Charlotte Nelson, MD, FACP

ACP/AAPL Capstone Project Summary

An analysis of lab and imaging utilization by our hospitalist group over four quarters

The St Patrick Hospital Hospitalist Program is a 26 provider group working in a mid-sized tertiary hospital. It was my sense that there was significant variability within the group, in terms of lab and imaging utilization. With the increasing focus on high value care, I felt that it was important to better understand our group's ancillary utilization patterns. If wide variability were demonstrated, we might be able to influence practice habits. Additionally, assuming there was demonstrable variability, I wanted to know whether such variability had impact upon or correlated with certain measurables, such as mortality and length of stay.

The literature discusses "appropriate utilization" of laboratory testing in particular, although it is not clear that there are any established comparative benchmarks. Most articles which I encountered focused on the related larger issue of total cost of care and trying to influence practice patterns by helping providers understand the costs of the tests they are ordering.

In order to focus my review, I queried our lab director and lab manager regarding high cost tests, the ordering of which they were concerned about, but also our overall volume of tests. As an example, in parallel with this project, the lab had seen and was hoping to better understand a recent spike in the ordering of Hemoglobin A1c tests. Additionally, in the prior year or so, our group had been counseled regularly to avoid the inpatient ordering of very costly, unreimbursed and low value-added genetic panels. The lab provided me with a letter to the editor from the Archives of Pathology which described an educational campaign around test costs aimed particularly at resident physicians which significantly reduced that groups laboratory spending.

Methods:

I looked at data from Quarter 3 (Q3) and 4 of 2016 and Q1 and 2 of 2017. Lab data was provided to me by a clinical informaticist working within the lab. The information came in the form of a data base organized into a pivot table and searchable by provider. I was reliant upon the accuracy of the data provided. For example, I identified several errors in Q4 2017 data which were ultimately rectified by the informaticist. Ultimately, I had to trust that the final iteration given to me was accurate.

Imaging data was collected and then manually organized into a large database by the radiology manager. This database included all ordering providers at our facility, requiring me to create a pivot table in order to focus in only on the hospitalist group.

I was able to use our scheduling software (Qgenda) to document worked shifts by provider over each quarter. I combined our swing shift and night shifts into one "shift type," as these are both admitting shifts, versus our day shift, which is primarily rounding and discharge responsibilities. Our EHR does not track patient encounters, so as a surrogate I chose the number of charges dropped per provider per quarter. This necessity has limitations in that there could be more than one charge per provider for a given patient on a given day, but this is a rare enough occurrence to be felt negligible.

Mortality and length of stay data was provided to me by our Quality Department. This data is pulled from Vantage software. Historically, attribution difficulties have made analyses difficult, particularly for hospitalists - that is shared or team-based- care. Under current definitions, length of stay and mortality is attributed to the discharging hospitalist who may, in some instances, not have been the provider delivering the bulk of a particular patients care, thus raising questions about the validity of the attribution. It is hoped that this difficulty is somewhat "washed out" by capturing data for four quarters.

Finally, I analyzed our group in three cohorts.

1) General hospitalists (18-19 depending on the quarter) – rotate through all teams (other than ID).
2) Advanced practice providers (5) – two split their time between a daily rounding team and an evening cross cover team, one does only evening cross cover, and two work only days.
3) ID/IM Hospitalists (3)– work only daytime rounding and clinic shifts.

Assumptions/Limitations:

-For the group of 19 providers, it is assumed that over the course of 4 quarters the case mix index and number of floor vs ICU patients will balance out.
- Shifts are not necessarily distributed evenly – one provider works predominantly nights, a couple work no nights. I do not have data down to the level of what labs were ordered on what shift.
-"Daily labs" includes CBC, BMP, and CMP. However, there are instances when these tests are ordered outside of a routine daily lab situation.
-Breaking down radiology studies per encounter seemed less helpful, as not all patients will have radiologic studies. It seemed more appropriate to simply look at provider-based patterns of utilization .
-We have one team which functions as a teaching service, with the majority of the testing volume ordered by resident physicians. As such, the attendings will have lower testing volumes when on that service.
- We have one team which is primarily focused on cardiology patients. This could skew the # of echocardiograms and troponins ordered.

Results:

Please see the attached spreadsheets. The majority of this analysis is directed at the providers who rotate through each of our teams.

The lowest utilizers of laboratory services remained consistent across all four quarters examined, with the exception of one provider from Q1 2017, who is typically one of our higher utilizers. This provider worked a majority of daytime rounding shifts (22 of her 24 shifts) during that outlier quarter, the likely explanation for this aberrancy.

There were also some consistencies among the 3-4 highest utilizers each quarter. One provider was in the high category 3 of 4 quarters; he joined our group in Q4 part time and Q1 2017 full time. A long time (5 year) provider with us also appeared in the top utilizer group 3 of 4 quarters.

Length of stay was markedly lower for low vs high lab utilizers across all four quarters, with only one exception (Seagraves, Q2 2017). There is a less consistent but similar trend for mortality data.

Q3 2016 High Utilizers

Provider	LOS	Mortality (O/E)
Floyd	3.69	1.62
Neathawk	3.99	0
Osuchowski	3.32	0

Q3 2016 Low Utilizers

Provider	LOS	Mortality (O/E)
Curtis	2.54	0.93
Caramore	3.16	1.15
Paddock	2.82	0

Q4 2016 High Utilizers

Provider	LOS	Mortality (O/E)
Adams	3.14	0.36
Floyd	3.10	0
Watson	1	0
Osuchowski	3.79	0.26

Q4 2016 Low Utilizers

Provider	LOS	Mortality (O/E)
Seagraves	2.89	1.05
Paddock	2.47	1.21
Curtis	2.94	0.50
Caramore	2.74	0.58

Q1 2017 High Utilizers

Provider	LOS	Mortality (O/E)
Watson	5.09	1.28
Neathawk	3.69	0.80
Floyd	4.07	0.33

Q1 2017 Low Utilizers

Provider	LOS	Mortality (O/E)
Welch	2.85	0
Seagraves	3.21	1.02
Curtis	3.29	0.99

Q2 2017 High Utilizers

Provider	LOS	Mortality (O/E)
Watson	2.58	0
Welch	3.75	1.49
Knight	3.17	1.56

Q2 2017 Low Utilizers

Provider	LOS	Mortality (O/E)
Seagraves	3.69	0.27
Curtis	2.70	0.58
Paddock	2.67	2.95
Caramore	2.81	0.46

Analysis of radiology utilization patterns was far less straightforward. I chose to look at the number of CT scans (excluding CT guided procedures), MRI's, echocardiograms (complete and limited; transesophageal echocardiograms were excluded), and nuclear medicine studies. I did not look at ordering patterns for plain radiographs. I also opted not to tally the total number of radiology studies.

The low volume of nuclear medicine studies, true across all quarters and providers, rendered analysis of this particular type of test valueless.

Echocardiogram ordering showed great variability from quarter to quarter, and analysis of this was complicated by the probable loading of echocardiogram ordering toward our single cardiology-heavy team. This was a variable which could not be controlled for nor readily broken out.

Higher use of CT scans appeared to parallel the higher lab utilization patterns among the same providers. The volume of MRI's ordered was much lower than that of CT's.   There was one provider who was amongst the highest for all four quarters in terms of MRI ordering. A second provider was also consistently among the top three in three of four quarters. Reinforcing the observation that lab and imaging utilization patterns run in parallel, our providers ordering the fewest CT studies were the same providers who had the lowest lab utilization histories.

I ran a list of "expensive tests". These included antiphospholipid antibody panel, celiac panel, hyperacoagulable panel, paraneoplastic panel, SPEP, and respiratory pathogens panel. Over a given quarter, individual providers never ordered any of these more than once or twice.   As such, there was both insufficient data for analysis and little reason for concern about overuse of high-cost laboratory tests. As stated earlier, we had been receiving regular updates from the lab director regarding the reimbursement challenges and low added-value to many of such molecular tests, and it is probable that these educational efforts influenced ordering habits for such studies during the course of this project.

Our midlevel partners are noted to order considerably fewer tests than their physician counterparts, regardless of the shifts they are working.   The only exception to this was with Mayer in Q2 2017, during which time her total tests/encounter were quite high, but this was during a very low-encounter quarter for her.

For most of my analysis, Cunnington's data has been excluded. This PA had very few "encounters" – ie, his shift type, which is largely cross-coverage, left him with very few charges.

Our infectious disease hospitalists consistently ordered significantly more tests than did the general hospitalists, thought to be largely accounted for by their higher number of encounters. When looking at the figure of tracked tests/encounter, however, they really are not much higher than the rest of the group, and this difference is likely accounted for by the specialized and reference testing which the specialized nature of their practices required. Among the three of them, two demonstrated very similar patterns, while the third was a bit of an outlier with

significant more testing done. Presentation of this data to them may present an opportunity to influence some change.

Discussion:

Gaining a better understanding of our own utilization patterns and practices has been a topic of interest for our group for some time. This project gave me the opportunity to learn not only where to obtain this data, but how to analyze it in meaningful ways. The need for manual report writing and running, true even with a sophisticated EHR, was a substantial obstacle to projects such as mine. Attribution continues to be a challenge for hospitalists, particularly when looking at quality markers such as length of stay and mortality. That said, I did identify trends which will give us opportunities for education and improvement.

Firstly, among our providers ordering the fewest tests, I did not demonstrate any adverse consequences in terms of length of stay or mortality. Our higher utilizers tended to be a consistent and smaller group than the low-utilizers, and there was no apparent corollary mortality benefit among them. Furthermore, LOS was higher among the higher lab utilizers. These observations offer opportunities for all manner of speculation about the subject and influence of practice styles and philosophies, goals of hospitalization, diagnostic confidence and experience, utility of and reliance or over-reliance upon ancillary data, in the inpatient setting.

 I think that presenting data such as mine, perhaps even more powerfully coupled with differences in costs between utilization groups, could be very instructive and behavior-changing. As has been suggested in the literature, at least among resident physicians, the sharing of cost information with ordering providers can influence practice habits. Going forward, my group plans to not only run and evaluate these reports regularly, but track individual and group utilization patterns over time.

My group was aware that I was working on this project; might this have influenced ordering habits even before the information was presented? This is a potential limitation.

I would like to be able to incorporate cost data into my analysis in a more precise way. In a future state, we could have cost/case data for individual providers, by diagnosis, and on the group level.

 With regard to the tests which I tracked, there are clear opportunities for cost savings. For example, the provider who ordered 62 procalcitonins in Q2 2017 incurred a cost of $1814 (excluding phlebotomy costs). The next highest provider ordered 41 tests, at a cost of just under $1200. The average number of procalcitonins ordered, that same quarter, by 8 low utilizers averaged 5.5 tests per quarter and a cost of $161, a greater than 10 fold difference. Actual billed costs/pricing to patients and payors is, of course, far higher.   The magnitude of these costs differences over many providers and many tests in the course of a year is huge.

In trying to find a figure that I could compare and which was normalized for shifts, I looked at the "daily labs/day shift" figure.   The cost of a CBC, BMP, and CMP is all very close to $6.25 each. Therefore, looking again at Q2 2017, the provider who had the least number of daily labs/shift had an average of 6.30 tests for a cost of $39.38 per day. For the provider who ordered the greatest number of daily labs/day shift at 23.62/day, the cost was $147.63. This sizeable cost differential, ostensibly between otherwise similar patients, could influence provider behavior in future.

There is certainly much more information which can be gleaned from further iterations and dissection of the data gathered. Now that the reports have been built, I will be able to access this information in an ongoing fashion.   The data will be presented to my group and, as already referenced, I plan to do a follow up analysis after 2 quarters to look for any significant changes in ordering habits.

It is my hope that my project and such data will give my group opportunities for discussion, standardization and improvement, while making the case for meaningful and more easily acquired clinical cost and utilization data. Given the continued emphasis on high value care, we will all be forced to give such data more scrutiny.

References:

Arch Pathol Lab Med – Vol 138, March 2014

Variation in Physician Spending and Association with Patient Outcomes. JAMA Intern Med doi: 10.1001/jamainternmed.2017.0059

Physician Spending and Patient Outcomes (letter to the editor). JAMA online March 2017

Internal Medicine Resident Engagement with a Laboratory Utilization Dashboard: Mixed Methods Study. Journal of Hospital Medicine Vol 12:9 743-746.