Star Assessment as a Predictor of PSSA Score

Maggie Fields, Research Intern
Jeremy Raff, Director of Data Analytics

Introduction

The Star Assessment is a "periodic progress-monitoring assessment" designed to give educators accurate estimates of student ability so that they can fine-tune instruction before a high-stakes state assessment (Renaissance Learning, 2022). 2023 was the first year that the School District of Lancaster used the Star assessment. It was implemented in order to gauge students' math and reading levels prior to the PSSA. This report summarizes how well Star was able to predict 2023 PSSA scores for third through eighth grade students. Ultimately, this analysis will guide administrators and teachers in future interpretations of Star data and provide insight into the success of the Star assessment after the first year of implementation.

Research Design

The purpose of this study is to determine the Star assessment's ability to accurately predict a student's PSSA score. Ascertaining predictability is often accomplished through a regression analysis in which there is a dependent variable (in this case the PSSA scale score) and an explanatory variable (the Star score or level). In this study, we use linear regression models for numerical test scores (continuous variables), and logistic regression models for benchmark levels (discrete variables). We also use multiple-linear regression models to account for more than one explanatory variable. For example, how both a student's economic situation and their Star score might affect their PSSA score. The visuals in this report include graphs with regression lines. These lines represent the linear relationship between the explanatory and dependent variables and the predictions of PSSA score that the model would make based on Star score. Each relationship has a correlation coefficient that classifies the strength of the relationship numerically. Correlation coefficients fall between –1 and 1, with 1 being a perfect positive correlation and –1 being a perfect negative correlation. Generally, a correlation coefficient between .40 and .59 is considered a moderate correlation, a coefficient between .60 and .79

is considered strong and a coefficient between .80 and 1 is very strong (Akoglu, 2018).

The sensitivity and specificity of the Star is also included to highlight how often a student's benchmark score on Star and PSSA matched. Sensitivity represents the true positivity rate or how often someone who was expected to receive a positive result (for example, a student who scored proficient for their Star Proficiency Level) also received a positive result on the actual assessment (proficient on the PSSA as well). Specificity refers to the true negativity rate or how often a student who was expected to receive a negative result, did receive a negative result on the actual assessment (for example, when a student who did *not* score proficient for their Star Proficiency Level also did *not* score proficient on the PSSA).

Variables

Both numerical score and benchmark categories were used to analyze Star's ability to predict PSSA scale scores. *Star Proficiency Level* identifies whether students are identified as Advanced, Proficient, Basic, or Below Basic per the Star state benchmark level. *Star Score* refers to the numeric scaled score within the Star assessment. There are two key variables related to PSSA data. First, *PSSA Score* refers to the scaled numeric score students receive for each subject on the PSSA exam. *PSSA Proficiency Level* refers to the category which identifies students as Advanced, Proficient, Basic, or Below Basic.

School identifies the specific building each student attended, and *Grade* indicates their current grade level. Several demographic variables are included as well. These include *Economically Disadvantaged* for students who qualify for free or reduced lunch in the district's school information system. *English Language Learners* are identified as the same status by the district. The variable *Special Education* indicates if a student has an individualized education plan (IEP). Finally, *Chronically Absent* indicates whether students were absent more than 10% of days in the 2022-2023 school year.

Sample Description

Data for this analysis comes from the 2023 PSSA preliminary test file. This includes all third through eighth-grade students who completed the ELA PSSA and the math PSSA assessment at School District of Lancaster. Importantly, this is *not* the final accountability file with attribution for district students, but all test takers associated with the district who also have an IRLA scale score available. These students must have also completed a related ELA or Math Star assessment.

This cohort as a whole consists of 4,273 students across thirteen elementary schools and five middle schools from within the district. 3,915 students are included in the sample for the math assessment analysis and 3,863 students are included in the sample for the reading assessment analysis. Across both samples, students are primarily Hispanic (61.4%), followed by Black or African American (15.4%), White (11.4%), Multi-Racial/Two or More (5.2%), and Asian (4.3%). 90% of students are economically disadvantaged, 20.6% are English language learners, 27.1% have an IEP, and 19.7% are chronically absent.
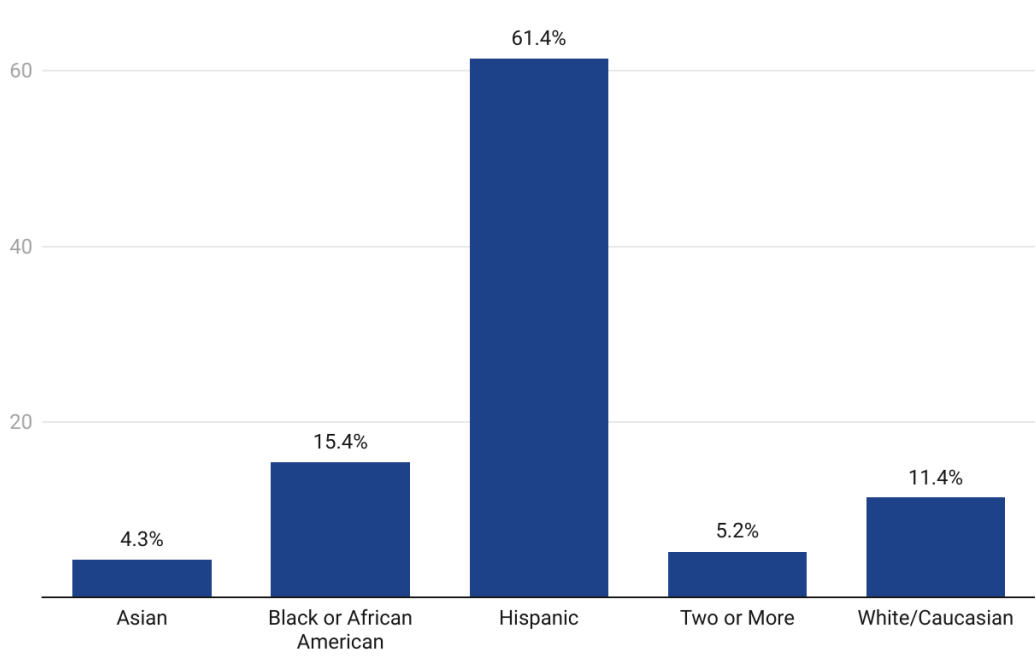

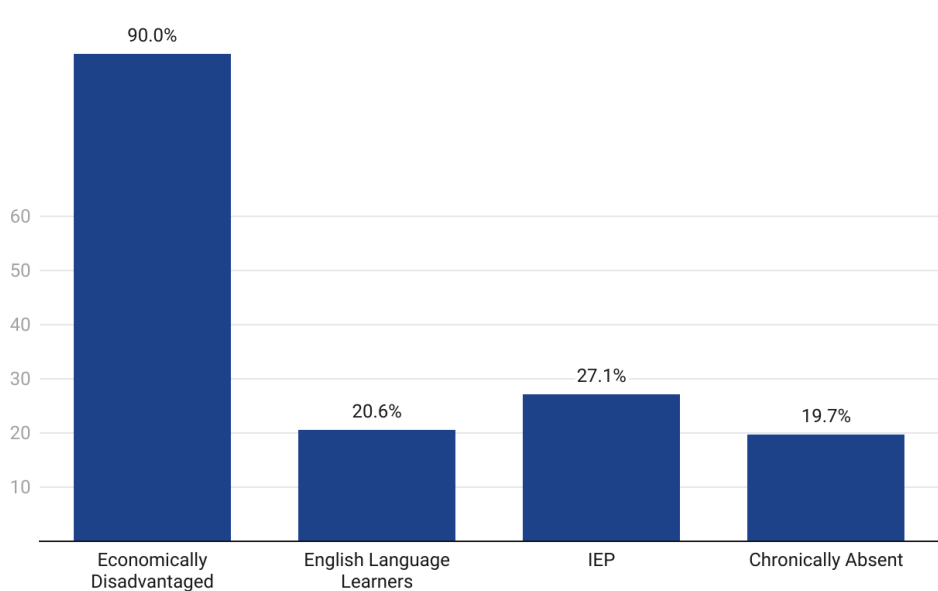Figure 1: Racial breakdown by % of sample

Figure 2: Breakdown of subgroups by % of sample

Results

Benchmark Categories: Sensitivity and Specificity

Math

Both the Star Assessment and the PSSA use the standard categories Below Basic, Basic, Proficient, and Advanced to classify students. The Star assessment specifically utilizes this as part of the state proficiency level provided in the platform. For students who were noted Advanced on the math Star, only **38.8%** scored Advanced on the math PSSA and for those who did not qualify as Advanced on the math Star, **99%** did not score Advanced on the math PSSA. Therefore, specificity is much higher than sensitivity for this group (for this example and other benchmark levels, see the Appendix, Figures 16-19). The math Star correctly estimated Basic, Proficient, and Advanced levels on the math PSSA for only 38-53% of students. For Below Basic, 83% of students had the same benchmark level on both math Star and math PSSA. This indicates that the math Star assessment has a general tendency to overestimate.

	Underestimated	Correct	Overestimated
Advanced	0.00%	38.78%	61.22%
Proficient	7.09%	43.72%	49.19%
Basic	13.68%	53.10%	33.23%
Below Basic	16.19%	83.81%	0.00%

Figure 3: Percent of students at each Math Star level whose Math PSSA level was underestimated, overestimated, or correctly predicted.

Reading

Out of students who were classified as Below Basic on the reading Star, **49.5%** scored Below Basic on the ELA PSSA and **50.5%** did not. For this group, only about half the time did a student's Star level match their PSSA level (for this example and other benchmark levels, see the Appendix, Figures 20-23). For each reading benchmark level, specificity is substantially higher than sensitivity. In other words, when students score a certain Star benchmark category, they do not often get the same benchmark category on the PSSA. From Figure 4 we can see that for each benchmark level, only about 45-65% of students scored at the same level on the reading Star and the reading PSSA.

	Underestimated	Correct	Overestimated
Advanced	0.00%	45.74%	54.26%
Proficient	8.65%	59.33%	32.02%
Basic	23.74%	64.80%	11.46%
Below Basic	50.46%	49.54%	0.00%

Figure 4: Percent of students at each ELA Star level whose ELA PSSA level was underestimated, overestimated, or correctly predicted.

Overall Analysis

Across the entire sample, math Star and PSSA scores had a *moderate* correlation coefficient of **.471**. Star and PSSA reading scores for the sample overall had a *strong* correlation coefficient of **.703**. This cohort, however, includes a wide variety of students in different grades, school buildings, and demographic groups. To gain a better understanding of the Star's predictive capability and how different types of students perform on the Star and PSSA, a closer look at these variables is warranted.
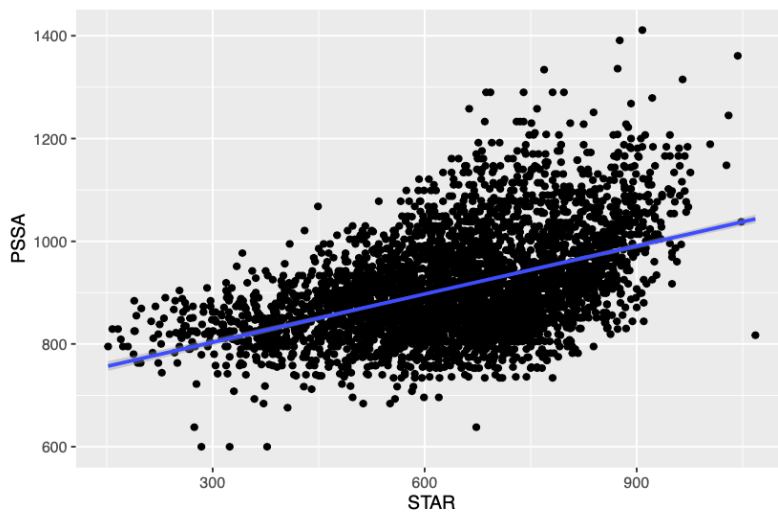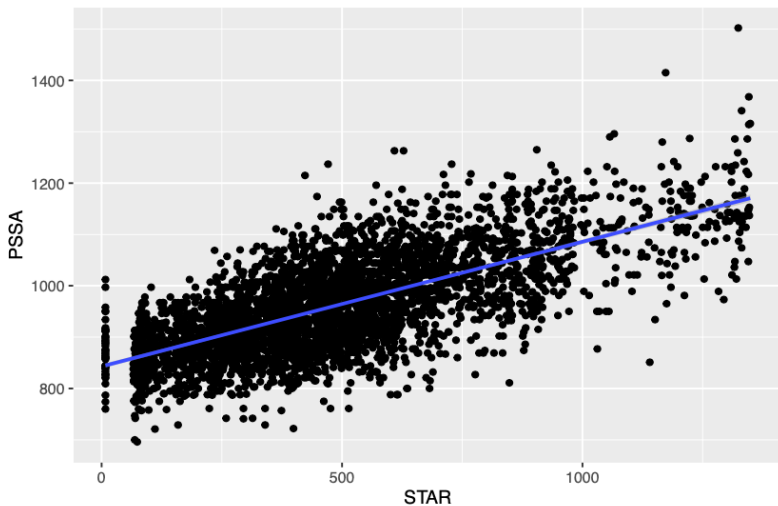

Figure 5: Star and PSSA math scores overall

Figure 6: Star and PSSA reading scores overall

Multivariate Analysis

By the application of variables to the overall regression model, we can see just how much the predictability of PSSA scores is affected by a student's identity. The variables included in the visuals below are economic disadvantaged, English language learner status, chronic absenteeism, and IEP status. Each of these variables' relationship to PSSA score is statistically significant with a p-value of less than .05. The regression analysis for math scores has an adjusted R-squared of .27, meaning only about 27% of the variance in the output was accounted for by the explanatory variables. The multiple regression for reading scores has an adjusted R-squared of .52, so about 52% of the variance was accounted for in the model.
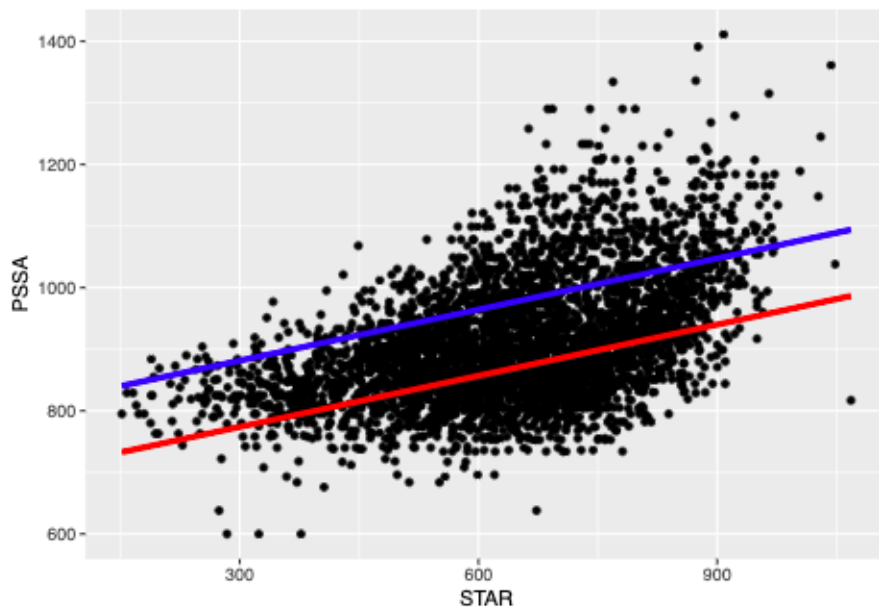
Math

Figure 7: Star and PSSA math scores overall, red multiple regression line

Explanatory Variable	Regression Coefficient
STAR	0.28
Chronically Absent Status	−24.59
Economically Disadvantaged Status	−61.34
English Language Learner	−10.82
IEP Status	−10.81

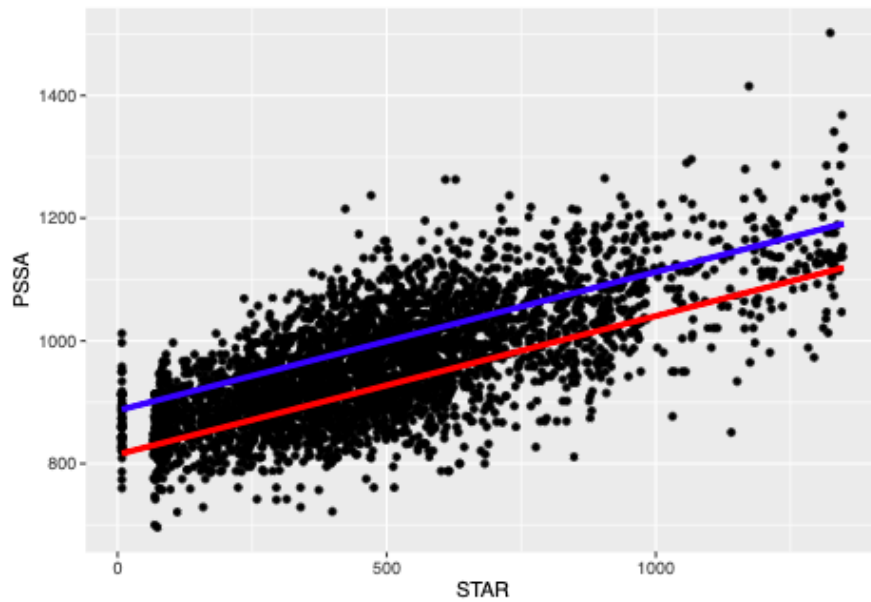Figure 8: Coefficients for the multiple regression line for math scores.

Reading

Figure 9: Star and PSSA reading scores overall, red multiple regression line

Explanatory Variable	Regression Coefficient
STAR	0.23
Chronically Absent Status	−13.02
Economically Disadvantaged Status	−27.32
English Language Learner	−6.66
IEP Status	−10.81

Figure 10: Coefficients for the multiple regression line for reading scores.

Correlation Analysis: Demographic Subgroups

Among students in the identified demographic subgroups, correlations between Star and PSSA were stronger for reading assessments than math. Reading correlation coefficients were between .**55** and .**78**, falling into the *moderate* and *strong* correlation categories. Math correlation coefficients for these same subgroups were between .**37** and .**57**, falling into the *weak* to *moderate* categories.

	Correlation Coefficient
Chronically Absent Status	0.370
Economically Disadvantaged Status	0.455
English Language Learner	0.383
IEP Status	0.574

Figure 11: Correlation coefficients for demographic subgroups' **math** assessments.

	Correlation Coefficient
Chronically Absent Status	0.634
Economically Disadvantaged Status	0.684
English Language Learner	0.551
IEP Status	0.779

Figure 12: Correlation coefficients for demographic subgroups' **reading** assessments.

Correlation Analysis: School Building

Across all school buildings, correlations between Star and PSSA were either *moderate* or *high*. There were no drastic differences between buildings.

Middle School	Reading Score Correlation Coefficient	Math Score Correlation Coefficient
Jackson	.699	.669
Lincoln	.789	.712
Martin	.761	.699
Reynolds	.823	.720
Wheatland	.743	.626

Figure 13: Correlation coefficients between assessments for middle school building.

Elementary School	Reading Score Correlation Coefficient	Math Score Correlation Coefficient
Burrowes	.835	.731
Carter & MacRae	.706	.674
Fulton	.754	.653
Hamilton	.691	.628
King	.679	.603
Lafayette	.741	.621

Martin	.686	.669
Price	.701	.661
Ross	.729	.619
Smith Wade-El	.797	.678
Washington	.732	.703
Wharton	.777	.633
Wickersham	.718	.726

Figure 14: Correlation coefficients between assessments for elementary school building.

Correlation Analysis: Grade

Correlation coefficients for reading assessments were generally higher than correlation coefficients for math assessments across grade levels. For all grades, there were *strong* correlations between reading Star and PSSA, and correlations between math Star and PSSA were either *strong* or *moderate*.

Grade	Reading Score Correlation Coefficient	Math Score Correlation Coefficient
Third	.821	.775
Fourth	.796	.766
Fifth	.807	.728
Sixth	.767	.700
Seventh	.837	.682
Eighth	.779	.686

Figure 15: Correlation coefficients between assessments across grade level.

Limitations

This study and its results are limited in some respects. These limitations should be taken into account when applying these results in practice.

- 2023 was the first year of Star implementation. This data is likely to change and adjust as teachers and administrators become more comfortable with the Star assessment.
- This cohort was affected by the COVID-19 pandemic. It is likely that their data will look different than groups in the future who were unaffected by the pandemic.
- This report does not account for causation. This analysis highlights the different predictive capabilities of Star across school buildings and demographic groups, but does not provide specific reasons as to why these differences may have occurred.
- This analysis used the most recent Star score prior to PSSA testing that was available for each student. Therefore, it is possible that Star scores in this sample were pulled from multiple different months or years. This could have an effect on the Star score's relationship to PSSA score.
- The multivariate analysis does not explain all the variance present in the PSSA scores. It is likely that there are other variables at play that have not been taken into consideration. These may include but are not limited to student behavior, grades, and community or social factors.

Application and Recommendations

The purpose of this report is to provide data and observations on the relationship between the Star assessment and PSSA so that teachers and administrators can improve their understanding of student achievement and alter their expectations for students. There are a few key takeaways and possible ways that the results of this study can be applied in the future.

- Most Star benchmark levels only accurately predict PSSA benchmark levels about half the time. Teachers should take this into account and adjust instruction accordingly. Separating students or tailoring instruction based on Star benchmark levels may not help a child succeed on the PSSA.
- In the future, the cut-offs for these benchmark levels should be evaluated. It would be valuable to consider the amount of students who may be on the cusp of a different level. This may offer insight as to why the Star and PSSA levels do not match for many students. Star district benchmarks could be adjusted accordingly.
- Reading Star can be trusted as a good predictor of ELA PSSA and when students have **high** scores on Math Star, it is a good predictor of Math PSSA. However, when students have **low** scores on Math Star, this assessment is *not* reliable as a predictor of PSSA.
- Teachers should be careful as Star math scores are not directly correlated to PSSA outcomes. Students struggling on the math Star should be evaluated separately from students who score higher on the math Star. More specifically, for students in a

demographic subgroup, particularly English language learners or chronically absent students, the math Star and PSSA have a *weak* correlation. Students falling into these groups should be monitored separately and their math Star score should not be trusted to predict their PSSA score.

Appendix

	PSSA Math Level: Advanced	PSSA Math Level: Not Advanced
STAR Math Level: Advanced	38.8%	61.2%
STAR Math Level: Not Advanced	1.0%	99.0%

Figures 16: Percent of math Star Advanced level group that scored Advanced PSSA level or not.

	PSSA Math Level: Proficient	PSSA Math Level: Not Proficient
STAR Math Level: Proficient	43.7%	56.3%
STAR Math Level: Not Proficient	9.0%	91.0%

Figures 17: Percent of math Star Proficient level group that scored Proficient PSSA level or not.

	PSSA Math Level: Basic	PSSA Math Level: Not Basic
STAR Math Level: Basic	53.1%	46.9%
STAR Math Level: Not Basic	19.1%	80.9%

Figures 18: Percent of math Star Basic level group that scored Basic PSSA level or not.

	PSSA Math Level: Below Basic	PSSA Math Level: Not Below Basic
STAR Math Level: Below Basic	83.8%	16.2%
STAR Math Level: Not Below Basic	20.5%	79.5%

Figures 19: Percent of math Star Below Basic level group that scored Below Basic PSSA level or not.

	PSSA ELA Level: Advanced	PSSA ELA Level: Not Advanced
STAR ELA Level: Advanced	45.7%	54.3%
STAR ELA Level: Not Advanced	2.3%	97.7%

Figures 20: Percent of reading Star Advanced level group that scored Advanced PSSA level or not.

	PSSA ELA Level: Proficient	PSSA ELA Level: Not Advanced
STAR ELA Level: Proficient	59.3%	40.7%
STAR ELA Level: Not Proficient	15.8%	84.2%

Figures 21: Percent of reading Star Proficient level group that scored Proficient PSSA level or not.

	PSSA ELA Level: Basic	PSSA ELA Level: Not Basic
STAR ELA Level: Basic	64.8%	35.2%
STAR ELA Level: Not Basic	37.0%	63.0%

Figures 22: Percent of reading Star Basic level group that scored Basic PSSA level or not.

	PSSA ELA Level: Below Basic	PSSA ELA Level: Not Below Basic
STAR ELA Level: Below Basic	49.5%	50.5%
STAR ELA Level: Not Below Basic	6.8%	93.2%

Figures 23: Percent of reading Star Below Basic level group that scored Below Basic PSSA level or not.

References

Renaissance Learning. (2022). *Star Assessments™ for Math Technical Manual*.
https://renaissance.widen.net/view/pdf/kxqxbuhbef/SMRPTechnicalManual.pdf?t.download=true&u=zceria

Akoglu H. (2018). User's guide to correlation coefficients. *Turkish journal of emergency medicine*, *18*(3), 91–93. https://doi.org/10.1016/j.tjem.2018.08.001