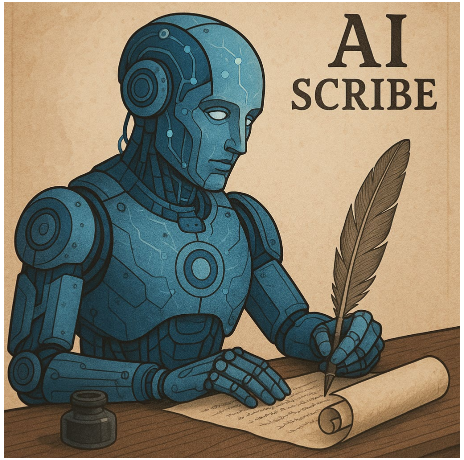
Use of Ambient AI Scribes to Reduce Administrative Burden and Professional Burnout: Lessons Learned and Future Directions

December 18, 2025

Kristine Olson, MD, MSc
Assistant Clinical Professor
Department of Medicine
Yale School of Medicine

Daniella Meeker, PhD
Associate Professor of Biomedical Informatics and Data Science
Yale School of Medicine
Chief Research Information Officer
Yale New Haven Health System

Lee H. Schwamm, MD
Sr Vice President and Chief Digital Health Officer
Yale New Haven Health System
Associate Dean, Digital Strategy & Transformation
Professor of Neurology and of Biomedical Informatics and Data Science
Yale School of Medicine



Disclosures

Mr. Troupe reported being senior director of clinical success at Abridge AI.

Dr. T. Shah reported being chief clinical officer of Abridge AI during the conduct of the study.

Dr. Jones reported receiving personal fees from Intrinsic Brands outside submitted work.

Dr. Schwamm reported serving as a volunteer member of a research advisory committee for Abridge to facilitate the use of ambient AI for clinical researchers.

No other disclosures reported.

Learning Objectives

- Understand the basics of ambient artificial intelligence-(AI) scribes.
- Give three examples how the ambient-AI platform reduced burdens and burnout.
- Describe how the ambient-AI could improve the patient-physician relationship and enhance quality care.
- List three additional ways AI technology could reduce administrative burdens, enhance clinical efficiency, and support professional satisfaction.

Background

Physicians are in short supply and high demand. Clinic-based ambulatory care physicians spend more than half of their workday documenting in the electronic medical record (EHR), and only a quarter of the time is face-to-face with patients. The proportion of time spent documenting continues to escalate, and it is associated with burnout, reduction in work effort.

Aims

Could 30-days of the Abridge ambient-AI scribe reduce burnout, cognitive task load, time spent documenting after-hours, and increase undivided attention on patients, create notes that patients can understand, and enable adding patients to the clinic schedule if urgently needed?

AAMC2024, Arndt et al 2017, Arndt et al. 2024, Sinsky et al 2017, Holmgren AJ et al. 2023, Holmgren et al 2022, Gardner RL 2019, Sinsky CA,et al 2024, Melnick ER, et al. 2021, Sinsky CA 2017, Doan-Wiggins L et al. 1995, Medicine NAo.

K

Benchmarks: Ambient and remote human scribes reduce proportion with burnout by 18-22%

Contents lists available at ScienceDirect

Healthcare

journal homepage: www.elsevier.com/locate/healthcare

ELSEVIER

The effect of remote scribes on primary care physicians' wellness, EHR satisfaction, and EHR use

Mark A. Micek[a,*], Brian Arndt[b], Jeffrey J. Baltus[b], Aimee Teo Broman[c], Joel Galang[d,1], Shannon Dean[e,2], Matthew Anderson[a,3], Christine Sinsky[f]

burnout reduced from 71% to 51% after remote human scribes

Original Investigation | Health Informatics

Ambient Documentation Technology in Clinician Experience of Documentation Burden and Burnout

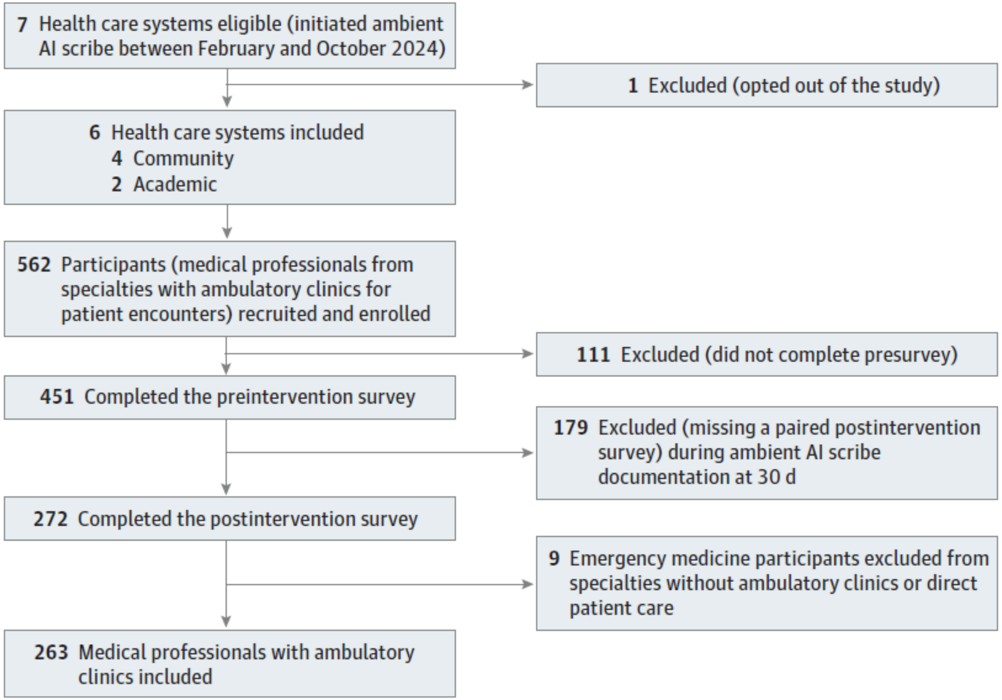
Jacqueline G. You, MD[1,2]; Reema H. Dbouk, MD[3]; Adam Landman, MD, MS, MIS, MHS[1,4,5] ; et al

burnout reduced from 52% to 30% after ambient scribes

Our survey includes potential mechanisms contributing to burnout reduction (e.g. cognitive load, focus on patient)

K

Figure. Inclusion and Exclusion Criteria

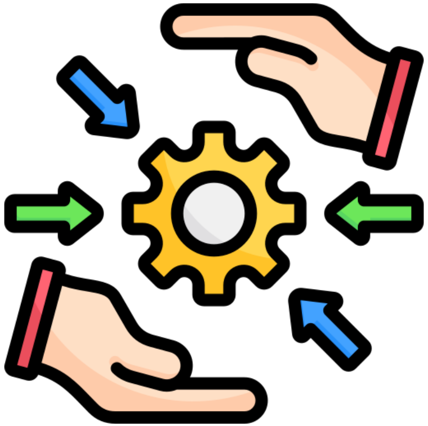


Setting and Participants

FEBRUARY TO OCTOBER 2024

Intervention: Use of the Ambient Scribe

- Selected the patient from their EHR schedule
- Obtained verbal consent to record the encounter
- AI instantaneously generated a standard medical note
- An online secure portal allowed viewing and editing
- Underlying transcripts or audio can be viewed or heard
- Text is imported into the clinician's EHR note template
- Original recordings and transcripts are erased from the system
- all sites used the same software version of the technology



Survey Design and Distribution

- YNHHS coordinated with Abridge to modify client satisfaction survey
- Abridge and Yale worked together to create a cohort of participating hospitals
- Vendor aggregated all the data and delivered a deidentified dataset to Yale researchers for statistical analysis and manuscript generation
- Vendor provided no financial support and had no editorial control over the publication
- Full statistical analysis plan was submitted via pre-registration to AsPredicted.org and available to readers and reviewers

D

Statistical Analysis

Pre-registered primary outcome: Change in the ordinal burnout score after using the AI scribe adjusting for intra cluster correlation among 6 health systems.
Assumptions:

- $\alpha = .05, \beta = 0.80$
- 194 Participants with complete burnout survey data
- Health systems' baseline ICC: 0.06
- Baseline SD Burnout: 0.9
- Analysis model: hierarchical linear model (HLM) with random intercept for health system

Additional analyses: paired (pre, post) t-tests on unadjusted values; univariate and multivariate HLMs adjusting for baseline characteristics random intercepts for nested clusters

Wharton CREDIBILITY LAB UNIVERSITY of PENNSYLVANIA

ASPREDICTED

80% power to detect less than single item shift in mini-Z

Insufficient power for dichotomized non-inferiority analysis (vs. human scribes)

D

Measures

Content	Question
Mini Z: Burnout	Using your own definition of “burnout”, please 1 - I enjoy my work. I have no symptoms of burnout 2 - I am under stress, and don’t always have as much energy as I did, but I don’t feel burned out 3 - I am beginning to burn out and have one or more symptoms of burnout (e.g. emotional exhaustion) 4 - The symptoms of burnout that I’m experiencing won’t go away. I think about work frustrations a lot 5 - I feel completely burned out. I am at the point where I may need to seek help
NASA TLX: Mental Demand	how mentally demanding is it to write your notes?
NASA TLX: Temporal Demand	how hurried / rushed is the pace of your note writing?
NASA TLX: Effort	how hard do you have to work to accomplish your level of note-writing performance?
Undivided Attention	I'm able to give patients my undivided attention during the encounter.
Work Outside Work	The average amount of time I spend per week writing notes outside of clinic hours is:
Patient Access	I feel that I could add at least one more patient encounter to my clinic session if urgently needed.
Patient Access – Branch	[ask if Agree or Strongly Agree to Access Question above] I estimate the number of patient encounters I could add to my clinic session is: 1 patient 2 patients 3 patients >3 patients
Patient Comprehension	My notes are as comprehensive and complete as I would like them to be for clinical purposes.
Pre-Intervention Feedback	Do you have any other feedback or thoughts you'd like to share

Table 1. Demographics of 263 Participants Completing the Preintervention and Postintervention Surveys *

Characteristic	Participants, No. (%)
Health system site	
1	44 (16.7)
2	17 (6.5)
3	9 (3.4)
4	19 (7.2)
5	69 (26.2)
6	105 (39.9)
Clinician type	
Medical doctor	232 (88.2)
Advanced practice practitioner	29 (11.0)
Unknown	2 (0.8)
Practice model	
Academic	168 (63.9)
Medical group employed	90 (34.2)
Community private practice	5 (1.9)
Specialty	
Family practice or internal medicine/pediatrics	55 (20.9)
Adult general internal medicine	38 (14.4)
Adult specialty care	46 (17.5)
Pediatrics	38 (14.4)
Neurology or psychiatry	14 (5.3)
Obstetrics and gynecology	27 (10.3)
Surgery	45 (17.1)
Experience level	
Years in practice, mean (SD)	15.1 (9.3)
Years in practice by group	
≥1 to ≤5	44 (16.9)
>5 to ≤10	46 (17.6)
>10 to ≤15	69 (26.4)
>15 to ≤20	36 (13.8)
>20	66 (25.3)
Sex	
Female	141 (53.6)
Male	120 (45.6)
Not reported	2 (0.8)

* The subset of providers who provided burnout data (n =194) were demographically similar

Unadjusted pre- and post-intervention scores reported as means (standard error, SE) and arithmetic difference of means (SE) transformed to 10-point scales.

K

Baseline Clinical Documentation Methods

- Manual typing (82.9%)
- Templates or dot-phrases (85.2%)
- Dictation (46.8%)
- Human scribes (16.3%)
- Previous experience with another ambient-AI scribe solution (1.5%)

Table 2. Univariable and Multivariable Models of the Association of the Intervention With Self-Reported Burnout

Model[a]	No. of participants	Mean (SE), %		Difference, percentage points	OR (95% CI)	*P* value
		Baseline	Follow-up			
Univariable	186	51.4 (4.1)	37.5 (4.1)	−13.9 (3.7)	0.30 (0.15-0.59)	<.001
Multivariable[b]	184	51.9 (3.3)	38.8 (3.3)	−13.1 (3.3)	0.26 (0.13-0.54)	<.001

Abbreviation: OR, odds ratio.

[a] Hierarchical mixed-effects logistic regression models with random intercepts for clinicians nested in sites.

[b] Multivariable models adjusted for degree, practice model, specialty, years in practice, sex, and site.

Table 3. Comparison of Secondary Outcome Measures Before and After Use of the Ambient AI Scribe

Outcome	No. of participants	Mean (SE) score[a]			*P* value
		Baseline	Follow-up	Difference	
Burnout	186	4.59 (0.15)	4.12 (0.15)	0.47 (0.12)	<.001
Note-related cognitive task load					
Any	243	7.10 (0.09)	4.46 (0.12)	2.64 (0.13)	<.001
Temporal demand	249	7.01 (0.11)	4.35 (0.13)	2.66 (0.16)	<.001
Effort	248	7.31 (0.12)	4.71 (0.13)	2.60 (0.15)	<.001
Mental demand	254	6.84 (0.12)	4.38 (0.15)	2.46 (0.15)	<.001
Documentation after hours	263	4.95 (0.18)	4.05 (0.16)	0.90 (0.19)	<.001
Focused attention on patients	253	6.51 (0.16)	8.56 (0.11)	−2.05 (0.18)	<.001
Comprehensible care plans	254	7.34 (0.13)	7.79 (0.13)	−0.44 (0.17)	.005
Agreeable to add urgent patients	230	6.21 (0.21)	6.72 (0.20)	−0.51 (0.24)	.02
No. of additional patients (1 to ≥4)	91	2.19 (0.11)	2.16 (0.11)	0.02 (0.11)	.58

Abbreviation: AI, artificial intelligence.

[a] Unadjusted preintervention and postintervention paired *t* tests transformed to 10-point scales.

means (standard error, SE) and arithmetic difference of means (SE)

Exploratory Secondary Outcomes: Burnout By Demographic Features

	n	Baseline Mean (SE)	Follow-up Mean (SE)	Difference Mean (SE)	p-value diff>0	p-value diff !=0
BURNOUT-ONLY SAMPLE						
Total	194					
Health System Site:						
1	44	4.17 (0.30)	3.91 (0.29)	0.26 (0.60)	0.128	0.256
2	17	5.90 (0.55)	4.18 (0.55)	1.72 (0.60)	0.005	0.011
3	9	4.00 (0.75)	3.50 (0.88)	0.50 (0.63)	0.223	0.447
4	19	5.03 (0.50)	4.32 (0.47)	0.71 (0.42)	0.055	0.111
5	-	-	-	-	-	-
6	97	4.53 (0.53)	4.22 (0.42)	0.30 (0.30)	0.014	0.027
Missing	0					
Practice model:						
academic	91	4.56 (0.20)	4.23 (0.21)	0.32 (0.14)	0.011	0.023
community private practice	5	4.15 (0.55)	4.15 (0.55)	0 (0.71)	0.500	1.000
medical group employed	90	4.65 (0.23)	4.00 (0.22)	0.65 (0.20)	<0.001	0.001
Missing	8					

Unadjusted pre- and post-intervention scores reported as means (standard error, SE) and arithmetic difference of means (SE) transformed to 10-point scales.

K

Exploratory Secondary Outcomes: Burnout By Demographic Features

	n	Baseline Mean (SE)	Follow-up Mean (SE)	Difference Mean (SE)	p-value diff>0	p-value diff !=0
Specialty:						
Family Med/Med-Peds	48	4.75 (0.30)	3.77 (0.29)	0.98 (0.28)	<0.001	<0.001
Adult Gen Med	23	3.64 (0.42)	3.25 (0.32)	0.39 (0.34)	0.128	0.257
Adult Specialties	27	4.83 (0.33)	4.33 (0.28)	0.50 (0.28)	0.042	0.083
Pediatrics	27	5.42 (0.42)	5.08 (0.40)	0.33 (0.23)	0.081	0.161
Neurology/Psychiatry	9	3.75 (0.50)	3.75 (0.50)	0 (0.38)	0.081	1.000
OB/Gyn	19	4.55 (0.40)	3.96 (0.42)	0.59 (0.34)	0.048	0.062
Surgery	33	4.41 (0.37)	4.48 (0.45)	0.07 (0.29)	0.594	0.812
Missing	8					

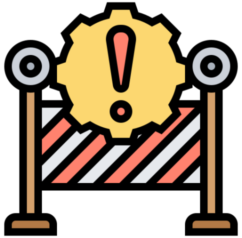
Unadjusted pre- and post-intervention scores reported as means (standard error, SE) and arithmetic difference of means (SE) transformed to 10-point scales.

Exploratory Secondary Outcomes: Burnout By Demographic Features

	n	Baseline Mean (SE)	Follow-up Mean (SE)	Difference Mean (SE)	p-value diff>0	p-value diff !=0
Degree:						
MD/DO/MBBS	179	4.64 (0.15)	4.12 (0.15)	0.52 (0.12)	<0.001	<0.001
NP/PA/APC	14	3.94 (0.59)	4.12 (0.65)	0.17 (0.31)	0.701	0.584
Other	1					
Years in practice, groups						
>=1 to <=5	38	4.67 (0.35)	4.43 (0.34)	0.24 (0.20)	0.127	0.254
>5 to <=10	40	4.49 (0.29)	4.26 (0.34)	0.23 (0.24)	0.176	0.352
>10 to <=15	54	4.38 (0.30)	4.00 (0.28)	0.38 (0.22)	0.048	0.095
>15 to <=20	24	4.94 (0.34)	3.81 (0.36)	1.13 (0.41)	0.006	0.011
>20	38	4.86 (0.40)	3.97 (0.37)	0.88 (0.32)	0.005	0.009
Sex:						
Female	102	5.08 (0.20)	4.62 (0.20)	0.46 (0.17)	0.004	0.009
Male	84	4.00 (0.20)	3.52 (0.20)	0.48 (0.16)	0.002	0.003
Missing	8					

K

Limitations



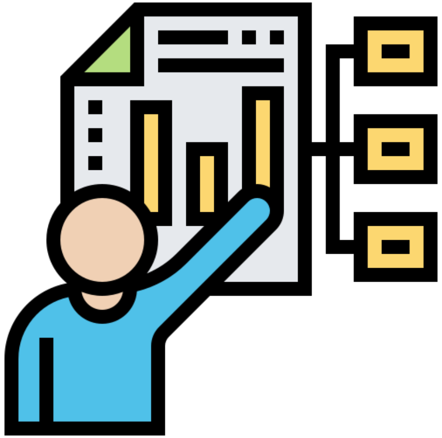
- This is a quality improvement initiative, not designed for research purposes.
- The dataset is one of convenience.
- The survey was not anonymous.
- Self-selection in recruitment and attrition may represent a biased perspective.
- There is no control group.
- We cannot characterize non-completers or non-responders.
- Participates may favor of new technologies.
- The findings are not paired with quantitative data from the EHR.

- However, the analysis does control for other factors including diversity in health systems (national sample of academic and community-based sites), professional degree, specialty, time in practice, and sex. These facts support potential generalizability to other health system ambulatory clinics.

Discussion & Conclusions:

After 30-days with an ambient-AI scribe, there was significant improvement in ...

- burnout amongst physicians and advanced practitioners with ambulatory clinics decreased significantly from **51.9% to 38.8%.**
- improvements in the cognitive task load, time spent documenting afterhours, focused attention on patients, and urgent access to care.



K

Acknowledgements:
AI Multicenter Study Team

Kristine Olson, MD MSc
Department of Medicine, Yale School of Medicine, New Haven CT
Office of the Chief Wellness Officer, Yale New Haven Health-Yale Medicine

Daniella Meeker, PhD
Department of Biomedical Informatics and Data Science, Yale School of Medicine, New Haven CT

Matt Troup, PA-C
Abridge AI., Inc

Timothy D. Barker, MD
Clinical Excellence Division, CHRISTUS® Health
Office of the System Medical Director Chief Medical Information Officer of Ambulatory Care, Christus Health

Vinh Nguyen, MD
Department of Family Medicine
Department of Clinic Informatics, MemorialCare Health System

Jennifer Manders, MD
Department of Surgery, The Christ Hospital Health Network

Cheryl D. Stults, PhD
Center for Health System Research, Sutter Health

Veena G. Jones, MD
Department of Digital Health, Sutter Health
Office of the Chief Medical Information Officer, Sutter Health

Sachin Shah, MD
Departments of Medicine and Pediatrics, University of Chicago
Office of the Chief Medical Information Officer, University of Chicago Medicine and Biological Sciences

Tina Shah, MD MPH
Division of Pulmonary and Critical Care, RWJ Barnabus Health, Newark Beth Israel Medical Center

Lee Schwamm, MD
Department of Bioinformatics and Data Sciences, Yale School of Medicine and Digital and Technology Solutions, Yale New Haven Health System, New Haven CT