**Supplement:** Survey answer explanations. Data presentation format is also shown below (Supplemental Table 1).

1. Question: Which hospital uses the most central lines?

Concept category: Unadjusted Data

Rationale: Number of central-line days is the measure of central line usage. Therefore, the respondent had to go to column 3 and find the hospital with the highest number.

Correct Answer: hospital F (90% correct)

2. Question: If hospital G’s number of actual infections doubled, what would its CLABSI rate be?

Concept category: Unadjusted Data

Rationale: Actual infections and CLABSI rate are related in that CLABSI rate = (actual infections) / (central-line use). Therefore, if the number of actual infections is doubled, the CLABSI rate would subsequently be doubled.

Correct Answer: 3.58. Any answer between 3.50 and 3.61 was accepted as a correct response (92% correct).

3. Question: Which hospital has the lowest CLABSI rate?

Concept category: Unadjusted Data

Rationale: The fourth column provided the raw CLABSI rates for each hospital. The respondent needed to go to this column and find the hospital with the lowest number.

Correct Answer: hospital D (90% correct)

4. Question: If hospital B had its number of projected infections halved, what is its SIR?

Concept category: Risk-Adjusted Data

Rationale: SIR, or standardized infection ratio, is the ratio of actual infections to projected infections. Therefore, if a hospital has its number of projected infections halved (the denominator of SIR), this will double the SIR.

Correct Answer: 2.4. Any answer between 2.0 and 3.0 was accepted as a correct response (70% correct).

5. Question: If hospital A doubled its central-line use but other practice patterns remained the same, how many actual infections would hospital A expect to have?

Concept category: Unadjusted Data

Rationale: Practice patterns determine the rate of CLABSI; if there is no change in practice patterns, there should be no change in rate regardless of the number of patients to whom those practices apply.

Central-line use, actual infections, and CLABSI rate are related in that CLABSI rate = (actual infections) / (central-line use). The question doubled the denominator of the rate, but held the rate constant, meaning that the numerator must also double. That is, the actual number of infections would need to double in order to maintain the same CLABSI rate.

Correct Answer: 14. Any answer between 13 and 15 was accepted as a correct response (73% correct).

6. Question: Suppose hospital A begins using a central line with an antibiotic coating that halves infections. What would hospital A's number of projected infections be?

Concept category: Risk-Adjusted Data

Rationale: Practice patterns do not impact risk-adjustment. The number of projected infections for each hospital is a risk-adjusted statistic that does not take practice patterns into account. Therefore, an alteration in prevention practices will result in the same number of projected infections. This required the respondent to look in column 5 for Hospital A and answer that same number, unchanged.

Correct Answer: 3.7 (38% correct)

7. Question: Which hospital is most effective at preventing CLABSI?

Concept category: Risk-Adjusted Data

Rationale: The most effective hospitals at preventing CLABSI are those with the lowest SIR, which is the risk-adjusted CLABSI statistic. This involved identifying the hospital with the lowest number in column 6.

Correct Answer: hospital C (85% correct).

8. Question: The presence of a gastrostomy (g) tube is a risk factor for CLABSI. If this variable is not accounted for in CLABSI reporting, how would this impact the interpretation of the number of infections projected by national experience? (**Check all that apply**)

Concept category: Risk-Adjusted Data

Rationale: The number of infections projected by national experience is a risk-adjusted statistic that takes into account known factors to impact the development of CLABSI. Unaccounted risk factors lead those projections to be wrong for an individual hospital. If an unaccounted risk factor is abundant, then more infection will happen at that hospital, and the number of projected infections will be too low. If an unaccounted risk factor is rare, then fewer infections will happen at that hospital, the number of projected infections will be too high.

Correct Answer: Underestimation in a hospital with many g-tubes, overestimation in a hospital with few g-tubes. Correct answer defined as selection of either response or selecting both responses (69% correct).

9. Question: Which hospital’s patients are the most predisposed to developing CLABSI?

Concept category: Risk-Adjusted Data

Rationale: Patients are most predisposed to developing a condition when they are at the highest risk for the condition irrespective of practice patterns. The number of projected infections by national experience is the data that indicate the predisposition for each hospital’s patients to develop CLABSI. Thus, the hospital with the highest number of projected infections (the highest number in column 5) is the answer.

Correct Answer: hospital F (83% correct)

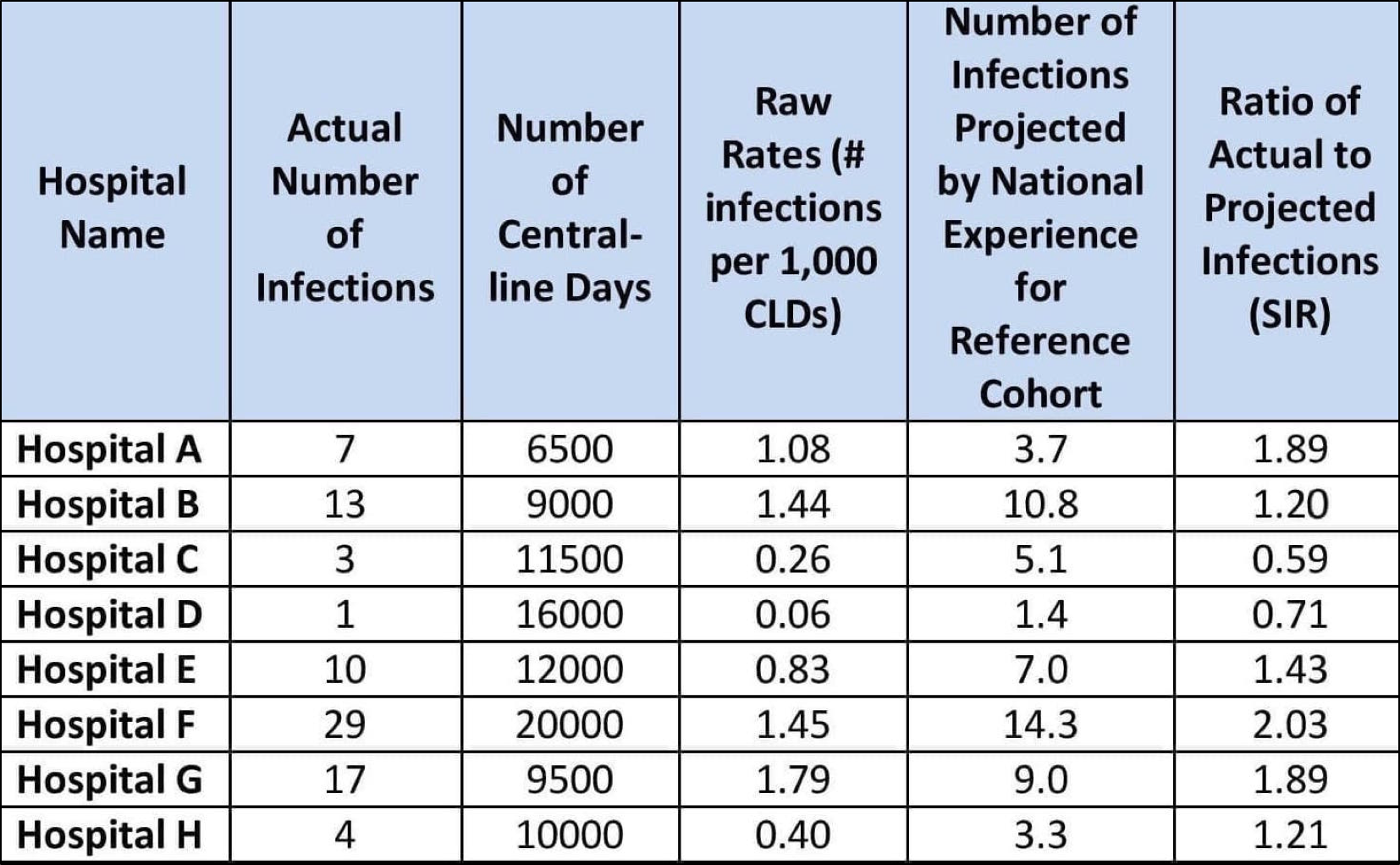
10. Question: Suppose hospitals A and H have the exact same CLABSI prevention practices. Which hospital will have the higher number of CLABSI?

Concept category: Risk-Adjusted Data

Rationale: Practice patterns do not impact risk-adjustment. Therefore, in the case where two hospitals have the same prevention practices, the hospital with patients who are most predisposed to develop CLABSI will have the higher rate. Thus, the higher number of projected infections will result in the higher number of CLABSI – Hospital As “3.7” in column 5 is greater than Hospital H’s “3.3”.

Correct Answer: hospital A (48% correct)

**Supplemental Table 1:** Data Presentation Format



**Supplemental Table 2: Expert Answers to Question, “In your opinion, what are the three biggest problems for reliability of quality metric data at your hospital?"**

|  |  |  |  |
| --- | --- | --- | --- |
| **Respondent** | **Response 1** | **Response 2** | **Response 3** |
| A | Estimation of risk of population | Lack of accounting for socioeconomic and health literacy | CAUTI criteria are not reliable, when fever is related to other process. |
| B | risk stratification | inclusion of MBI-LCBI | counting the presence of more than 1 line as 1 line |
| C | Predicted infections based on gamed national data | Little to no validation of benchmark data |  |
| D | training of data collector | data available in files |  |
| E | High number of chronic ECF admits with ventilators/G tubes | Lack of understanding of random variation from the mean | Lack of understanding of risk adjustment |
| F | |  | | --- | | miscounting line days | | MBI confounding |  |
| G | |  | | --- | | reliability of CLABSI diagnosis | | risk adjustment | validity of CLABSI diagnosis |
| H | |  | | --- | | inadequate risk adjustment | | inability to account for other medical conditions (ie, abdominal trauma) that can lead to bacteremia in pts with central lines |  |
| I | |  | | --- | | lack of risk adjustment for comorbid conditions | | lack of objectivity/reproducibility of measure | poor choice of measures |
| J | |  | | --- | | Interface between MedMined and NHSH issues | | EPIC Denominator data integrity issues | CLABSI Definitions are flawed, & NHSN appeals are fruitless with the non-clinical staff behind computers in Atlanta |
| K | |  | | --- | | difficulty with applying surveillance definitions | | incomplete medical record to appy definitions | inter-rater reliability |
| L | |  | | --- | | not all pts are the same | | some preventable infections are more preventable than others | over use of devices can give lower rates, so device use data should be included |
| M | |  | | --- | | less human resource | | less budget | data collections |
| N | |  | | --- | | Antibiotics often started before blood cultures taken | | Staff issues | No national projected value for risk adjustment |
| O | |  | | --- | | errors in counting central line days | | disagreement on CLABSI cases | errors due to downloading automated data |
| P | inter-rater reliability |  |  |
| Q | |  | | --- | | Blood draw technique | | NHSN contaminant list not updated | Not enough site specific infection criteria in NHSN include blood culture as a criterion of infection |
| R | I don't think we have 3 big problems with our data. I do wonder how brutally honestly other centers apply the criteria, especially in overcalling 2ndary clabsi's when the primary site doesn't meet NHSN criteria for infection |  |  |
| S | Inaccurate denominator data | definition changes over time |  |
| T | Colonization | presence of complicated abdominal sepsis | inadeqaute blood cultures |
| U | |  | | --- | | Presence of multiple site of infection | | Patients with neutropenia | Improper culture technique |
| V | |  | | --- | | automated measures not capturing true preventable infecton - MRSA and c diff lab ID | | surveillance defintion not making clnical sense and difficult to explain to front line staff. | Lack of proper risk adjustment. |
| W | |  | | --- | | Time needed to collect and validate data | | Interpreting data for actionable initiatives | Prioritizing infection among all other competing priorities |
| X | |  | | --- | | reliability of events (i.e. cultures of catheters) | | quality of information on denominator (i.e. dates of exposures) | risk adjustment on main counfounders |
| Y | |  | | --- | | explaining the SIR | | the SIR does not account for short-term vs long-term lines | NHSN should also track CLABSI/pt-day to better capture efforts to reduce line use |
| Z | |  | | --- | | lack of risk adjustment | | surveillance definitions do not match clinical picture |  |
| AA | |  | | --- | | criteria vs patient acuity differences | | staff training and stability | data input and extraction limitations |
| BB | The issue of rate calculation: per 1000 catheter days vs per 1000 patients with central line |  |  |
| CC | |  | | --- | | Surveillance definitions are problematic, esp for CAUTI, C diff | | Risk factor adjustments are limited | MRSA Lab ID events - picks up many infections POA |
| DD | We don’t use quality metric data |  |  |
| EE | Absence of SIR | ambiguity of CLABSI definition rather than CR-BSI |  |
| FF | |  | | --- | | Subjectivity in definitions | | Using line days instead of patient days as denomniator | Not adjusting for culturing practices |
| GG | NHSN definitions |  |  |
| HH | Gaming the system | Gaming the system | Gaming the system |
| II | Detection (over) | Line-days gathering | Other risk factors eg: neonates with intestinal failure |
| JJ | |  | | --- | | Other states do not audit CLABSI data routinely, notably these states have lower CLABSI rates...national expected numbers therefore may be underestimates | | Contaminated Blood draws with 2 of 2 vials positive because of improper collection from same stick | Getting Acuurate Line days (usually underestimated) |
| KK | |  | | --- | | Different interpretations affect measurement | | Not all facilities are "honest" in reporting data | Linking money to quality outcomes skews acuracy of data reporting |
| LL | |  | | --- | | Gaming (unconscious or conscious) | | Definitional issues related to organisms included in the MBI criteria | Definitional issues related to organisms included in the common skin contaminant category |
| MM | |  | | --- | | (Mis)classification of primary vs. secondary BSI | | Limited clinical variables for risk adjustment model | Penalizes teaching hospitals |
| NN | |  | | --- | | different interpretation of CLABSI def | | all risk factors for CLABSI are not accounted for | different patient populations |
| OO | improvements in actual number are not always captured in SIR |  |  |
| PP | Subjectivity in applying case definitions (e.g. presence of an alternative source of infection) | Lack of benchmark data |  |
| QQ | Not risk adjusted | Infection data not validated | Denominator data not validated |
| RR | Surveillance definitions | Information systems i.e. line day tracking | Limited Staffing |
| SS | |  | | --- | | intepretation in the measures that have some variability | | accuracy of information from EMR/data mining software compared with manual counts | intepretation of the metrics from non-epidemiology trained leadership, board, etc |
| TT | Applying definitions | Getting valid denominators | Risk adjustment |
| UU | Data collection process | Compliance to protocol | Low hand hygiene rates |
| VV | |  | | --- | | The NHSN definitions are still not realistic and valid, in my view. | | National adjudication skews the data |  |
| WW | |  | | --- | | Human factors in data gathering/recording | | Appropriately generalizable benchmarks |  |
| XX | |  | | --- | | inadequate adjustment for risk | | reasonable QI measures may or may not affect rates | people are trying to reduce rates, but do not focus on improving processes |
| YY | |  | | --- | | Lack of adequate risk adjustment | | Expectation from external agencies for 0 infections | Surveillance bias between facilities |