

Understanding thoracic surgeons' perceptions of administrative database analyses and guidelines in clinical decision-making



Kimberly A. Shemanski, MD,^a Albert Farias, PhD,^b Dustin Lieu, BS,^a Anthony W. Kim, MD,^a Sean Wightman, MD,^a Scott M. Atay, MD,^a Robert J. Canter, MD, MAS,^c and Elizabeth A. David, MD, MAS^a

ABSTRACT

Objectives: This study explored cardiothoracic surgeons' perceptions of health services research and practice guidelines, particularly how both influence providers' clinical decision-making.

Methods: A trained interviewer conducted open-ended, semistructured phone interviews with cardiothoracic surgeons across the United States. The interviews explored surgeons' experiences with lung cancer treatment and their perceptions of health services research and guidelines. Researchers coded the transcribed interviews using conventional content analysis. Interviews continued until thematic saturation was reached.

Results: The 27 surgeons interviewed mostly were general thoracic surgeons (23/27) who attend tumor board weekly (21/27). Five themes relating to physician perceptions of health services research and guidelines emerged. Databases analyses' inherent selection bias and perceived deficit of pertinent clinical variables made providers skeptical of using these studies as primary decision drivers; however, providers thought that database analyses are useful to supplement other data and drive future research. Likewise, providers generally felt that although guidelines provide a useful framework, they often have difficulty applying guidelines to individual patients. An analysis of provider characteristics revealed that younger physicians in practice for fewer years appeared more likely to report using guidelines, and physicians who were aged 50 years or more and not purely academic surgeons appeared to find database analyses less impactful.

Conclusions: Health services research, including database analyses, comprise much of the surgical literature; however, this study suggests that perceptions of database analyses and guidelines are mixed and questions whether thoracic surgeons routinely use either to inform their decisions. Researchers must address how to present compelling data to influence clinical practice. (*J Thorac Cardiovasc Surg* 2021;161:807-16)

Physicians rely on many resources, including national guidelines and health services research (HSR), to make clinical decisions. HSR is defined by the Agency for

From the ^aDivision of Thoracic Surgery, Department of Surgery, Keck School of Medicine of the University of Southern California, Los Angeles, Calif; ^bDepartment of Preventative Medicine, Keck School of Medicine of the University of Southern California, Los Angeles, Calif; and ^cDepartment of Surgery, University of California, Davis, Sacramento, Calif.

This project was supported by generous donations from a grateful patient designated for research use.

This study was deemed exempt by the University of California Davis's Institutional Review Board.

Read at the 100th Annual Meeting of The American Association for Thoracic Surgery: A Virtual Learning Experience, May 22-23, 2020.



Thoracic surgeons have mixed perceptions of HSR and guidelines.

CENTRAL MESSAGE

Thoracic surgeons have mixed perceptions about HSR and guidelines, and certain attributes affect how likely providers are to use HSR and guidelines in clinical practice.

PERSPECTIVE

HSR comprises much of surgical literature because of the difficulty of studying surgical topics in a randomized, controlled setting. However, this study suggests that perceptions of HSR and guidelines are mixed and questions whether thoracic surgeons routinely use either in practice. Researchers must address how to present compelling data to influence clinical practice.

See Commentaries on pages 817, 818, 819, and 820.

Healthcare Research and Quality as a multidisciplinary field of investigation that studies how social, financial, organizational, technologic, and personal factors interact to

Received for publication April 9, 2020; revisions received July 10, 2020; accepted for publication Aug 22, 2020; available ahead of print Sept 17, 2020.

Address for reprints: Elizabeth A. David, MD, MAS, Division of Thoracic Surgery, Department of Surgery, University of Southern California, 1510 San Pablo St, HCC 1 Suite 514, Los Angeles, CA 90033-4612 (E-mail: Elizabeth.david@med.usc.edu).

0022-5223/\$36.00

Copyright © 2020 by The American Association for Thoracic Surgery

<https://doi.org/10.1016/j.jtcvs.2020.08.112>

Abbreviations and Acronyms

COPD	= chronic obstructive pulmonary disease
HSR	= health services research
NCCN	= National Comprehensive Cancer Network
NSCLC	= non-small cell lung cancer
RCT	= randomized controlled trial
ThORN	= Thoracic Surgery Outcomes Research Network



Scanning this QR code will take you to the table of contents to access supplementary information. To view the AATS Annual Meeting Webcast, see the URL next to the webcast thumbnail.



affect healthcare delivery.¹ HSR includes many research designs, including retrospective database analyses, which comprise a large portion of surgical literature due to the ethical and logistical difficulties of conducting randomized controlled trials (RCTs) for surgical treatments.

Physicians' adherence to guidelines and dependence on HSR remains largely unknown; however, studies suggest that rates of guideline-concordant care may be lower than expected. Stokes and colleagues² reported that although patients with stage IA and IB non-small cell lung cancer (NSCLC) received high rates of guideline-concordant care (97.4% and 97.9%, respectively), patients with higher stages of disease received progressively lower rates of guideline-concordant care (stage IIA 49.2%, stage IIB 47.5%, and stage IIIA 32.2%).² Given that a survival benefit was seen for patients receiving guideline-concordant care (5-year overall survival 74.6% vs 55.0%; log rank; $P < .001$), it is vitally important to understand physicians' reasons for straying from the guidelines. The article by Stokes and colleagues² was a large, retrospective database analysis using the National Cancer Database, and although many articles are published yearly using database data, it is unclear how these publications affect clinical practice. By understanding what physicians consider strengths and weaknesses of these resources, we hope to learn how to disseminate new information in an impactful way.

Our study is the first to explore cardiothoracic surgeons' perceptions of guidelines and HSR using qualitative research techniques to enrich findings.³ Figure 1 provides a graphical summary of this study.

MATERIALS AND METHODS

A trained interviewer conducted 45- to 60-minute, open-ended, semi-structured telephone interviews with cardiothoracic surgeons across the United States. Interviews were recorded and then transcribed verbatim.

The transcripts were de-identified before analysis. Participants were recruited from the Thoracic Surgery Outcomes Research Network (ThORN), which is a multi-institutional cooperative group striving to increase evidence-based understanding of general thoracic surgical diseases through health services and outcomes research. For the purposes of diversifying the participant pool, additional participants were recruited using snowball-sampling techniques as follows: Each interview concluded with a request for participant referrals, specifically for providers who were not general thoracic surgeons in academic practice because the majority of ThORN members fall within those categories.⁴ A research assistant tracked communication with potential participants, and 3 emails were extended to each referral offering them the opportunity to participate. The University of California, Davis Institutional Review Board deemed the study exempt, and all participants gave informed consent. Participants received a \$100 gift card upon completing the interview (Appendix E1 shows interview script).

The interview was divided into 3 sections. The first section explored participants' experiences with treating advanced-stage NSCLC. The second section asked providers to develop treatment plans for hypothetical, case-based clinical scenarios, and the third and final section discussed providers' perceptions of guidelines and HSR. This analysis focused on the third section of the interview.

The interviews were analyzed using inductive reasoning.⁵ Seven researchers independently reviewed each interview transcript and generated codes for recurring content. The researchers then met as a team to analyze the interviews using conventional content analysis, which is defined as a methodology that subjectively interprets the content of text data and systematically classifies the data into themes.⁶ Studies have suggested that this method is suited particularly well for intensely examining language and deriving explicit or inferred meaning to understand the topic of interest deeply.^{5,6} The team created consensus codes for thematic elements related to the perception and use of HSR and guidelines in clinical practice, particularly in treatment decision-making. NVivo 12 software (QSR International, Melbourne, Australia) was used to organize the analysis. Interview participants were recruited until thematic saturation was reached, which was defined as when additional respondents' perspectives no longer introduced novel themes.⁵ A brief demographic survey was administered with the interviews. A secondary analysis was performed using descriptive statistics to explore if any of the themes appeared more frequently or more emphatically in any of the demographic groups.

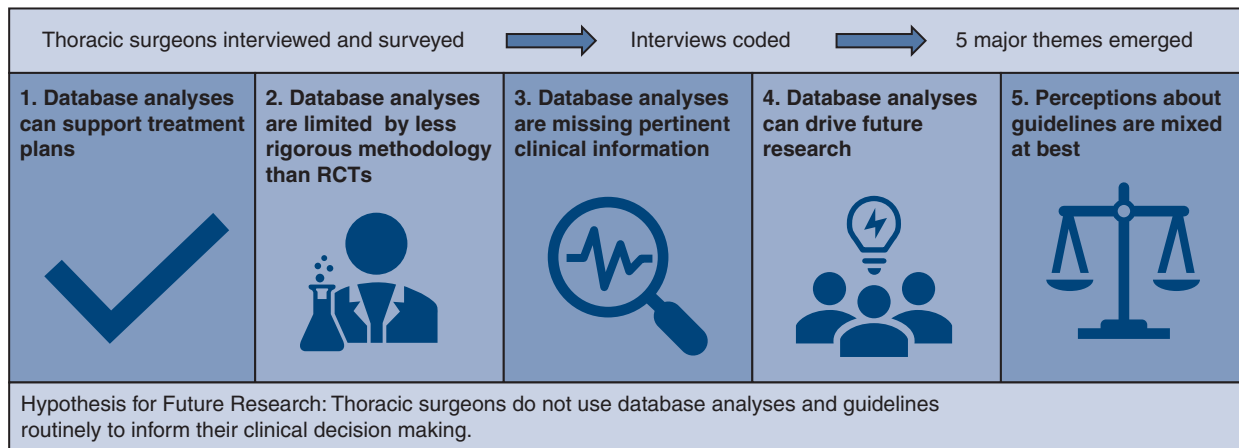
RESULTS

A total of 27 cardiothoracic surgeons were interviewed before reaching thematic saturation. Participant characteristics are summarized in Table 1. The participants mostly were general thoracic surgeons (23/27), aged 50 years or less (19/27), and academic surgeons (18/27), and attended multidisciplinary tumor board weekly (21/27). Time in practice was divided evenly between 10 years or less (13/27) and more than 10 years (14/27). Participants equated HSR with database analyses, and the terms will be used interchangeably here. Five broad themes relating to physician perceptions of database analyses and guidelines emerged from the data and highlighted mixed perceptions of both (Tables 2-6). Participant quotes support each theme.

Theme 1. Database Analyses Are More Useful to Support Treatment Plans Rather Than to Create Them

Many participants thought that although database analyses did not contain all of the relevant information

Understanding Thoracic Surgeons' Perceptions of Administrative Database Analyses and Guidelines



Shemanski, et al. April 2020.



FIGURE 1. Thoracic surgeons across the United States were interviewed regarding their perceptions of guidelines and HSR, namely, retrospective database analyses. A total of 27 surgeons were interviewed before reaching thematic saturation, and the 5 themes developed from participant responses demonstrated that thoracic surgeons have mixed perceptions of database analyses and guidelines. *RCT*, Randomized controlled trial.

necessary to create treatment plans, the studies were useful to supplement their decisions and validate other research (Table 2). One provider explained that although database studies may not act as “primary drivers,” “You can use information from [database] studies to help inform the decision.” Another echoed, “Outcomes data can certainly help drive better understanding of guidelines...or [help clarify] certain parts.” Furthermore, the large populations analyzed in database studies also increase the generalizability of their findings.^{7,8} In this sense, when database studies are “in agreement with [RCTs...they can give the data] extra credibility,” or simply put, database data can “supplement higher level data.” However, at least 1 physician cautioned using database analyses simply to justify a personal opinion rather than to inform decisions objectively. In the provider’s words, it can be “very easy for individual specialties to use the administrative databases to drum up support for their own cause.”

Theme 2. Database Analyses Are Limited by Less Rigorous Methodology Than Randomized Controlled Trials, Which Makes Them Particularly Susceptible to Selection Bias

Participants noted database analyses are not as methodologically rigorous as RCTs, which 1 provider noted are “obviously the gold standard.” Specifically, providers thought that a lack of randomization and the retrospective nature of database analyses predispose them to a problematic degree of selection bias, and that they “may not be the best to compare therapies” (Table 3). One provider emphasized that there is “no way of knowing what went

into the process of sending patients forward for surgery,” and that essentially, the patients “self-selected” themselves for surgery by being healthy enough to tolerate it. The participant continued to explain that in these instances, “You’re comparing apples to oranges.” In this sense, selection bias is closely linked to confounding bias with multiple providers mentioning that when patients “self-select” themselves for surgery it unevenly lumps favorable confounding variables in the surgical cohorts. This sentiment created concern that database studies could not tease out treatment effects from inherently better prognoses.

Confounding variables mentioned by the participants included improved performance status and pulmonary function and favorable responses to systemic and radiation therapies, all likely seen in higher proportions in the surgical cohorts. One provider illustrated this using pulmonary function as an example. This provider explained that patients with non-small cell lung cancer (NSCLC) and chronic obstructive pulmonary disease (COPD) may have severely decreased pulmonary function that prohibits surgical resection. These patients have a “significant risk of being dead in the next 2 to 5 years just from their COPD, irrespective of their cancer.” The surgeon reasoned that comparatively, patients with NSCLC with COPD and acceptable pulmonary function for surgical resection may have improved survival, but “[to attribute survival benefit to surgical treatment in that instance] is completely wrong.” One provider summarized this theme nicely with “The people with stage IV you operate on do better. Well, of course, but we’re only operating on a really small, select group of the very best of that cohort.”

TABLE 1. Characteristics of cardiothoracic surgeon interview participants

Characteristics	n (%)
Participants (n = 27)	
Age (y)	
≤50	19 (70.37)
>50	8 (29.63)
Surgeon affiliation	
Academic	18 (66.67)
Mixed	2 (7.41)
Other	7 (25.93)
Years in practice	
0-10	13 (48.15)
11-20	10 (37.04)
21-30	3 (11.11)
>30	1 (3.70)
% of practice that is general thoracic surgery	
>75%	23 (85.19)
51%-75%	2 (7.41)
25%-50%	0 (0.00)
<25%	2 (7.41)
% of practice that is thoracic oncology	
>75%	2 (7.41)
51%-75%	16 (59.26)
25%-50%	8 (29.63)
<25%	1 (3.70)
% of practice that is advanced stage thoracic oncology	
>25%	5 (18.52)
10%-25%	13 (48.15)
<10%	9 (33.33)
Tumor board attendance	
Weekly	21 (77.78)
Biweekly	4 (14.81)
Monthly	1 (3.70)
Other	1 (3.70)
Cases presented at tumor board	
>25	16 (59.26)
12-25	6 (22.22)
<12	5 (18.52)

A trained interviewer conducted 45- to 60-minute semistructured, open-ended interviews with cardiothoracic surgeons across the United States. A total of 27 cardiothoracic surgeons were interviewed before reaching thematic saturation with most participants aged 50 years or less, academic, general thoracic surgeons who attend tumor board at least once weekly. Years in practice was evenly divided between and less than 10 years and 10 years or more.

Although most participants perceived selection bias negatively, some described how providers could use it when making decisions. For example, 1 surgeon noted that database analyses could be useful if providers use the “cherry-picked selection bias” as inclusion criteria for surgical candidates. In another participant’s words, “Selection bias actually is an important part of understanding whom to choose.” Database analyses’ retrospective nature, lack of randomization, and inherent selection bias are weaknesses that may limit their impact on clinical decision making.

TABLE 2. Theme: Database analyses are more useful to support treatment plans rather than to create them

Interpretation	Examples
Database research does not drive clinical management but can supplement decisions and validate other research.	“I think that there are some questions that can be answered well using administrative databases...but they should not shape guidelines. They shouldn’t shape, necessarily treatment recommendations.”
	“I wouldn’t rely necessarily on an administrative database study to guide that decision, but you can use information from those studies to help inform the decision.”
	“[Database data] can supplement higher level data.”
	“So, each individual one has less weight than a randomized control trial...but done right I think that they certainly can push you in the right direction.”
	“Those are the cases where you start to look around for data to support one approach or another or data to inform decision making. For example, um, offering someone a pneumonectomy after chemoradiation. You know, I wouldn’t rely necessarily on a administrative database study to guide that decision, but you can use information from those studies to help inform the decision.”

The first theme was database analyses are more useful to support treatment plans rather than to create them. Providers thought that database analyses were better suited to supplement other data rather than to act as the primary drivers of their decisions. This theme is supported by quotes from the data.

Theme 3. Databases Do Not Capture All the Pertinent Clinical Information Needed to Make Important Treatment Decisions

Another critique of database analyses included a perceived deficit of relevant information that providers use to make clinical decisions (Table 4). One provider elaborated, “Administrative databases...have great numbers of patients, but are relatively thin on the details and granularity.” Another said “None of these administrative data sites have pulmonary function testing, pulmonary pressures, or cardiac echo studies. They don’t have the very tests that we use to determine [surgical candidacy].” Other surgeons included tumor location, genetic information, and resectability in their lists of missing critical variables.

TABLE 3. Theme: Databases are limited by less rigorous methodology than randomized controlled trials, which makes them particularly susceptible to selection bias

Interpretation	Examples
The retrospective nature of database analyses and their lack of randomization give them inherent selection bias that weakens them in comparison to randomized controlled trials.	“Retrospective studies may not be the best to compare therapies...”
	“Administrative databases probably have very highly selected populations.”
	“We can improve survival by taking the cancer out with surgery, ‘just look at how much better they do than this other group of patients,’ but this other group of patients contains a fairly high percentage of people who are not eligible for surgery. So, you’re not comparing apples to apples, you’re comparing apples to oranges.”
	“You can’t trust these studies because they’re retrospective – selection bias plays a huge role. There is no way of knowing what went into the process of sending patients forward for surgery.”
	“So you’re already selecting out a group of patients that is going to live longer...in part because the patient is healthy enough to be able to tolerate [surgery], right, so, you have a little bit of a skewed patient population.”

The second theme was databases are limited by less rigorous methodology than randomized controlled trials, which makes them particularly susceptible to selection bias. All participants agreed database analyses have selection bias, and most viewed this as a weakness that makes it hard for database studies to tease out treatment effect from inherently better prognoses seen in surgical versus nonsurgical candidates. Although some described how providers could use selection bias to their advantage. This theme is supported by quotes from the data.

Many providers thought that determining resectability was a complex, interactive process that could not be captured by databases. For example, 1 provider referenced a patient who was not a candidate for anatomic resection because of his pulmonary function, but because the lesion was peripheral the provider opted to treat it with a wedge resection rather than stereotactic body radiotherapy. The rationale for this decision was based on this surgeon’s previous experience where tissue damage from radiation was

TABLE 4. Theme: Databases do not capture all the pertinent clinical information needed to make important treatment decisions

Interpretation	Examples
Databases do not collect many of the relevant variables that would make their findings applicable to treating individual patients.	“I mean the administrative databases typically do not provide enough data to figure out why those patients were selected...”
	“...until those administrative datasets actually contain the information that we use to make our decisions, comparing groups with that data, I think has tremendous bias.”
	“None of these administrative data sites have pulmonary function testing. They don’t have pulmonary pressures. They don’t have cardiac echo studies. So they don’t have the very tests that we use to determine if a patient is a candidate for surgery.”
	“We don’t have the variables that we need to use the existing data to answer these questions. We really need to have a national database that includes all of these.”
	“Our current database system doesn’t have the kind of transparency that would be required to really, I think, better understand some of these very specific patient stages.”

The third theme was databases do not capture all the pertinent clinical information needed to make important treatment decisions. Providers thought that databases were particularly thin on the data that we use to determine surgical candidacy, such as pulmonary function testing and tumor location, among other variables. They reasoned that without these data, it is difficult to apply database analyses to their patients. This theme is supported by quotes from the data.

the functional equivalent of a lobectomy. Another expounded on patients that do not “fall nicely into 1 category...but by the books are the same stage.” For example, a database would categorize a patient with stage IIIA NSCLC due to a “couple of cancer cells in 1 level 5 lymph node and a left upper lobe tumor” the same as a patient with a “relatively large tumor and bulky, multi-station mediastinal nodal disease.” These are “clearly different tumors with different prognoses.”

Simply put, one said, “[Database analyses] don’t contain the data that we use to decide who’s a surgical candidate and who is not.” The participant continued, “Until those administrative data sets actually contain the information that we use to make our decisions, comparing groups with that data I think has tremendous bias.”

TABLE 5. Theme: Database analyses can drive future directions in research

Interpretation	Examples
Database research reveals interesting trends in data that require further investigation with more rigorous methodology.	“It also serves as a pilot for a larger project if it’s worth doing you know, you can look at those administrative databases.”
	“Well, it’s good information to lead questions for future research, but it’s hard to use that data as best practice.”
	“The single-institution, retrospective studies and the database-type studies play less of a role in shaping my clinical decision-making. They are more hypothesis-generating in general.”
	I mean, even if [database studies] aren’t good research, it might still just trigger my brain to think about it. You know, I might be like, “Well, it wasn’t really well done, but you know, it’s a good idea to think about.”
	“I think data like that would warrant or help to justify a randomized controlled trial.”

The fourth theme was database analyses can drive future directions in research. Even providers who viewed database analyses cautiously thought that these studies are good at detecting patterns in retrospective data that warrant more rigorous investigation, or in other words they can narrow the focus of more costly and time consuming randomized controlled trials. This theme is supported by quotes from the data.

Theme 4. Database Analyses Can Drive Future Directions in Research

A repeated opinion was that database studies were good at detecting patterns in retrospective data that warranted further investigation with more rigorous methodology (Table 5). In the words of 1 participant, an administrative database study can serve as a “pilot for a larger project if it’s worth doing.” Other providers agreed saying, “[Database analyses] generate some interesting hypotheses” and “[Database data] help justify RCTs.” Even providers who commented on the negative aspects of database analyses, such as a lack of relevant details or selection bias, agreed that these studies could be useful to “lead questions for further research.”

Theme 5. Perceptions About the Usefulness of Guidelines Are Polarized

Providers differed substantially on their overall perceptions of guidelines with their opinions tending to be either entirely positive or negative and rarely moderate (Table 6).

The majority of criticisms comprised 3 categories: Guidelines are outdated; they rarely apply to individual

TABLE 6. Theme: Perceptions about the usefulness of guidelines are polarized

Interpretation	Examples
Physicians have very positive or very negative perceptions of databases but few neutral stances.	Positives: “I’ll rely heavily more on NCCN guidelines that have kind of weeded through all the data...Someone’s already weeding through all those papers that are out there and figuring out the ones that are definitely more believable or more impactful.” “I think NCCN guidelines are lovely.” “I like them. I think they make sense to me, but I also think they do a good job of putting together sort of a dummy-proof [guide].” “A lot of our treatment decisions – we pull up the NCCN guidelines and walk through them with all of our patients.”
	Negatives: “It’s not clear that guidelines are going to matter, because it’s not clear that anyone actually follows them” “I think our guidelines tend to follow behind what thought leaders are able to put forth.” “It’s always hard to find the NCCN guidelines. I don’t know why they’re so hard to find.”

The fifth theme was perceptions about the usefulness of guidelines are polarized. Providers responses differed substantially, and their perceptions were rarely moderate. Most criticisms were in 3 categories: guidelines are outdated; they rarely apply to individual patients; and they are not usable. Support mostly were in 2 categories: Guidelines act as a filter and they organize information. This theme is supported by quotes from the data. NCCN, National Comprehensive Cancer Network.

patients; and they are not usable. In support of guidelines being outdated, 1 participant thought that “Our guidelines tend to follow behind what thought leaders are able to put forth.” The general consensus was that promising treatment advances are reported in the literature and should be considered for patient use before being incorporated into guidelines. When asked for general thoughts on guidelines, 1 participant said, “They are just that. They are guidelines, but they don’t dictate decisions on an individual patient.” Other phrases used to describe how guidelines influence, or more accurately, do not influence treatment decisions were, “We’re in the era of personalized medicine;” “[Treatment is] a case-by-case basis;” “[Treatment is] highly individualized;” and guidelines have a “whole bunch of grey areas.” Last, when providers were probed about the

usability of guidelines, many thought they were cumbersome. One said, “The trouble is they are so comprehensive...[unlike the lung cancer staging card], you can’t put the National Comprehensive Cancer Network (NCCN) guidelines in your pocket.” Another said, “It’s always hard to find the NCCN guidelines.” A third elaborated further, “It’s not clear that guidelines are going to matter because it’s not clear that anyone actually follows them.”

In contrast, some providers had favorable reviews of the guidelines. Support was mostly in 2 categories: Guidelines act as a filter and organize information. In support of guidelines acting as a filter, 1 provider appreciated that with guidelines someone had already “weeded through” all of the available papers and incorporated “the most impactful ones.” Proponents thought that patients should be offered guideline-concordant care because “[Guidelines] are very well thought out and evidence-based.” Unlike the providers who thought that the guidelines were cumbersome, many applauded their ability to organize data concisely, stating that the guidelines made treatment recommendations “dummy-proof.” One provider even called the guidelines “lovely.”

Physician Characteristics

A secondary analysis of the physicians’ characteristics revealed several patterns. Generally, with respect to guidelines, physicians who were younger or in practice fewer years appeared more likely to follow guidelines (Table 7). Expanding on this, all providers who suggested that guidelines routinely influenced their clinical decision making were aged 50 years or less, and the majority of providers aged 50 years or less agreed with this sentiment. The majority of participants in practice 10 years or less had positive perceptions of guidelines, and the majority of participants in practice more than 10 years had mixed or negative perceptions of guidelines. Providers who were not solely academic surgeons (mixed; part of their practice is academic or other) were slightly more likely to follow guidelines.

Generally, with respect to HSR, physicians who were aged more than 50 years or not purely academic surgeons (mixed or other) appeared to rely less on database analyses in their clinical practice (Table 7). Expanding on this, among providers who were aged 50 years or less, there was a fairly even distribution of those who used database analyses to make clinical decisions and those who did not, but the majority of providers aged more than 50 years did not use database analyses in their practice. Likewise, among academic surgeons there was a fairly even distribution of those who used database analyses to make clinical decisions and those who did not, but the majority of providers who practice outside of the academic setting (mixed or other) did not use database analyses in their practice. Additionally, attending tumor board weekly made it slightly less likely that database analyses influenced providers’ clinical decision making.

DISCUSSION

RCTs have long been regarded as the “gold standard” of research design; however, in this study surgeons conceded that many clinical questions are “really hard” to study via RCTs, and “the only way we can learn enough is to just continue to study them in any way that we can.” This has led to using HSR to evaluate treatments. It follows that database analyses then would contribute to guideline formation, but it is not clear that they do. The NCCN ranks the quality of its references on a 5-point scale, with well-designed RCTs earning a 4 or 5 and comprising the bulk of guidelines’ supporting evidence. Well-designed, nonrandomized trials and low-quality, randomized trials each earn a 3 of 5 and are used less frequently.⁹ Despite a role for HSR and the existence of guidelines for almost every major cancer, this study suggests that perceptions of database analyses and guidelines are mixed at best. Database analyses’ inherent selection bias and inability to capture patients’ complete clinical pictures make providers skeptical of using these studies as primary decision drivers; however, providers think that database analyses are useful to supplement other data and to create hypotheses for further study. Likewise, physicians weighed the pros and cons of guidelines and generally think that although they provided a useful framework, they often fell short when applied to individual patients.

The physician perceptions of database analyses and guidelines revealed in this study have real-world applications. First, database analyses can complement data from RCTs to fill in the knowledge gaps created by stringent trial criteria and protocol. Carls and colleagues¹⁰ tapped this potential with their retrospective database analysis that identified patient adherence as the likely reason that real-world effectiveness differed from RCT efficacy for several blood glucose lowering medications.⁸ In this case, results of the RCT allowed providers to choose effective medications, and the results of the database analysis highlighted the need for patient education to optimize the medications.¹⁰ This reflects the theme that database analyses are useful to support treatment decisions and supplement higher-level data.

Second, database analyses can inform RCTs. High-quality RCTs are expensive to conduct and can have a long lag time between inception and results; therefore, it is useful to have data to narrow their focus.⁷ For example, Behera and colleagues¹¹ performed a retrospective database analysis using the National Cancer Database to assess survival outcomes among patients with stage III (N2+) NSCLC who received trimodality therapy with surgery, chemotherapy, and radiotherapy versus chemotherapy and radiotherapy alone. Despite improved survival in the trimodality therapy cohort, their study also found that less than 10% of patients with stage III (N2+) received this.¹¹ This contradiction supports our study’s suggestion that physicians’ may not be relying on database analyses in their

TABLE 7. Representative Responses for Strong Patterns Related to Physician Characteristics

Guideline-related patterns			
Provider age			
<i>Younger physicians appeared more likely to follow guidelines.</i>			
≤50 y	“For a lot of our treatment decisions we pull up the NCCN guidelines and walk through them with all our patients.”	>50 y	“Well, guidelines are guidelines, right? So, they’re meant to be guidelines, and if you have justifications for not following the guidelines, then I think that’s reasonable, right? I do a lot of things that don’t follow the guidelines.”
Provider time in practice			
<i>Physicians who were in practice less years appeared more likely to follow guidelines.</i>			
≤10 y	“Well, 9 of 10 times, I think we’re offering guideline-concordant care.”	>10 y	“They are guidelines, but they don’t dictate decisions on an individual patient.”
	“I think [guidelines] are huge, and I think that most patients should be offered guideline-concordant care. They’re very well thought out, evidence-based guidelines.”		
Health services research-related patterns			
Provider age			
<i>The majority of physicians who were greater than 50 years old appeared to rely little on HSR. Among younger providers, there was a fairly even distribution of those who used HSR to make clinical decisions and those who did not.</i>			
≤50 y	No associated pattern	>50 y	“Well, it’s good information to lead questions for future research, but it’s hard to use that data as best practice.”
			“So, I’m a little bit circumspect on administrative data...I don’t think they guide decision making.”
			“I’m not sure that [retrospective database analyses] help that much.”
Practice setting			
<i>The majority of providers who practice outside of the academic setting (mixed or other) appeared to rely little on HSR. Among academic surgeons, there was a fairly even distribution of those who used HSR to make clinical decisions and those who did not.</i>			
Academic	No associated pattern	Mixed/Other	“We shouldn’t change treatment paradigms based on those retrospective administrative database-based studies.”
			The databases-type studies play less of a role in shaping my clinical decision making.”
			“I think that often times we push the conclusions from those [database] studies a little bit far. They should not shape guidelines.”

A subgroup analysis of the physicians’ characteristics revealed several patterns. Generally, with respect to guidelines, physicians who were younger or in practice fewer years appeared more likely to follow guidelines. With respect to HSR, physicians who were aged more than 50 years or not purely academic surgeons (mixed or other) appeared to rely less on HSR in their clinical practice. *NCCN*, National Comprehensive Cancer Network; *HSR*, health services research.

clinical practice. It presents an opportunity for a focused RCT to revisit possibly outdated results from a 2009 RCT conducted by Albain and colleagues,¹² which suggested no additional overall survival benefit from adding surgical resection to chemotherapy and radiation for patients with stage III (N2+).¹² One provider alluded to this role of database analyses by saying the following:

What needs to continue to happen with NCCN guideline development, administrative database data, and clinical trials is really [acknowledging that] we now have surgical tools that 15 to 20 years ago we didn’t have. These tools allow much lower impact surgical therapy, so maybe it’s

time to revisit where surgery should be involved with [advanced stage] patients.

Interestingly, there were patterns among provider characteristics and perceptions of database analyses and guidelines. We offer several speculative explanations for the reported patterns. Generally, younger physicians have been in practice for fewer years and may have less personal experience to rely on, possibly due to their stricter adherence to guidelines. At least 1 participant supported this notion by reflecting, “I guess it’s possible that the guidelines have shifted, and I’m no longer following them. I don’t get them out with each case anymore. I used to when I was a

fellow and when I first started as an attending.” Similar reasoning could explain why physicians who were aged more than 50 years appeared less likely to rely on database analyses. This pattern was corroborated by a study that reported patients with inflammatory breast cancer were more likely to receive nonguideline-concordant treatment if their physicians graduated from medical school more than 15 years ago.¹³ Second, physicians who are not purely academic surgeons (mixed or other) may rely less on database analyses simply because academic surgeons may have more exposure to all forms of research in their daily practice.

Study Limitations

Limitations of this study include a small sample size; however, interviews were conducted until thematic saturation was reached, so it is unclear how much more insight, if any, more participants would have provided. In fact, Hennink and colleagues¹⁴ studied adequate sample sizes for qualitative research and found that studies typically reached code saturation after 9 interviews and meaning saturation after 16 to 24 interviews. They defined code saturation as the identification of the full range of thematic issues and meaning saturation as the development of a deep understanding of the issues.¹⁴ Additionally, despite attempts to diversify our cohort with targeted recruitment emails and snowball sampling, the majority of participants were aged 50 years or less and academic, general thoracic surgeons who attend tumor board weekly. Years in clinical practice was the only significant variable that was divided nearly evenly between participants. Future research should further investigate the patterns found in this study using appropriately matched groups. A predominance of academic, general thoracic surgeons likely reflects primarily recruiting from ThORN. Given the organization’s mission of increasing evidence-based understanding of health services and outcomes research, one would expect selecting participants from this group to introduce a favorable bias toward HSR; however, the themes uncovered in this study suggest provider ambivalence toward database analyses making selection bias less problematic. In fact, ambivalence toward database analyses among a group of highly engaged researchers who focus on HSR may even strengthen our findings.

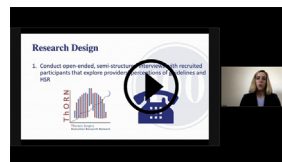
CONCLUSIONS

Qualitative research serves to create hypotheses rather than to test them, and the patterns and themes revealed in this study have generated several hypotheses that require further consideration.⁵ Notably, this study raises the question of whether thoracic surgeons routinely use database analyses and guidelines to inform their clinical decision making. The findings in this qualitative analysis will allow us to create a focused survey that quantitatively explores this hypothesis. Ultimately, future research must investigate

how best to present compelling, new data to influence clinical practice.

Webcast

You can watch a Webcast of this AATS meeting presentation by going to: <https://aats.blob.core.windows.net/media/20AM/Presentations/Understanding%20Thoracic%20Surgeons%20Perc.mp4>.



Conflict of Interest Statement

The authors reported no conflicts of interest.

The *Journal* policy requires editors and reviewers to disclose conflicts of interest and to decline handling or reviewing manuscripts for which they may have a conflict of interest. The editors and reviewers of this article have no conflicts of interest.

The authors thank Qualitative Health Research Consultants and Ruibei Li, BA.

References

1. American Institute of Research and Quality. An organizational guide to building health services research capacity. *Quality*. 2011;1-41.
2. Stokes SM, Massarweh NN, Stringham JR, Varghese TK. Clinical-pathologic correlation and guideline concordance in resectable non-small cell lung cancer. *Ann Thorac Surg*. 2019;108:837-44.
3. Bressan V, Bagnasco A, Aleo G, Timmins F, Barisone M, Bianchi M, et al. Mixed-methods research in nursing – a critical review. *J Clin Nurs*. 2017;26:2878-90.
4. Kirchner J, Charles K. Enhancing the sample diversity of snowball samples: Recommendations from a research project on anti-dam movements in Southeast Asia. *PLoS One*. 2018;13:1-17.
5. Malterud K. Qualitative research (en medicina): standards, challenges, and guidelines. *Qual Res Ser*. 2001;358(panel 2):483-8.
6. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res*. 2005;15:1277-88.
7. Camm AJ, Fox KAA. Strengths and weaknesses of ‘real-world’ studies involving non-vitamin K antagonist oral anticoagulants. *Open Heart*. 2018;5:e000788.
8. Blonde L, Khunti K, Harris SB, Meizinger C, Skolnik NS. Interpretation and impact of real-world clinical data for the practicing clinician. *Adv Ther*. 2018;35:1763-74.
9. National Comprehensive Cancer Network. *Non-small cell lung cancer Evidence Blocks*. Plymouth Meeting, PA: NCCN; 2020.
10. Carls GS, Tuttle E, Tan RD, Huynh J, Yee J, Edelman SV, et al. Understanding the gap between efficacy in randomized controlled trials and effectiveness in real-world use of GLP-1 RA and DPP-4 therapies in patients with type 2 diabetes. *Diabetes Care*. 2017;40:1469-78.
11. Behera M, Steuer CE, Liu Y, Fernandez F, Fu C, Higgins KA, et al. Trimodality therapy in the treatment of stage III N2-positive non-small cell lung cancer: a national cancer database analysis. *Oncologist*. 2020;25:e964-75.
12. Albain KS, Swann RS, Rusch VW, Turrisi AT, Shepherd FA, Smith C, et al. Radiotherapy plus chemotherapy with or without surgical resection for stage III non-small-cell lung cancer: a phase III randomised controlled trial. *Lancet*. 2009;374:379-86.
13. Denu RA, Hampton JM, Currey A, Anderson RT, Cress RD, Fleming ST, et al. Influence of patient, physician, and hospital characteristics on the receipt of

guideline-concordant care for inflammatory breast cancer. *Cancer Epidemiol.* 2016;40:7-14.

14. Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: how many interviews are enough? *Qual Health Res.* 2017;27:591-608.

Key Words: qualitative research, guidelines, retrospective database analyses, non-small cell lung cancer

Discussion

Presenter: Dr Kimberly Shemanski



Dr Virginia R. Litle (*Boston, Mass.*). I am excited that you are bringing a novel area of presentation to the sessions because usually it's all just a lot of clinical work. Thank you for bringing something new that will be applicable and of interest to all subspecialties.

The purpose of clinical guidelines as you know is to provide a roadmap to safety and assist clinicians in navigating a potentially tortuous or changing road. Guideline development has evolved beyond just a small panel of experts without the use of a systematic review as it was 40 years ago to use research evidence that resulted in multiple different approaches. Approximately 20 years ago, to reduce bias, came the application of the GRADE system, which is grading a recommendation, assessment, development, evaluation. The Institute of Medicine provided these criteria for trustworthy guidelines.

There are different kinds of guidelines for cardiothoracic surgeons depending on the clinical question, but with respect to the NCCN guidelines, which was your apparent focus, not all content is derived from a retrospective database. So, how can the conclusions in your study be tempered to reflect this fact?



Dr Kimberly Shemanski (*Los Angeles, Calif.*). You bring it up a couple of great points for discussion. We didn't have a lot of time to delve into the interview script here, but the interview questions specifically asked about the NCCN guidelines, which is why the hypotheses generated by this study

refer to perceptions about that particular set of guidelines.

Additionally, the phrasing of the questions very pointedly differentiated between guidelines and retrospective database analyses. So the themes about one are not meant to be applicable to the other. The participants actually provided a lot more commentary on HSR than guidelines and

the distribution of the themes reflects that, with themes 1 through 4 applying to HSR and theme 5 applying to guidelines.

You bring up an interesting point that not all of the content in the NCCN guidelines is from retrospective database analyses, but as we pointed out earlier, retrospective database analyses comprise a lot of surgical literature, so it's unclear how much HSR contributes to guideline formation and it appears that physicians have mixed perceptions of HSR and guidelines. So when you take those 2 things together, you start questioning how much of an impact the bulk of what we publish really has on clinical practice. I should clarify that when I said they had mixed perceptions, it was anywhere from people questioning whether the guidelines matter because it's not clear that anyone follows them, to calling them lovely and saying that they are very well thought out and they pull them up with every patient in their clinic. So I think the main point is that there's a lot of disconnect in a lot of places that just requires further evaluation.

Dr Litle. It is great that you're trying to educate the surgeons about HSR. My second question is, 6 of 27 interviewees did not attend a weekly tumor board and NCCN guidelines are typically applied in that group format discussion. How did this 22% of the sample not cause selection bias on your results?

Dr Shemanski. This is a fair question because as surgeons we're used to evaluating quantitative data, which relies pretty heavily on a statistical argument, but qualitative data are a little different in the sense that what we're really looking for is thematic saturation, which is essentially when participants' responses become redundant to the point that new participants don't introduce any additional themes. So exact numbers and percentages are a little less relevant with qualitative research; what's more important is what's being said and how that's being echoed by other participants.

Additionally, we didn't have providers specify where they were using the guidelines, whether that was tumor board or elsewhere, and some providers actually volunteered that where they use a lot of their guidelines are actually in clinic to assist with shared decision-making.

So it's difficult for us to comment on how tumor board attendance may or may not have biased perceptions of the guidelines, but the one thing that we can say from all of this is that a lot of studies have suggested that there is a relationship between guideline-concordant care and improved patient survival.

Given this study, guideline use remains pretty controversial and it's important for us to understand why providers may stray from the guidelines, and when they do, how to fix that. And most importantly, how to present new research and new information so that it's impacting clinical practice.

APPENDIX E1. INTERVIEW SCRIPT**Section III. Knowledge and Opinion on Guidelines**

Thank you for your thoughts. Now I'd like to ask about your thoughts on the current NCCN guidelines.

1. In your professional knowledge and experience, what types of NSCLC patients may be recommended for therapeutic surgery, according to guidelines?
 - a. From your professional experience and knowledge, is surgery associated with improved survival for any of these patients?
 - b. In your view, are there patients who fall into a "grey area" with respect to these guidelines?
 - c. What do you think of these guidelines? Should they be amended?
2. Recent research has shown that despite the generally poor prognosis for patients with advanced NSCLC, isolated patients with greater disease burdens who have had surgery as part of a multimodality treatment regimen and that it is common for these patients to have superior survival outcomes when compared to treatments without surgery.
 - a. What are your thoughts about these cases?
 - i. **[if selection bias mentioned]** What does selection bias mean to you?

- ii. **[if selection bias not mentioned]** Do you think selection bias might play a role in extended survival in these cases?

- b. What role does outcomes research play in your decision-making?
 - i. More specifically, how it relates to treatments you offer NSCLC patients?
 - ii. What role does it play in your decision to recommend surgery to a patient?

Thank you very much for sharing your opinions. You have helped us a great deal by sharing your thoughts and opinions about surgical decision making for advanced stage NSCLC.

- a. If you were me and wanted to know as much as possible about how surgeons make decisions about the treatment of NSCLC, is there anything else you would ask?
- b. Is there anything else that you would like to add to what you said already?
- c. We are also hoping to interview individuals that are not academically affiliated. We are permitted to obtain up to 2 referrals from each participant. If you have one or 2 people that you think would be willing to do the interview can you please provide me their name and the best way to contact them?