Paper PD20

Data Storytelling: A Crucial Leadership Skill in Today's Clinical Research Landscape

Bhavesh Patkar, Syneos Health, Hyderabad, India
Brian Thornton, Spaulding Clinical Research, West Bend, Wisconsin, USA

ABSTRACT

In an era of increasing data complexity, the ability to tell a clear and compelling story with data has become a core leadership skill. While technical expertise ensures accurate analyses, it is effective data storytelling that bridges the gap between data and decision-making. Leaders who can translate numbers into meaningful narratives drive alignment, influence stakeholders, and accelerate outcomes.

This paper explores data storytelling as a critical capability for today's biometrics leaders. It highlights how to frame data within context, tailor messages for diverse audiences, and balance technical accuracy with accessibility. Through illustrative examples and practical strategies, the paper demonstrates how effective data storytelling can transform clinical data into actionable insight, enabling leaders to communicate with clarity, build trust, and drive meaningful impact in a data-driven clinical research environment.

INTRODUCTION

Data storytelling is often viewed as a modern skill, shaped by advances in analytics, visualization tools, and increasing volumes of complex data. In reality, the principles of effective data storytelling have been central to impactful decision-making for well over a century. One of the earliest and most influential examples comes from Florence Nightingale during the Crimean War.

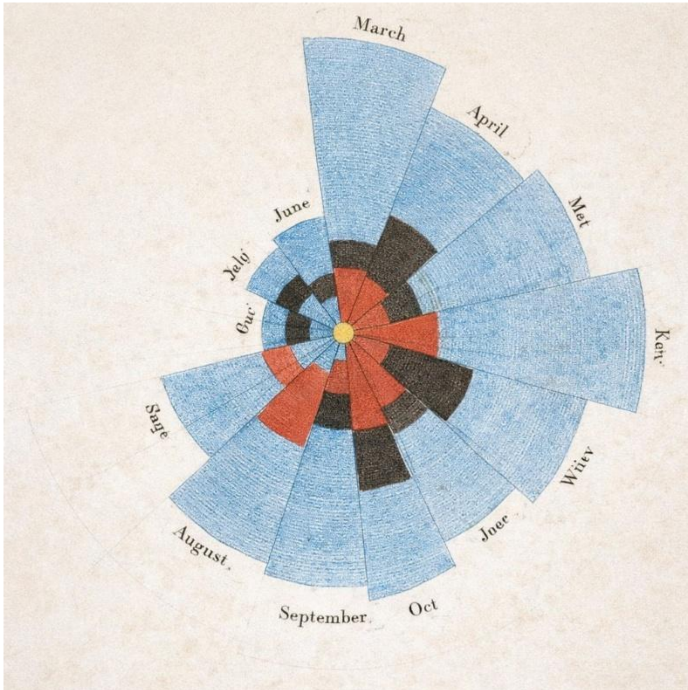
Faced with alarmingly high mortality rates among British soldiers, Nightingale systematically collected and analyzed data on causes of death. Rather than relying solely on tables of figures or written reports, she translated the data into clear visual narratives—most notably her polar area diagrams—that revealed a compelling and previously overlooked story. The majority of deaths were not due to battlefield injuries, but to preventable causes such as infection and poor sanitation. By presenting this information in a way that senior military leaders could quickly understand, Nightingale influenced policy changes that dramatically reduced mortality rates.

Nightingale's work illustrates a critical lesson that remains highly relevant today: accurate data alone is rarely sufficient to drive action. It is the ability to frame data within a clear and persuasive narrative that enables leaders to bridge the gap between analysis and decision-making. Her success was not rooted solely in statistical rigor, but in her ability to communicate insights in a form that aligned with the needs and perspectives of her audience.

In today's clinical research landscape, leaders in biometrics face a similar challenge. Advances in data collection, statistical methodology, and technology have resulted in unprecedented volumes of high-quality data. However, the increasing complexity of analyses can make it difficult for stakeholders—ranging from clinicians and operations teams to senior executives and regulators—to interpret results efficiently and confidently. As a result, valuable insights may fail to translate into timely or effective decisions.

This paper positions data storytelling as a crucial leadership skill for professionals working in clinical research. It explores how leaders can use storytelling principles to provide context, tailor messages for diverse audiences, and balance technical accuracy with accessibility, ultimately transforming complex clinical data into actionable insight.

Figure 1. Florence Nightingale's polar area diagram illustrating causes of mortality among British soldiers during the Crimean War (public domain).



WHAT IS DATA STORYTELLING IN CLINICAL RESEARCH?

Data storytelling in clinical research extends beyond the presentation of results or the creation of visually appealing charts. At its core, data storytelling is the deliberate integration of data, context, and narrative to convey insight in a way that supports understanding and informed decision-making. As demonstrated by Florence Nightingale's work, effective data storytelling combines analytical rigor with purposeful communication, ensuring that key messages resonate with the intended audience.

In a clinical research setting, data are abundant and technically complex. Statistical outputs, listings, and tables are essential for accuracy, traceability, and regulatory compliance. However, these elements alone do not automatically translate into insight. Data storytelling bridges this gap by framing analyses within a meaningful context, clarifying why results matter, and guiding stakeholders toward appropriate interpretation without compromising scientific integrity.

FROM DATA TO INSIGHT

A useful way to understand data storytelling is to distinguish between data, information, and insight. Data represent raw observations, such as patient measurements, adverse event counts, or laboratory values. Information emerges when data are organized and summarized through analyses, tables, or figures. Insight is achieved only when information is interpreted within context and connected to a specific question or decision.

In clinical research, leaders are often responsible for making decisions based on insights rather than raw data. These decisions may involve assessing benefit–risk profiles, prioritizing development activities, responding to safety signals, or communicating results to regulatory authorities. Data storytelling enables biometrics leaders to elevate analyses from descriptive outputs to actionable insights by explicitly linking results to their implications.

STORYTELLING WITHOUT OVERSIMPLIFICATION

A common concern among statisticians and programmers is that storytelling may lead to oversimplification or loss of technical nuance. In regulated environments, maintaining accuracy, transparency, and reproducibility is non-negotiable. Effective data storytelling does not remove complexity; instead, it manages complexity by presenting information at an appropriate level for the audience while preserving access to underlying detail.

Just as Nightingale's visualizations did not replace the underlying mortality data but illuminated its key message, modern data storytelling should complement—not substitute—detailed analyses. Well-crafted narratives guide stakeholders to the most relevant findings while allowing deeper exploration when required.

AUDIENCE AND PURPOSE

An essential element of data storytelling is an understanding of audience and purpose. The same analysis may need to be communicated differently to clinicians, operational teams, senior executives, or regulators. Data storytelling involves tailoring the narrative to align with the audience’s priorities, expertise, and decision-making responsibilities, while ensuring that scientific conclusions remain consistent across communications.

In this way, data storytelling becomes a leadership skill rather than a presentation technique. It reflects an ability to anticipate questions, address uncertainty transparently, and communicate with clarity and intent.

DATA STORYTELLING AS A LEADERSHIP PROCESS

Effective data storytelling does not occur by chance. It is the result of a deliberate process that guides how data are framed, presented, and interpreted. For leaders in clinical research, this process provides a structured approach to transforming complex analyses into clear, decision-focused narratives.

The data storytelling process described in this paper consists of six interconnected steps: context, appropriate display, elimination of clutter, drawing attention, design, and story.

1. **Context**: Context establishes the foundation for effective data storytelling by clarifying the clinical question, study objectives, and decision being supported. Without context, even accurate analyses may be misinterpreted or overlooked.
2. **Appropriate Display**: Selecting an appropriate display ensures that data are interpreted correctly. Different analytical questions require different visual representations, and leadership judgment is required to choose displays that best support the intended message.
3. **Eliminate Clutter**: Clutter distracts from insight. Excessive labels, unnecessary gridlines, and overly dense visuals increase cognitive burden and obscure key findings. Effective leaders prioritize clarity through purposeful omission.
4. **Draw Attention**: Once clutter is reduced, attention can be intentionally directed to clinically meaningful patterns, trends, or thresholds. This step guides stakeholders toward relevant insights without biasing interpretation.
5. **Design**: Design choices—including layout, color, and consistency—support comprehension and reinforce professionalism. Thoughtful design signals intent and respect for the audience’s time and attention.
6. **Story**: The final step is the narrative itself, connecting context, visuals, and data into a coherent message that supports understanding and decision-making while maintaining scientific integrity.

BALANCING TECHNICAL ACCURACY AND ACCESSIBILITY

Balancing technical accuracy with accessibility is a central challenge in clinical research communication. While rigor and traceability are essential, excessive complexity can obscure insight and delay decision-making.

Risks of Over-Simplification

Oversimplification may result in loss of nuance or misinterpretation. Leaders must ensure that clarity does not come at the expense of accuracy or transparency.

Risks of Over-Complexity

Conversely, overly detailed presentations can overwhelm stakeholders and mask important findings. Accessibility should be viewed as a component of impact, not a compromise of rigor.

Layered Communication

Layered communication provides information at multiple levels of detail, enabling stakeholders to engage with data according to their role and expertise while preserving accountability and traceability.

Transparency and Trust

Clear communication of assumptions, limitations, and uncertainty builds trust and reinforces leadership credibility.

REAL-WORLD (ILLUSTRATIVE) EXAMPLES

This section presents illustrative, fictional examples based on common scenarios encountered in clinical research and biometrics leadership. Each example contrasts ineffective and effective data storytelling using the same underlying data, demonstrating how leadership choices in framing, emphasis, and narrative influence interpretation and decision-making.

EXAMPLE 1: SAFETY SIGNAL COMMUNICATION

Scenario:
A higher incidence of elevated liver enzymes is observed in the treatment arm during a clinical study.

Ineffective Storytelling:
In Figure 2a, safety data are presented as a comprehensive adverse events summary table with equal emphasis across all reported events. While the table is complete and statistically accurate, it requires the reader to manually scan rows and columns to identify potentially meaningful differences between treatment groups. Temporal patterns are not visible, and clinically relevant thresholds are not emphasized. As a result, identifying whether a true safety signal is emerging—or whether observed differences reflect background variability—requires substantial effort and interpretation.

This type of presentation places a high cognitive burden on stakeholders and may lead to inconsistent conclusions. Some readers may overlook the developing signal entirely, while others may overreact to isolated numerical differences without sufficient context.

Figure 2a. Adverse events summary table with equal visual weight across events and no temporal context.

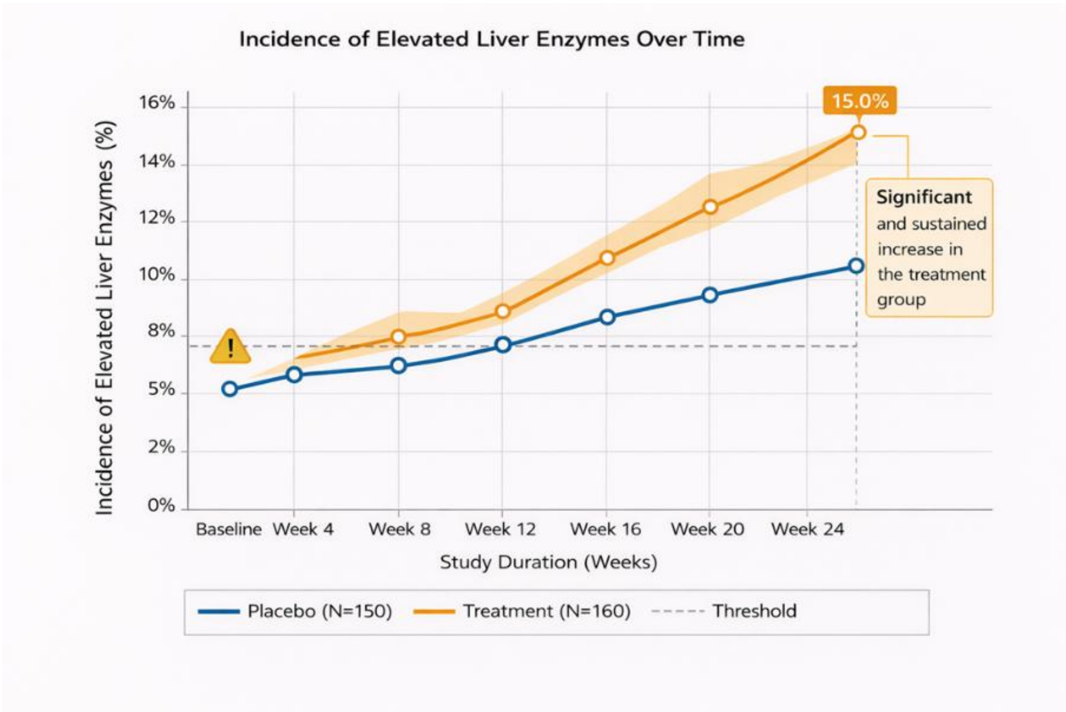
Adverse Events Summary

Adverse Event	Placebo (N=150) n (%)	Treatment (N=160) n (%)	Total (N=310) n (%)
Headache	23 (15.3)	28 (17.5)	51 (16.5)
Nausea	17 (11.3)	30 (18.8)	47 (15.2)
Fatigue	14 (9.3)	22 (13.8)	36 (11.6)
Dizziness	10 (6.7)	18 (11.3)	28 (9.0)
Rash	8 (5.3)	15 (9.4)	23 (7.4)
Diarrhea	12 (8.0)	20 (12.5)	32 (10.3)
Vomiting	6 (4.0)	14 (8.8)	20 (6.5)
Insomnia	4 (2.7)	11 (6.9)	15 (4.8)
Chest Pain	7 (4.7)	10 (6.3)	17 (5.5)
Elevated Liver Enzymes	16 (10.7)	24 (15.0)	40 (12.9)
Hypertension	7 (4.7)	18 (11.3)	14 (4.5)
Cough	15 (10.0)	9 (5.6)	33 (10.6)
Back Pain	13 (8.7)	9 (5.6)	21 (6.8)
Fever	13 (8.7)	9 (5.6)	21 (6.8)
Constipation	6 (4.0)	8 (5.0)	20 (6.5)
Anxiety	11 (7.3)	10 (6.3)	25 (8.1)
Other Adverse Events	29 (19.3)	39 (22.5)	65 (21.0)

*TEAE = Treatment-Emergent Adverse Event
Data presented as n (%).

Effective Storytelling:
In contrast, Figure 2b reframes the same data by focusing explicitly on the adverse event of interest and displaying incidence over time. By presenting elevated liver enzyme incidence longitudinally and in relation to a clinically meaningful threshold, the visualization makes both the magnitude and trajectory of the signal immediately apparent. Minimal annotation draws attention to the sustained divergence between treatment and placebo groups while preserving transparency around comparator trends and uncertainty.

Figure 2b. Incidence of elevated liver enzymes over time with clinical threshold and targeted annotations.



Leadership Impact:
This shift from a dense summary table to a focused, time-based visualization enables timely and proportionate decision-making. Leaders can quickly assess whether enhanced monitoring, protocol review, or further investigation is warranted, without unnecessary alarm or delay. The example illustrates how effective data storytelling transforms accurate but opaque data into actionable clinical insight.

EXAMPLE 2: BIOEQUIVALENCE INTERPRETATION

Scenario:
A bioequivalence analysis compares capsule and caplet formulations for two drugs using standard pharmacokinetic parameters, including AUC and Cmax.

Ineffective Storytelling:
In Figure 3a, bioequivalence results are presented as a comprehensive statistical table summarizing least squares geometric means, ratios, and confidence intervals for multiple parameters and drugs. Although analytically correct and complete, the table assigns equal visual weight to all values and requires readers to mentally apply acceptance criteria, scan confidence intervals, and infer conclusions. For stakeholders without deep familiarity with bioequivalence conventions, determining whether bioequivalence is achieved—and for which parameters—can be time-consuming and error-prone.

Figure 3a. Dense bioequivalence summary table requiring manual interpretation of acceptance criteria.

Results – *Summary of Bioequivalence Main Analysis*

Summary of Bioequivalence Main Analyses
PK Population

Analyte	Capsules (Treatment A, Test)		Caplets (Treatment B, Reference)		LSGMR (Test/Reference)	90% Confidence Interval	
Parameter (units)	N	LSGM	N	LSGM	%	Lower	Upper
Drug 1 (ng/mL)							
AUC_{0-t} (h*ng/mL)	45	82497.75	43	75735.54	97.38	96.28	100.41
C_{max} (ng/mL)	45	33087.42	43	23451.09	137.45	131.00	145.31
Drug 2 (ng/mL)							
AUC_{0-t} (h*ng/mL)	45	28393.43	43	23532.57	104.92	102.30	108.23
C_{max} (ng/mL)	45	8675.81	43	7657.31	112.71	106.32	117.54

Abbreviations: CV%, percent coefficient of variation; ln, natural logarithm; LS, least squares; LSGMR, least squares geometric mean ratio; PK pharmacokinetic.
The primary analyses included comparisons of test treatment (capsules [Treatment A]) and reference treatment (Caplets [Treatment B]) using a mixed-effects analysis model to evaluate the effect of treatment, sequence and period as fixed effects, and subject nested within sequence as a random effect for PK parameters C_{max} and AUC_{0-t}. PK parameters were ln-transformed for analysis and have been back transformed in table presented above. Only the treatments of interest were included in the model (Treatment A and Treatment B). An available case analysis was performed using all observed data.
Bioequivalence is determined by the 90% confidence interval of the LSGMR for C_{max} and AUC_{0-t}, if LSGMR falls completely within the range 80.00-125.00%. This indicates statistical equivalence between Capsules (Treatment A) and Caplets (Treatment B).

Effective Storytelling:
Figure 3b presents the same results in a decision-focused format that explicitly surfaces the conclusions relevant to regulatory and development decisions. Visual emphasis distinguishes parameters that meet bioequivalence criteria from those that do not, while concise narrative summaries translate statistical outputs into clear, interpretable statements aligned with the accepted bioequivalence range. Importantly, this approach preserves traceability to the underlying analysis while eliminating the need for readers to derive conclusions themselves.

Figure 3b. Decision-focused bioequivalence summary highlighting parameters that meet and do not meet acceptance criteria.

Results – *Summary of Bioequivalence (Primary Analysis)*

Analyte	Capsules (Treatment A, Test)		Caplets (Treatment B, Reference)		LSGMR (Test/Reference)	90% Confidence Interval	
Parameter (units)	N	LSGM	N	LSGM	%	Lower	Upper
Drug 1 (ng/mL)							
AUC_{0-t} (h*ng/mL)	45	82497.75	43	75735.54	97.38	96.28	100.41
C_{max} (ng/mL)	45	33087.42	43	23451.09	137.45	131.00	145.31
Drug 2 (ng/mL)							
AUC_{0-t} (h*ng/mL)	45	28393.43	43	23532.57	104.92	102.30	108.23
C_{max} (ng/mL)	45	8675.81	43	7657.31	112.71	106.32	117.54

Abbreviations: LSGMR, least squares geometric mean ratio.
The primary analyses included comparisons of test treatment (Capsules [Treatment A]) and reference treatment (Caplet [Treatment B]) using a mixed-effects analysis model to evaluate the effect of treatment, sequence and period as fixed effects, and subject nested within sequence as a random effect for PK parameters above. PK parameters were ln-transformed for analysis and have been back transformed in table presented above.

For **Drug 1**, LSGMR for **AUC (97.38%)** and CI (96.28, 100.41) are within the accepted bioequivalence range of 80.00% to 125.00%, indicating equivalent exposure between Capsules and Caplets. However, the LSGMR for **Cmax (137.45%)** and CI (131.00, 145.31) are outside the bioequivalence acceptance range.

For **Drug 2**, LSGMR for **AUC (104.92%)** and **Cmax (112.71%)** and respective CI (102.30, 108.23) (106.32, 117.54) are within the accepted bioequivalence range of 80.00% to 125.00%.

Leadership Impact:
By reframing dense statistical output into a decision-ready summary, leaders can quickly assess the implications of the analysis and focus discussion on next steps rather than numerical interpretation. This example demonstrates how effective storytelling accelerates understanding without sacrificing scientific rigor or transparency.

EXAMPLE3: FINANCIAL PERFORMANCE AND BUDGET OVERSIGHT

Scenario:
Biometrics leaders are responsible for managing project budgets and explaining variances to senior stakeholders during mid-year financial reviews.

Ineffective Storytelling:
In Figure 4a, cost performance is presented as a detailed month-by-month table comparing budgeted and actual costs, with future months included despite the absence of actual data. While the table is numerically accurate, all months are presented with equal visual weight, and the distinction between historical performance and future periods is subtle. Identifying when budget deviations began, whether the variance is accelerating, and what this implies for year-end performance requires manual calculation and interpretation.

Figure 4a. Month-by-month budget versus actual cost table with limited emphasis on trend or risk.

Cost Summary – 2025 (shown in $K)

Actuals available through June only

Month	Budget Cost	Actual Cost	Difference (Actual – Budget)
Jan	100.0	100.0	0.0
Feb	101.0	100.5	-0.5
Mar	102.0	101.5	-0.5
Apr	103.0	106.0	+3.0
May	104.0	108.0	+4.0
Jun	105.0	110.0	+5.0
Jul	106.0	—	—
Aug	107.0	—	—
Sep	108.0	—	—
Oct	109.0	—	—
Nov	110.0	—	—
Dec	111.0	—	—
Total	**1,266.0**	**626.0 (YTD)**	**+11.0 (YTD)**

Effective Storytelling:
In Figure 4b, the same information is reframed to separate confirmed historical performance from forward-looking projections. Actual costs through June are clearly distinguished from projected costs for the remainder of the year using visual cues, and the narrative explicitly links observed variances to specific, time-bound drivers such as temporary resource needs, accelerated project timelines, and the initiation of a new critical project. Forward-looking projections incorporate these assumptions, enabling assessment of whether the budget overrun is likely to persist or resolve.

Figure 4b. Budget versus actual cost trend with projections and explanatory context.

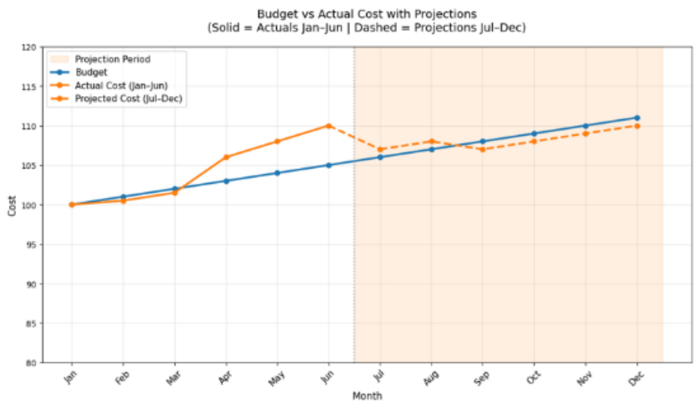
Cost Summary – 2025 (shown in $K)

Actuals available through June only

Month	Budget Cost	Actual Cost	Difference (Actual – Budget)
Jan – Jun (Actual)	615.0	626.0	+11.0
Jul – Dec (Projection)	651.0	649.0	-2.0
Total	**1,266.0**	**1275.0**	**+9.0**

Reasons for Budget Overrun

- ***Project Lynx** – Legacy project with additional unanticipated ad-hoc analysis, completed in June ($4k)*
- ***Project Zebra** – New Critical project not in original budget, will complete in August ($5k)*
- ***Project Aardvark** – Accelerated Start, originally slated for Sep-Dec, now completed in Apr-Jun. Net zero impact to budget for the year (cost shifted)*



Leadership Impact:

This reframed presentation enables proactive financial management by shifting discussion from retrospective variance explanation to forward-looking decision-making. Leaders can evaluate the effectiveness of corrective actions, anticipate remaining risk, and make informed choices about resource allocation. The example highlights how effective data storytelling supports accountability and transparency while enabling timely, confident leadership decisions.

PRACTICAL STRATEGIES FOR DEVELOPING DATA STORYTELLING SKILLS

Developing data storytelling capability requires intentional practice and leadership reinforcement. The following strategies support consistent application across analyses, reviews, and presentations.

- **Adopt a storytelling mindset:** *Treat analyses as decision inputs rather than endpoints.*
- **Ask the right questions early**:
  What decision does this support?
  Who is the audience?
  What is the primary message?
  What uncertainty or limitation must be communicated?

- **Apply principles consistently:** *Use context, focus, and clarity across tables, figures, and narratives; reduce clutter and guide attention deliberately.*
- **Coach and model:** *Leaders reinforce effective communication through review practices, feedback, and examples.*

THE ROLE OF AI AS A STORYTELLING COPILOT

Artificial intelligence can enhance data storytelling by supporting structure, clarity, and iteration. AI tools can assist with summarization, drafting narratives, and tailoring communication for different audiences.

However, accountability for interpretation and decision-making must remain with human leaders. AI-generated content should be critically reviewed to ensure alignment with source data, study objectives, and regulatory expectations. Appropriate guardrails include documentation of AI use, validation against original analyses, and adherence to PHUSE guidance on transparency and disclosure. When applied responsibly, AI can serve as a storytelling copilot that enhances—not replaces—human judgment.

CONCLUSION

In an increasingly complex clinical research environment, data storytelling has emerged as a critical leadership capability. While technical rigor remains essential, the ability to contextualize data, communicate clearly, and guide interpretation determines whether analyses translate into meaningful decisions.

This paper has demonstrated how data storytelling bridges the gap between analysis and action through practical frameworks and illustrative examples. By intentionally developing these skills, biometrics leaders can enhance credibility, build trust, and increase the impact of their work—ensuring that high-quality data lead to high-quality decisions.

AI USE DISCLOSURE

AI tools were used to assist with drafting and refining portions of this paper, including improving clarity, structure, and technical explanations. All content was reviewed, edited, and validated by the authors, who retain full responsibility for the accuracy and integrity of the material presented.

REFERENCES

- Knaflic C. *Storytelling With Data: A Data Visualization Guide for Business Professionals.* Wiley; 2015.
- Schwabish J. *Better Data Visualizations: A Guide for Scholars, Researchers, and Wonks.* Columbia University Press; 2019.
- Cairo A. *The Truthful Art: Data, Charts, and Maps for Communication.* New Riders; 2016.
- Few S. *Show Me the Numbers: Designing Tables and Graphs to Enlighten.* Analytics Press; 2012.
- Yau N. *Data Points: Visualization That Means Something.* Wiley; 2013.
- McCandless D. *Information Is Beautiful.* Harper Design; 2012.
- Duarte N. *Resonate: Present Visual Stories That Transform Audiences.* Wiley; 2010.
- Anderson C. *TED Talks: The Official TED Guide to Public Speaking.* Houghton Mifflin Harcourt; 2016.
- Tufte E. *The Visual Display of Quantitative Information.* Graphics Press; 2001.
- Davenport T, Harris J. *Competing on Analytics: The New Science of Winning.* Harvard Business Review Press; 2007.
- Kahneman D. *Thinking, Fast and Slow.* Farrar, Straus & Giroux; 2011.
- Bolman LG, Deal TE. *Reframing Organizations: Artistry, Choice, and Leadership.* Jossey-Bass; 2017.
- Glasziou P, Ioannidis JPA, Guyatt C. *Clinical Epidemiology: How to Do Clinical Practice Research.* 4th ed. Elsevier; 2013.
- OpenAI. (2025). *ChatGPT (GPT-4o, 2 July version)* [Large language model]. https://chat.openai.com/

CONTACT INFORMATION

Author Name: Bhavesh S. Patkar
Company: Syneos Health
Email: bhavesh.patkar@gmail.com

Author Name: Brian Thornton
Company: Spaulding Clinical Research
Email: brian.thornton@spauldingclinical.com

DISCLAIMER

The opinions expressed in this paper are solely those of the authors and do not necessarily reflect the views of their respective organizations. Brand and product names are trademarks of their respective companies.