology and all coefficients will be made publicly available.
- Deleted: provider

A-5

155 status or ethnicity have been shown to be predictors of poorer outcomes. The Steering
156 Committee opined that the counterpoint to this position is that it could improperly credit
157 institutions that provide substandard care to these populations. Thus prior to making a final
158 recommendation, the Steering Committee asked that the developer supply additional details to
159 allow the TAP to conduct a more detailed review of the CACR, including the implied moral
160 hazard, and determine, based solely on its technical merits, whether its use would be
161 appropriate in a measure advanced for endorsement. Having secured additional documents
162 and examples detailing the model, the TAP reconsidered the soundness and generalizability of
163 the CACR and agreed that it is a valid, robust probability model that could feasibly be
164 implemented on a national scale. While its concerns related to the inclusion of socioeconomic
165 variables remained unchanged, the TAP acknowledged that the long-standing socioeconomic
166 variable debate is an issue of moral hazard, and is thus not technical, but rather philosophical in
167 nature. With the technical merits of the measure clarified and confirmed by the TAP, the
168 Steering Committee recommended the advancement of the measure while acknowledging the
169 ongoing concern regarding use of socioeconomic variables in this and any measure.

Deleted: P

Deleted: provide

Deleted: ly used

Deleted: risk model

Deleted: . It

Deleted: the model

Deleted: , and could

171 Overall inpatient hospital average length of stay (ALOS) and ALOS by DRG service
172 category. During its deliberations, the Steering Committee viewed this as an important and
173 necessary area for measurement and approached this standard together with the all-cause
174 readmission index informed by the TAP evaluation. It was in agreement with that body
175 regarding their assessment of the weaknesses of these measures, including a risk adjustment
176 model based on resource-based DRGs, the need for additional testing, and a failure of the LOS
177 measure to effectively account for outliers - which also compromised its alignment with the
178 readmission measure. After follow-up consultation with the TAP, the Steering Committee
179 addressed the treatment of outliers by recommending that the arithmetic mean be replaced by
180 the geometric mean, thus eliminating the need for the exclusion. Additionally, acknowledging
181 that DRGs are inherently resource-related and their use in risk adjustment is an extrapolation,
182 and in accordance with NQF's intellectual property policy, the Steering Committee made clear
183 that its recommendation for endorsement is contingent on use of CMS-DRGs rather than APR-
184 DRGs, since the former are in the public domain. In doing so, it recognized that there are
185 populations that will not be captured (i.e., pediatric and obstetric patients), and that as MS-
186 DRGs will replace CMS-DRGs in the next few years, measure maintenance will be necessary to

Deleted: Inpatient

Formatted: Font: Book Antiqua, 11 pt

Formatted: Font: Book Antiqua, 11 pt

Formatted: Font: Book Antiqua, 11 pt

Deleted: T

Deleted: and

Deleted: strengths and

Deleted: T

Deleted: was addressed after consultation with the TAP

Deleted: outlier

Deleted: A

Deleted:

Deleted: necessitating

187 ensure currency of the measure. Further, the Steering Committee notes that this measure is not
188 suitable for mental health, substance abuse and transplant patients groups the developer has
189 agreed to exclude. Finally, it recommends that whenever reported this measure be paired with
190 the all-cause readmission index.

Deleted: and that there are populations that cannot be captured; i.e., pediatrics and obstetrics.
Deleted: it
Deleted: ,
Deleted: Additionally
Formatted: Indent: First line: 18 pt

191 During the report review phase, the measure developer requested reconsideration of use of
192 the geometric mean based on 1) concern that the geometric mean inappropriately raises the
193 average, allowing poor performers to appear to perform 'better' because they are being
194 compared to an inflated mean; 2) the position that the arithmetic mean with exclusion of outliers
195 represents a normative approach to addressing outliers and 3) the developer's belief that the
196 arithmetic mean with outliers is the industry standard. After an opportunity to review documents
197 from its prior deliberations, the Steering Committee recommends advancing the measure to vote
198 with the arithmetic mean with outlier exclusions. It further recommends that the developer
199 continue testing to ensure that this approach consistently results in a less biased mean.

201 All-cause readmission index. The importance and the strengths and weaknesses of this
202 candidate standard parallel those of the 'inpatient hospital average length of stay measure'
203 discussed above. A concern specific to this measure was the fact that same day readmissions
204 are excluded from the denominator population. In its deliberations, the TAP determined that
205 an occurrence of true, unplanned readmissions (as opposed to planned transfers to another
206 acute care facility) above 20 percent of all same-day readmissions would cause it to rule against
207 advancing the measure and requested that the developer provide this information. With the
208 developer's documentation that, when excluding intended transfers, the rate at its highest is 11
209 percent, the Steering Committee supported the measure, again with the stipulation that it use
210 the CMS-DRGs for risk adjustment. As previously noted, the Steering Committee
211 recommended that this and the 'inpatient hospital average length of stay' measure be reported
212 together.

Deleted:
Deleted: of the measure
Deleted: measure
Deleted: an
Deleted: sion

213
214 **30-day all-cause risk standardized readmission rate following heart failure hospitalization.**

215 This disease-specific measure is limited to the Medicare fee-for-service population and, by
216 including all causes for readmission, expands the potential for improving care beyond the
217 specified diagnosis. It employs the same hierarchical risk adjustment methodology used in the
218 recently NQF-endorsed heart failure, acute myocardial infarction, and pneumonia 30-day

Deleted: uses

219 mortality measures. The developer reports that its data demonstrates that for this patient
 220 population about 25 percent of hospital readmissions within 30 days are due to a recurrence of
 221 heart failure, suggesting that 30 days is a reasonable time frame for the measure. Further,
 222 validation of administrative against medical record data showed that administrative data captured
 223 most of the risk and that there was high correlation between the two. Two concerns were
 224 discussed by the Steering Committee in some detail: 1) the way in which the measure is reported
 225 could reduce its usefulness by identifying outliers only at the extremes of high and low; and 2)
 226 hierarchical models reflect low volume hospitals at the population mean rather than reflect the true
 227 performance of the hospital. In recommending the measure, the Steering Committee pointed out
 228 that there is some lower limit to the measure's utility in terms of its interpretation, in terms of
 229 improvement, or in terms of patient or payer choice. Accordingly, it strongly recommended that
 230 when publicly reported, a volume threshold should be identified below which results are only
 231 marginally affected by a hospital's own data and that results below that threshold should be
 232 suppressed (i.e., not reported).

233

234 **Severity-standardized average length of stay – routine care, Severity-standardized average**
 235 **length of stay – special care, Severity-standardized average length of stay - deliveries.** These
 236 three measures are recommended for endorsement as individual consensus standards but were
 237 considered together because of their fundamental methodologic similarities. A readmission risk
 238 adjuster, formerly a fourth measure, is to be incorporated as a part of each of the three to
 239 discourage inappropriately early release of patients to improve LOS scores.

240 The Steering Committee again agreed with TAP's assessment of the measures' strengths and
 241 weaknesses – specifically that the measures address an important topic, employ a feasible data
 242 source, and that the risk variables are well chosen and applied, but that the measures could
 243 benefit from additional testing. While the measures target an insured, commercial population,
 244 the Steering Committee was of the opinion that they are generalizable if the appropriate data
 245 sets can be accessed. The Steering Committee ultimately recommended the measures with two
 246 adjustments, to which the developer has agreed: 1) removing pregnancy as a complication in
 247 the 'deliveries' measure; and 2) removing ing from the list exclusions 'cases where accommodation
 248 revenue codes are missing'.

249

250 Patient Safety, Adult and Pediatrics

- Deleted: ed
- Deleted: data
- Deleted: only
- Deleted: that
- Deleted: ive of
- Deleted: choice
- Deleted: ,
- Deleted: ;
- Formatted: Line spacing: 1.5 lines
- Deleted: D
- Deleted: measures
- Deleted:
- Formatted: Font: Not Italic
- Formatted: Body Text, No widow/orphan control, Don't hyphenate
- Formatted: Font: Book Antiqua, 11 pt, Not Italic
- Formatted: Font: Book Antiqua, 11 pt, Not Italic
- Formatted: Font: Book Antiqua, 11 pt, Not Italic
- Formatted: Font: Book Antiqua, 11 pt, Not Italic
- Formatted: Font: Book Antiqua, 11 pt, Not Italic
- Deleted: to the measures
- Formatted: Font: Book Antiqua, 11 pt, Not Italic
- Deleted: e

251
252 **Accidental Puncture or Laceration (Children and Adults)**. The two measures were considered
253 together based on fundamental methodologic similarities though the risk adjustments and
254 comorbidities differ based on the populations. The strengths considered were that this is an
255 important issue, these coded events have high accuracy, good predictive value particularly with
256 surgical cases, and user feedback indicates strong feasibility. It was noted that reporting
257 caution often results in coding only when additional care is required. The Steering Committee
258 stressed that, in reporting the pediatric measure, it is important that comparisons should be
259 among pediatric populations only whether care is provided in general acute care hospitals or in
260 children's hospitals. The Steering Committee also recommended that the developer add a
261 volume standard to the pediatric measure specifications to which the developer agreed.

262
263 **Death in low mortality DRGs**. The Steering Committee recommended this measure advance
264 as originally submitted, as a rate. The impetus for using a rate was to be able to risk adjust,
265 therefore to consider the rate within the context of the hospital size and population/patient-
266 specific factors. The TAP had recommended it be changed to a count because the low rates
267 were difficult to interpret. There was considerable discussion about the relationship between
268 this measure and the NQF-endorsed serious reportable events; however, in the end all
269 acknowledged that the serious reportable events do not meet the criteria of measures. A
270 concern registered within the Steering Committee, resulting in the one dissenting vote, was the
271 question of positive predictive value in that the most recent studies addressing this are from the
272 1980s.

273
274 **Iatrogenic pneumothorax, Iatrogenic pneumothorax in non-neonates**: As with all the AHRQ
275 quality indicators advanced in this project, these two measures are in public use at present. The
276 Steering Committee confirmed that the case mix adjustment in these measures is on the patient
277 level and, without further concern recommended both measures advance.

278
279 **Decubitus ulcer. (Pediatrics)** This measure directly addresses the quality of care received since
280 decubiti are largely preventable. It is currently being publicly reported, relates to the NQF-
281 endorsed™ nurse sensitive measure of pressure ulcer prevalence, and complements NQF-
282 endorsed Safe Practice 27, which requires evaluation of each patient upon admission and

283 regularly thereafter for risk of developing pressure ulcers. Nonetheless, the consideration of
284 the measure posed challenges for the Steering Committee. The measure excludes certain
285 patients with an ulcer present on admission (POA) or transferred from a long term care facility
286 to avoid penalizing providers that care for these patients. However POA coding is not in
287 widespread use at present though in the near future, it will become part of the specifications of
288 the measure. In terms of weakness, the measure captures an infrequent occurrence and its
289 results might not be easily understood by consumers. Of note, the software used to calculate
290 the measure stratifies the population into high and low risk and the incidence of decubiti is
291 much greater in the high risk population; however, the ability of the public to interpret such
292 information was of concern. The TAP did not recommend advancing the measure, primarily
293 because of the potential for public misinterpretation of the results. In recommending the
294 measure for endorsement, the Steering Committee noted that while the value of the measure
295 argues for its advancement, public interpretation of the measure does pose challenges. For that
296 reason, it recommended that guidance regarding interpretation of measure results be included
297 in public reporting.

298
299 Transfusion reaction – adult and pediatric. The value to the consumer of expressing these
300 measures as a rate when such events occur rarely was raised by the TAP and echoed by the
301 Steering Committee. Additionally, it was noted that one of the NQF-endorsed serious
302 reportable events calls for reporting such events individually (counts) as they occur. The
303 Steering Committee recommended advancing the measures contingent on their being reported
304 as counts, rather than a rate.

305
306 Death among surgical inpatients with serious, treatable complications. This was one of two
307 similar measures considered by the Patient Safety TAP, which ultimately recommended this
308 measure based on determination that the measure was the more ‘actionable’ of the two. While
309 it captures a limited set of five complications, this fact focuses it on a set of complications that
310 permits hospitals to pinpoint the issues that contributed to events and enable focused corrective
311 actions quickly and directly. Additionally, this measure has been reconciled with the NQF-
312 endorsed measure by the same name in the nurse-sensitive set; each now use the same
313 numerator and denominator and will continue to be aligned and maintained by AHRQ as a
314 single measure. The Steering Committee recommended advancing the measure in this set as

A-10

315 | specified and further recommended [that the measure, as specified, be advanced as a materially](#)
316 | [changed update to the NQF-endorsed nurse sensitive measure of the same name. Additionally,](#)
317 | [it recommended](#) that a parallel measure for pediatrics be explored by the developer.

318
319 | [Acute stroke mortality rate.](#) This measure remains relevant to consumers because of the
320 | ongoing concern regarding timely treatment despite the fact that no more than 15 percent of
321 | stroke deaths occur in hospital. While there is general agreement that a 30 day mortality
322 | measure would be useful, an in-hospital mortality measure can be collected in real time where
323 | 30 day data is generally not available in under a year. The Steering Committee recommended
324 | advancing the measure as specified and further recommended that the developer explore the
325 | development of two companion measures related to 1) patient functional status post-stroke and
326 | 2) 30 day mortality. A comment from the audience expressing concern that the risk adjustment
327 | model results in the measure inappropriately including hemorrhagic and subarachnoid events
328 | rather than only ischemic events was heard and while it did not affect the Committee's action, it
329 | has been provided to the developer for review and appropriate action.

330
331 | [Hip fracture mortality rate.](#) This measure is widely used and the data is easy to collect though
332 | there was some concern expressed that it could be unfair to hospitals if the fracture occurred
333 | days prior to admission as this reduces the institution's ability to prevent morbidity/mortality.
334 | It was noted that there is literature that shows some variability in admission delays; however,
335 | most patients are admitted within 12 hours of fracture. Studies that have looked at percentages
336 | of admissions within specified periods are not population-based and since administrative data
337 | does not include time of fracture, the issue of time delay has not been evaluated. There was
338 | discussion regarding whether there is data to suggest that some element around the incidence
339 | of hip fracture mortality varies systematically across hospitals such that some hospitals will be
340 | more vulnerable to rate elevation; there is none. It was noted that, if properly identified and
341 | coded, the measure's risk adjustment will pick up the co-morbidities resulting from admission
342 | delay. The Steering Committee recommended advancing the measure contingent on the age
343 | specification changing to 65 from 18 as the AHRQ advisory panel has recommended.

344
345 | [Bilateral cardiac catheterization rate.](#) By reporting information about the rates of bilateral cardiac
346 | catheterization, this measure offers an opportunity to evaluate potential overuse or

A-11

347 inappropriate use of the practice. Concluding that the concern regarding over/inappropriate
348 use is strongly supported by the evidence and that the measure is valid and reliable, the
349 Steering Committee recommended advancing the measure.

350
351 Blood cultures performed within 24 hours prior to or 24 hours after hospital arrival for patients
352 who were transferred or admitted to the ICU within 24 hours of hospital arrival.

353 Approximately 12 percent of patients who are admitted to the intensive care unit (ICU) either
354 directly or in transfer have the diagnosis of pneumonia. The national average for performance
355 on this measure is 91 percent with considerable variability based on data from approximately
356 30,000 admissions in the first quarter of 2007. There was discussion about whether there is
357 value in advancing a measure with a relatively high level of performance as well as whether
358 there is an association between performing blood cultures and patient outcomes. Information
359 available to the Committee indicated the evidence for an association is sparse and conflicting.
360 The Committee determined that the behavior called for by the measure is the right thing to do
361 and recommended advancing the measure with the additional recommendation that the
362 developer refine the language of the numerator and denominator for improved clarity.

363
364 Congestive heart failure mortality. This measure is similar to an NQF-endorsed congestive
365 heart failure (CHF) 30 day mortality measure; however, the endorsed measure applies only to
366 the Medicare Fee for Service population. The developer provided information that this in-
367 hospital measure and endorsed 30 day mortality measure have a correlation of approximately
368 90 percent. The Committee opined that using the two measures together could result in
369 learning more than either alone would allow and recommended advancing the measure.
370 Additionally, committee members recommended a parallel 30 day measure be explored by the
371 developer.

372
373 A clarification of importance to eight of the AHRQ measures was made regarding the use of
374 APR-DRGs with the measures. The intellectual property owner, 3M, has provided
375 documentation to AHRQ that a website that explains the methodology, algorithms, etc. will be
376 accessible to the NQF membership during the consensus development process and that it will
377 provide the necessary information regarding the APR DRG methodology so that it can be
378 included in the relevant measure specifications for the final public consensus report.

379

380 Pediatrics

381

382 **PICU Severity-adjusted Length of Stay, PICU Unplanned Readmission Rate, Review of**

383 **Unplanned PICU Readmissions.** These three measures are advanced for endorsement

384 individually with the recommendation that they be reported as either a pair – length of stay and

385 unplanned readmission rate or the bundle of three. [The committee does not recommend that](#)

386 [any be reported as ‘stand alone’ measures.](#) In arriving at this recommendation, the Committee

387 took into consideration information it had requested from the developer and noted that length

388 of stay and unplanned readmission rate together has the potential to illuminate issues around

389 premature discharge and post-discharge care. It also noted that review of unplanned

390 readmissions was an appropriate companion measure to the unplanned readmission rate, but

391 would have little value if reported only with length of stay. The risk-adjustment methodology

392 was discussed in terms of the need to ensure that the methodology used was standardized and

393 in the public domain. Committee members posed a series of questions to the developer related

394 to the risk adjustment methodology, whether the developer recommended the three as a bundle

395 or paired, the definition of ‘unplanned readmission’, and how the reliability and validity of the

396 unplanned readmission review could be enhanced. The developer has responded to the

397 Steering Committee with documentation that it is revising the measure specifications as needed

398 to reflect that:

- 399 • PICU Severity-adjusted Length of Stay will include additional specifications to indicate
400 that PRISM III is the only risk-adjustment methodology for use with the measure in that
401 it meets the criteria of residing in the public domain (with endorsement of the measure),
402 widespread use, validation in the U.S., and a mechanism in place for ongoing validation
403 and recalibration;
- 404 • it proposes that only the PICU Severity-adjusted Length of Stay and PICU Unplanned
405 Readmission should be reported publicly together;
- 406 • the definition of ‘unplanned readmission’ is clarified to mean a readmission to the PICU
407 within 24 hours of discharge or transfer for which there is no pre-existing
408 documentation of a planned readmission. Exclusion criteria: All planned readmissions
409 as identified by pre-existing documentation in chart; e.g., surgical note, physician note;
410 and

A-13

- 411 • it is working with The Leapfrog Group to operationalize an approach to enhance the
412 reliability and validity of the unplanned readmission review.

413

414 **Home management plan of care document given to patient/caregiver.** While the Steering
415 Committee recommended advancing this measure, it expressed concern that it requires only
416 that a document be provided; the measure can be met without the important components of
417 education and care coordination actually having taken place. The Steering Committee stressed
418 that education is an important component of care which could reasonably include a plan of care
419 but noted that there is no evidence that the simple presence of a written care management plan
420 affects outcomes. Over the course of consideration of this measure by the TAP, specifications to
421 the measure were modified to reflect that “arrangements for a follow-up appointment” were
422 initiated to demonstrate an effort at care coordination. The Steering Committee also considered
423 the fact that the measure includes evidence-based medication-related elements and that its use
424 to date has shown there is variability in providing a plan of care. This discussion lead to
425 acknowledgement that moving the field to include a plan of care would be a positive start. The
426 Committee determined that the measure points to a number of areas that should be targeted for
427 improvement including care coordination and patient education and that implementation of
428 measure could stimulate quality improvement. It challenged the developer to continue to refine
429 the measure to require evidence that educational effort has occurred and to separate the
430 evidence-based elements from those for which there is little or no evidence.

431

432 **Pediatric Heart Surgery Mortality, Pediatric Heart Surgery Volume.** The Steering Committee
433 noted that the TAP had taken the step of confirming that pediatric heart surgeons agreed with
434 the timeframe and the notion of public accountability. The Committee readily agreed to
435 appropriateness of advancing these two measures; its discussion of the measures centered
436 around the value of each of the measures individually and how the measures should be
437 reported. The issues discussed included:

- 438 • in some pediatric populations some of the procedures are so rare that it would be difficult
439 to interpret;
- 440 • reporting volume is relatively burden free but often the data elements necessary to address
441 mortality rate and the severity adjustment are inaccessible;

A-14

- 442 • whether the consumer can make a useable decision on the basis of volume or mortality
443 alone;
- 444 • states' regulation of facilities when relationships between volume and quality are identified
445 in literature, without looking at mortality;
- 446 • concern over potential generalization of volume and mortality rate pairing to the
447 physician level; and
- 448 • that the volume measure is a descriptive structural measure that has predictive value
449 about outcome and for which there is research which provides evidence of this value,
450 especially for high risk procedures.

451 After considering the issues, the Steering Committee strongly recommended that the measures
452 be reported together. However, when it is not possible to report mortality, volume may be
453 reported alone.

Deleted: should

Deleted: but that

Deleted: feasible

454

455 **PICU pain assessment on admission, PICU periodic pain assessment.** The Steering
456 Committee questioned why the two measures were limited to the PICU and was advised that
457 the measures were developed by pediatric intensivists with the focus on PICU. A concern for
458 the Steering Committee, which had been identified by the TAP also, is the lack of specificity
459 around the pain assessment tool to be used. Because there is a lack of consensus about the
460 appropriate tool to use for the range of ages encompassed by the measure, a specific tool(s)
461 cannot be included in the specifications, though the developer can and will provide a list of
462 generally accepted tools/scales with the measure on its website. Recommendations were made
463 that research be pursued to expand the PICU measures to the entire pediatric population and
464 that research be conducted into pain assessment tools in an effort to standardize tool(s) for use.

465

466 **PICU standard mortality ratio.** The Steering Committee questioned why deaths occurring after
467 transfer from the PICU were excluded and was advised that, as in the case of the pain
468 assessment measures, this measure was developed by pediatric intensivists with a focus on care
469 in the PICU. The Committee was advised that the developer will incorporate a palliative care
470 exclusion requested by the TAP; that exclusion reads, "*Children who were NOT admitted for the
471 purpose of critical care intervention or monitoring i.e. related to their real or potential risk of physiologic
472 instability, but instead because there was no other location in the hospital to provide these end of life
473 services.*" The Steering Committee questioned certain variables listed in the risk-adjustment,

474 including the use of PRISM III. In response to information that this model has been tested for
475 use in the PICU and is the most widely accepted and that adult risk-adjustments also adjust for
476 previous hospitalizations; the Steering Committee approved advancing the measure.

477

478 **Surgery and Anesthesia**

479

480 Prior to beginning the discussion of individual surgery and anesthesia-related measures, the
481 Steering Committee discussed pairing volume and mortality measures for publicly reporting.
482 After clarifying that there are no additional costs or undue effort to download the needed
483 software for the AHRQ quality indicators, the Committee recommended that measures of
484 mortality should always be paired with volume counterparts but when reporting mortality is
485 not feasible, that volume may be reported alone.

486

487 **Abdominal Aortic Aneurysm Volume (AAA), Abdominal Aortic Aneurysm (AAA) Mortality**

488 **Rate.** After hearing the strengths (ease of measure calculation and the minimal burden caused
489 by use of administrative data) and weaknesses (modest evidence that volume is a predictor of
490 adverse outcomes, heterogeneity of the population captured, data collection ill defined) of the two
491 measures, the Steering Committee had a lengthy discussion around the fact that in most cases,
492 when an AAA has ruptured/is rupturing, transfer to institutions with high volume of these
493 surgeries is not an option. It also noted that the differences between surgery for ruptured
494 aneurysm and elective aneurysm repair may not be properly differentiated in claims. The
495 developer noted that the risk-adjustment accounts for the differences in elective versus ruptured
496 aneurysms and stated that it is in the process of clarifying definitions to decrease mis-
497 classifications. As noted, the Steering Committee recommended that the mortality measure should
498 always be reported with volume, when feasible.

499

500 **Esophageal Resection Mortality Rate, Esophageal Resection Volume.** In advancing these two

501 measures, the Steering Committee considered the fact that esophageal resection is both rare
502 (three cases could be considered high volume) and elective and that there is no consensus
503 regarding minimum practice volumes. Its primary concern was the low volume. The Steering
504 Committee advances the measures with the recommendation that mortality and volume be

505 reported together, but due to the low numbers for the procedure that volume may be reported
506 alone if reporting mortality rate is not feasible.

507
508 **Incidental Appendectomy in the Elderly Rate.** In advancing this measure, the Steering
509 Committee agreed with the TAP that the procedure should not occur and that there is wide
510 variability among hospitals, thus an opportunity for improvement exists. It noted that the
511 developer is revising the definitions to better clarify what constitutes “incidental” in order to
512 increase the accuracy of the measure.

513
514 **Pancreatic Resection Mortality Rate, Pancreatic Resection Volume.** As with other measures
515 that rely on administrative data, the data collection burden is minimal and because of the nature
516 of the procedure, there is little miscoding. However, there are few evidence based processes
517 that can decrease mortality. Though the measures do not currently differentiate between a
518 Whipple procedure and total pancreatectomy, the risk-adjustment will have been revised to
519 address this prior to the time of endorsement. The Steering Committee advances the measures
520 with the recommendations that mortality and volume be reported together, but that volume
521 may be reported alone.

522
523 **Post operative Wound Dehiscence, Post operative Wound Dehiscence** The Steering
524 Committee voted to advance the two measures with relatively brief discussion agreeing that the
525 measures are strong, that there are processes of care that can reduce likelihood of dehiscence,
526 and that they are important; in fact that there should be zero tolerance.

527
528 **Foreign Body left after procedure - pediatric, Foreign Body Left in after Procedure – adult.**
529 Two key issues were of concern during deliberation of these measures: 1) the fact that
530 identification of the foreign body as ‘present on admission’ (POA) is an option rather than a
531 requirement and 2) presentation of the measure as a percentage rather than a count since the
532 events occur so rarely. With respect to the former, the developer agreed to change POA from an
533 option to a requirement. With respect to the latter, the concern regarding value of a measure for
534 public reporting that has very low rates was the central concern and the developer has agreed to
535 change the specifications of the measure to a count, rather than a rate, in future releases of the
536 software. The relationship of these measures to the NQF-endorsed™ serious reportable event

A-17

537 (SRE) that addresses unintended retention of a foreign body after surgery was discussed;
538 however, the measures provide the ability to audit the occurrence and to compare results across
539 organizations. The measures were recommended for advancement contingent on the
540 requirement of POA and reporting the results of the measure as a count. The importance of the
541 SREs was acknowledged and a recommendation was made that a project to endorse measures
542 for the full list of SREs be undertaken.

543
544 Failure to rescue. This measure, which addresses death among general surgery, orthopedic,
545 and vascular patients with or without documented complications, was considered by two TAPs,
546 Patient Safety and Surgery and Anesthesia, the second at the request of the Patient Safety TAP,
547 which asked that it be evaluated for appropriateness as a surgical mortality measure after
548 recommending advancing the similar measure, “Death among surgical inpatients with serious,
549 treatable complications”, as a patient safety measure. This measure was submitted with death
550 defined as death within 30 days from admission though the developer noted that it could be
551 defined as in-hospital mortality and that similar results had been seen with both definitions.
552 The discussion of the measure focused on three issues: 1) value and challenges associated with
553 capturing 30 day mortality; i.e., complexity associated with ascertainment of the data and
554 attribution of results; 2) complexity associated with the number of complications to be captured
555 and acted upon, acknowledging that the number of complications result in capturing 50 percent
556 more events; and 3) ability to reliably discern hospital-associated issues that can be acted upon
557 to improve rates. Obtaining the 30 day data was acknowledged as a challenge and will require
558 more effort on the part of hospitals. It was determined that the scope of the measure in terms of
559 the complications captured make it sufficiently different from the similar measure to warrant its
560 advancement in the form of two separate measures – an in-hospital mortality measure and a 30
561 day mortality measure.

562