

“Failure to Rescue”: An Imperfect Measure Well Suited to Complement an Imperfect World

Ralf J. Holzer, MD, MSc

Aiming for improved quality of care for our patients is a non-controversial goal. To be able to improve care, it is important to measure and compare outcomes, which is where it becomes complicated. The patient population in the pediatric/congenital catheterization laboratory (PCCL) is heterogenous with a wide variation in case complexity and acuity. This is often combined with a fairly low case volume when compared with adult centers. Furthermore, important adverse events (AE) are rare, and validated evidence-based measures of procedural efficacy are often lacking.

Registries have aided the process of comparing outcomes, by collecting a large amount of data, and providing participants with specific metrics, that facilitate comparisons of outcomes as well as ranking of hospital to each other. While some metrics like the standardized AE ratio (CHARM [Catheterization for Congenital Heart Disease Adjustment for Risk method]), or risk standardized AEs (IMPACT [Improving Adult and Congenital Treatment registry]) have been well studied and validated,^{1–3} the majority of metrics presently used are based on expert consensus, with associated shortcomings and limitations.

The importance of outcome metrics cannot be understated, as they are the backbone to facilitate and measure quality improvement efforts. However, adding and using metrics lightly without thorough validation, poses potential problems. This is even more so apparent in the context of discussions on public outcomes reporting. As an example,

institutional ranking systems such as *US News and World* report frequently combine many different quality metrics. The methods of how these metrics are chosen, combined, and weighted is a subject of regular controversy with a myriad of conflicting interests complicating this process.⁴ Factors such as risk averseness and the impact of salvage procedures on those rankings have not yet been addressed satisfactorily for many centers, and the congenital cardiac catheterization community can and should learn from similar discussions taking place surrounding STS (Society of Thoracic Surgeons) data for our surgical colleagues.

Summary of Methods and Findings

In this issue of the *Journal of the American Heart Association (JAHA)*, O’Byrne and colleagues report on “Failure to Rescue” (FTR), a new metric suggested as a quality marker for PCCL.⁵ FTR was defined as the occurrence of a catastrophic AE, after a proximal AE during a catheterization procedure. The authors used a data set derived from the IMPACT registry and its associated AE definitions. Catastrophic AEs included death within 2 days from cardiac catheterization, cardiac arrest, or initiation of mechanical/ECMO (Extracorporeal Membrane Oxygenation) support, or any unplanned cardiac, vascular, other surgery, or cardiac catheterization because of a catheterization complication. Proximal AE were defined as new arrhythmia, new valve regurgitation, cardiac tamponade, air embolus, embolic stroke, device malposition, device embolization, airway event, or initiation of dialysis. Cases with repeat catheterization or surgery because of a catheterization complication were reported as having both a proximal and a catastrophic AE. Covariates were identified that loosely mirrored parameters that have been used previously within IMPACT and CHARM, including patient and procedure specific parameters, as well as parameters reflecting hospital programmatic quality (case volume). Risk standardized ratios (RSR) were calculated for each participating hospital for FTR, proximal AE, all/pooled AE (with/without death), and catastrophic AE. Hospitals were ranked based on RSRs, and the ranking compared between FTR, all/pooled AE, and catastrophic AE.

The opinions expressed in this article are not necessarily those of the editors or of the American Heart Association.

From the NewYork-Presbyterian Komansky Children’s Hospital, New York, NY; Weill Cornell Medicine, New York, NY.

Correspondence to: Ralf J. Holzer, MD, MSc, FACC, FSCAI, Chief, Division of Pediatric Cardiology, Department of Pediatrics, Director Pediatric Cardiac Catheterization, NewYork-Presbyterian Komansky Children’s Hospital, Weill Cornell Medicine, 525 East 68th Street, Room F-677, New York, NY 10065. E-mail: rjh3001@med.cornell.edu

J Am Heart Assoc. 2019;8:e014356. DOI: 10.1161/JAHA.119.014356.

© 2019 The Authors. Published on behalf of the American Heart Association, Inc., by Wiley. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

The authors report on a data set containing 77 580 procedures, performed in 53 056 individual subjects, in 91 hospitals. Any AE occurred in 4.7% of cases, catastrophic AE in 1.2%, and proximal AE in 4.4%. The risk of catastrophic AE after proximal AE was 20.3%. The precedence rate (proportion of catastrophic AEs with a preceding proximal event) was 70%. The authors found that the adjusted risk of FTR was significantly lower at higher volume hospitals (odds ratio: 0.68), whereas the same did not apply for all/pooled AEs. However, increasing PCCL volume was also associated with significantly reduced odds of catastrophic adverse outcomes (odds ratio: 0.79). All/pooled AE and catastrophic AE had significant correlations with patient and procedure specific factors, findings that were not consistently seen for FTR. When comparing methods of hospital ranking, there was no correlation when comparing rankings by RSR for All AE and RSR for FTR, but there was a strong correlation between rankings based on RSR for catastrophic AE and RSR for FTR (Spearman $r=0.65$). Rankings based on hospital volume were significantly associated with FTR RSR, and also catastrophic AE volume RSR.

Discussion of Limitations

This study continues to explore the ability to identify a “perfect” metric to accurately reflect and rank the outcome of pediatric/congenital cardiac catheterization laboratories. FTR is an important concept, but some issues related to its use and the data presented in this paper warrant further discussion. In order of importance, these include the following:

FTR Versus Catastrophic AE

While FTR has some intuitive validity, it is difficult to see the advantage over the use of catastrophic AEs. The authors make a good point highlighting why FTR has advantages over pooled AE, yet the same cannot be said in relationship to just using catastrophic AE. Similar to other studies, the authors identified significant associations between patient and procedure level factors, and catastrophic AEs. Furthermore, in contrast to pooled AE, an increasing PCCL volume was associated with reduced odds of catastrophic adverse outcome (odds ratio: 0.79). This brings up the question what benefit can be derived from using FTR as opposed to catastrophic AE, given that catastrophic AEs have been shown in this study to correlate not just with patient and procedure factors, but also procedural volume? Independence of case mix was highlighted as a benefit of FTR, but this benefit falls flat when you use risk adjustment/risk standardization methods. Furthermore, FTR also had several patient and procedure factors significantly associated with it, but in a

less consistent manner. There was a positive association with pre-procedural receipt of inotropes or low mixed venous saturations, and a negative association of procedure risk categories 2 and 3 as opposed to 1, or surgery within the past 30 days. This seems to be somewhat inconsistent, as all those factors taken independently would signify a potential higher risk, yet some appeared to have a positive and some a negative association with FTR. It raises the question whether the (potential) lack of causality between proximal AE and catastrophic AE may have contributed to this inconsistency.

In addition, there was also a significant association between procedural volume rank of hospitals and catastrophic volume RSR, just as there was between procedural volume rank and FTR RSR. Taken all this into account brings up the question what additional advantage FTR can provide over just using catastrophic AE, especially given the limitations related to causality of the proximal AE. Furthermore, given the limitations of using a metric with a low occurrence, there is a small but valid difference of 38% between FTR (694 cases) and catastrophic AEs (960 cases).

Causality of Proximal AE

The authors state that “The proposed FTR metric studies the progress from AE to a catastrophic AE”. This is, however, incorrect as IMPACT does not provide any information on causality or attributability. In reality, most catastrophic AEs are imminently recognizable in the catheterization laboratory. One could argue that the lack of causality may not be all that important as long as the metric performs well. However, why not then use catastrophic AEs alone, rather than failure to rescue? If the authors want to analyze the progression of an AE from minor to catastrophic, then causality between the proximal and catastrophic event is absolutely crucial.

FTR and Hospital Procedural Volume

The authors chose hospital annual catheterization volume as a surrogate for the quality of a PCCL program. This is problematic as there has been little consistent evidence that this also applied to the congenital catheterization laboratory, where patients are usually discharged within 24 hours and any prolonged stay usually occurring in patients that were already admitted for other reasons. In fact, volume definitions in IMPACT are likely misleading, as a proportion of centers participating in IMPACT share a combined adult (structural and PCI [Percutaneous Coronary Interventions]) and pediatric catheterization program. Those centers may have a low pediatric volume, but the volume of overall catheterization procedures within the laboratory (and henceforth experience) may be higher than in isolated pediatric centers that have a higher volume documented in IMPACT. This is supported by

the fact that centers with an adult congenital contribution of >35% had a lower FTR than other centers. Furthermore, using volume as a surrogate for quality of a PCCL implies volume being directly related to catheterization laboratory experience and outcomes. It ignores that centers with a large transplant and pulmonary hypertension program may be performing a significant number of catheter procedures, but interventional procedures (that often carry the highest risk) may only account for a small portion of the overall catheterization volume. In contrast, a center without a large transplant and PHTN (Pulmonary Hypertension) program may perform only half the overall cases, but this may include a much larger number of interventional cases. None of those elements is captured in the present study. It is also worthwhile briefly commenting on operator experience that was mentioned by the authors: A study by Holzer and colleagues has shown that both, junior and senior operators have significantly higher associated AE.⁶ Both types of operators are more likely to be found at larger high-volume centers, rather than smaller or medium sized center. Taken all this into account, there has to be caution associating high procedural volume with a quality marker of the congenital catheterization laboratory.

FTR to Discriminate Outcomes in “Real-Life”

A metric used for ranking is only as good as it allows to discriminate between centers and operators. While the authors acknowledge that the use of FTR would make “quarterly reporting challenging”, this is a mild description of a major problem. As an example, a low-volume catheterization laboratory with 100 cases a year may only see 1 FTR event per year (in this study there were 694 FTR, 0.9%). To have any meaningful use it would require many years’ worth of rolling data, which makes it impossible to identify shortcomings in any reasonable time. Even catastrophic AEs are still rare (in this study: 960/1.2%). This repeats the problems IMPACT and other registries have faced when using some important but rare AE (such as device embolization): These metrics are too infrequent to allow any meaningful comparison and ranking. Instead, expanding the capture of AE to for example higher severity AE (level 3–5, as defined in C3PO⁷), strikes some balance between avoiding to incorporate lower level AE, while still providing a sufficient numerator to aid comparison between centers.

IMPACT Data Set

This study relied on the data set and AE definitions inherent to the IMPACT registry.⁸ This is the greatest strength, but also one of the greatest weakness of this study. Strength as the data set is large and therefore allows analysis of data, that would otherwise not be possible. Weakness, as it relies on specific AE definitions, that are selective and often based on

adult experience (such as initiation of dialysis). The AE do not allow a broad capture based on severity levels (like C3PO does), and higher severity AEs that may not fall into those specific categories (such as for example hypotension requiring inotropic support) would not be captured as a proximal AE.

Selection and Timing of Catastrophic AE and Case Exclusions

Including a post-procedural period of 2 days for death has the potential of capturing death attributable to factors other than the cardiac catheterization procedure. The authors tried to account for this by performing an additional analysis by excluding subjects that had undergone surgery within the previous 30 days with similar results. However, it is not cardiac surgery alone that has a potential impact in the post-catheterization period. A multicenter study looking at attributability of death within 30 days of cardiac catheterization, found only 10% of those deaths related to the catheterization procedure itself, 14% to cardiac surgery, and the remaining 76% to a variety of other factors (such as non-cardiac comorbidity).⁹

The authors also excluded emergent and salvage procedures, as those may have a potential impact in rewarding risk averse centers and operators, and penalizing those who take on higher risk cases. While exclusion of those cases has the potential to disregard potential problematic case selections that could also be classified as a marker of hospital quality, it is clearly important to avoid any disincentive to taking on higher risk cases. It is equally important though to avoid any suggestion of risk-overtness being linked to lower volume centers.

The inclusion of urgent cases may be problematic, as many of the highest risk cases may in fact often be classified as urgent, rather than emergent or salvage. These are often patients that may have been observed for several days with a strategy to perform a cardiac catheterization if the patient does not improve. When the procedure is finally performed, this would have to be classified as urgent but would not meet the classification of emergent or salvage, even though these patients usually pose an extremely high risk—Asoh and colleagues reported an overall mortality of 43% in patients undergoing cardiac catheterization after cardiac surgery and before discharge.¹⁰

FTR as a Quality Measure for a PCCL Program

There is a conceptual problem relating to the use of FTR to characterize a PCCL program. The authors point out that FTR was initially developed in a cohort of adult surgical patients.¹¹ However, this is a different patient cohort than pediatric congenital catheterization patients. In contrast to a surgical

patient, catastrophic AEs usually manifest themselves at the time of cardiac catheterization and/or are clearly linked to the procedure. There is rarely a post-procedural period where catastrophic problems suddenly manifest themselves that may or may not be related to the procedure, and where the skill of the staff to “shepherd” a patient through a hospital admission becomes important. One of the main justifications for the use of FTR was its perceived ability to be better associated with hospital factors, rather than patient and procedure factors. In reality though, for the majority of catastrophic AE, there is little relationship with post-catheterization care, and the outcome was/is determined by what happens in the catheterization laboratory.

Summary and Perspective

To summarize, O’Byrne and colleagues report on FTR, a new metric to complement the existing outcome metrics for the PCCL. On its face value, it provides data on the progression of an AE to a catastrophic outcome, which is an important aspect of the quality of care provided in a PCCL. As an example, the outcome of a tear of a vessel after balloon angioplasty or stenting is dependent on a multitude of factors, including how well the staff is prepared to have the right equipment and additional staffing immediately available, how many covered stents (and what sizes) the laboratory has at its disposal, whether you have blood available immediately, and how well you can mobilize surgical support, to name just a few.

In this article, the authors ask an important question, namely whether “risk-adjusted AR rate is the single, optimal quality measure for PCCL programs”? Clearly the answer has to be no, but FTR is not necessarily the answer either.

In a perfect world, the ideal metric used for PCCL programs would have several important characteristics:

1. Frequent “numerator” to facilitate:
 - Fine-graded discrimination between centers.
 - Analysis of short-medium time periods (3–6 months) rather than requiring many years of rolling data.
2. A composite of AEs and procedural efficacy.
3. Adjusted/standardized for patient and procedure factors.
4. Easy and unambiguous to interpret and understand by providers AND patients.
5. Accounts for potential long-term outcome (eg, implanting a small diameter stent with low profile may give great acute results and acute outcomes, until the surgeon has to operate to remove the stent after 2 years).
6. Procedural volume does not impact ability to achieve top (or bottom) ranking.
7. No disincentive to taking on higher risk cases.
8. It measures what the end-user wants to assess:

- Patient: overall outcome.
- Provider: individual components of a PCCL program.

When considering some of the desirable qualities of metrics listed above, it becomes clear that something like an “ideal metric” is impossible to achieve. It is almost like the ideal stent—you often have to find the best compromise of combining sometimes fairly opposite characteristics (such as flexibility and radial strength).

The authors listed several potential advantages of using FTR. They make an important point that the usage of FTR may remove a “disincentive for reporting these (potential equivocal)” proximal AEs, which holds true for current practice. It is, however, an issue that should eventually be able to get resolved with appropriate source document audits (which is already performed by IMPACT). Most of the proximal AE events should be fairly easily identifiable by reviewing records such as catheterization reports, technician report, and echocardiography reports. The authors also state that usage of FTR would remove a disincentive to take on high risk cases, but this also holds true for catastrophic AE, if one were to remove emergent and salvage procedures.

Despite the concerns relating to the use of FTR, none of the already existing metrics is perfect—not even close. Most of the outcome metrics presently used in IMPACT were defined based on expert consensus, and often have limited value and potential associated problems. For example, the ASD (Atrial Septal Defect) metrics use residual shunts and device embolization. Yet those do not take account of rim deficiencies and risk of erosions related to device size: By oversizing we reduce the chance of device embolization and residual shunts, yet we expose the patient to a higher risk of erosions, which is much more consequential to the patient, but does not get captured in the IMPACT metrics. The metrics also do not take account the surgical risk of 0.29% STS mortality: are patients exposed to a higher mortality because a center decides to send anyone with a slight rim deficiency for surgery?

Taking all this into context, FTR does have its place and can complement other existing metrics. The usage and validity of this metric could, however, be further enhanced by validating causality between proximal AE and catastrophic outcomes, through further studies. This could be easily performed, if a study were to retrospectively ask centers to review their FTR events, and assess causality/attribution.

Disclosures

None.

References

1. Bergersen L, Gauvreau K, Foerster SR, Marshall AC, McElhinney DB, Beekman RH, Hirsch R, Kreutzer J, Balzer D, Vincent J, Hellenbrand WE, Holzer RJ,

- Cheatham JP, Moore JD, Burch GE, Armsby L, Lock JE, Jenkins KJ. Catheterization for congenital heart disease adjustment for risk method (CHARM). *JACC Cardiovasc Interv.* 2011;4:1037–1046.
2. Jayaram N, Beekman RH III, Benson L, Holzer R, Jenkins K, Kennedy KF, Martin GR, Moore JW, Ringel R, Rome J, Spertus JA, Vincent R, Bergersen L. Adjusting for risk associated with pediatric and congenital cardiac catheterization: a report from the NCDR IMPACT Registry. *Circulation.* 2015;132:1863–1870.
 3. Jayaram N, Spertus JA, Kennedy KF, Vincent R, Martin GR, Curtis JP, Nykanen D, Moore PM, Bergersen L. Modeling major adverse outcomes of pediatric and adult patients with congenital heart disease undergoing cardiac catheterization: observations from the NCDR IMPACT Registry (National Cardiovascular Data Registry Improving Pediatric and Adult Congenital Treatment). *Circulation.* 2017;136:2009–2019.
 4. Galantowicz M, del Nido P, Hanley F, Gaynor JW. Correspondence to the pediatric cardiology & heart surgery rankings. U.S. New. Usnews.com. Published October 2017. Available at: <https://health.usnews.com/health-news/blogs/second-opinion/articles/2017-10-20/correspondence-on-the-pediatric-cardiology-heart-surgery-rankings> 2017. Accessed September 23, 2019.
 5. O'Byrne ML, Kennedy KF, Jayaram N, Bergersen LJ, Gillespie MJ, Dori Y, Silber JH, Kawut SM, Rome JJ, Glatz AC. Failure to rescue as an outcome metric for pediatric and congenital cardiac catheterization laboratory programs: an analysis of data from the IMPACT registry. *J Am Heart Assoc.* 2019;8:e013151. DOI: 10.1161/JAHA.119.013151.
 6. Holzer RJ, Gauvreau K, Kreutzer J, Moore JW, McElhinney DB, Bergersen L. Relationship between procedural adverse events associated with cardiac catheterization for congenital heart disease and operator factors: results of a multi-institutional registry (C3PO). *Catheter Cardiovasc Interv.* 2013;82:463–473.
 7. Bergersen L, Marshall A, Gauvreau K, Beekman R, Hirsch R, Foerster S, Balzer D, Vincent J, Hellenbrand W, Holzer R, Cheatham J, Moore J, Lock J, Jenkins K. Adverse event rates in congenital cardiac catheterization—a multi-center experience. *Catheter Cardiovasc Interv.* 2010;75:389–400.
 8. Martin GR, Beekman RH, Ing FF, Jenkins KJ, McKay CR, Moore JW, Ringel RE, Rome JJ, Ruiz CE, Vincent RN. The impact registry: improving pediatric and adult congenital treatments. *Semin Thorac Cardiovasc Surg Pediatr Card Surg Annu.* 2010;13:20–25.
 9. Backes CH, Bergersen L, Rome JJ, Bativala SP, Glatz AC, Ovunc B, David S, Rivera BK, Haque U, Kollins K, Yin H, Holzer RJ. Quality metrics in cardiac catheterization for congenital heart disease: utility of 30-day mortality. *Catheter Cardiovasc Interv.* 2015;85:104–110.
 10. Asoh K, Hickey E, Dorostkar PC, Chaturvedi R, van Arsdell G, Humpl T, Benson LN. Outcomes of emergent cardiac catheterization following pediatric cardiac surgery. *Catheter Cardiovasc Interv.* 2009;73:933–940.
 11. Silber JH, Williams SV, Krakaure H, Schwartz JS. Hospital and patient characteristics associated with death after surgery. A study of adverse occurrence and failure to rescue. *Med Care.* 1992;30:615–629.
-
- Key Words:** Editorials • cardiac catheterization • congenital heart disease • registry