Authors:
Laiping Wong-Stewart
Guizhou Hu
Hezhong (Mark) Ma

Case Study

Identifying Key Impairments with Generative AI

RGA

Introduction and purpose

Life insurance underwriting relies on multiple forms of medical evidence to assess an applicant's health status. These sources - including medical records, electronic health records (EHRs), lab test histories, and prescription data - offer valuable but often incomplete views of an individual's health profile. As underwriting shifts toward more digital-first processes, interest is growing in using AI-driven tools to organize, structure, and analyze this data more efficiently.

This paper presents a visibility study evaluating DigitalOwl, a generative AI tool developed to extract, standardize, and tag mentions of health impairments within common underwriting evidence sources and organize them into a structured format. Mentions are tagged based on key impairment categories, then assigned severity levels (high, medium, low, none).

Importantly, this is not a study of risk prediction, pricing accuracy, or mortality assessment. The study does not determine the clinical correctness or financial impact of its findings. Instead, it provides an exploratory assessment of what DigitalOwl can extract and structure from diverse underwriting evidences.

By clarifying the distinction between visibility and risk assessment, this paper aims to demonstrate how structured impairment output can support use cases such as triage development, digital evidence layering, and workflow optimization.

This study is not:

- A predictive modeling or mortality study
- A risk classification or underwriting judgement of accuracy
- A clinical or underwriting validation of data accuracy

This study is:

- A structured extraction and visibility analysis
- A demonstration of how impairments are tagged across evidence types
- A tool-focused perspective for exploring digital triage and automation potential

Methodology

We analyzed approximately 2,000 anonymized life insurance underwriting cases. Each file contained varying combinations of underwriting evidence sources. RGA's internal natural language processing (NLP) engine was used to segment documents by evidence type. They were then submitted to the DigitalOwl tool, which extracted and structured 40 predefined key impairments from the files.

Identification of key impairments could be broad. For example, cardiovascular disease (CVD) mentions included blood pressure tests, heart imaging, etc. Mentions with a severity rating of "none" and ambiguous medication references were excluded. Subsequently, we constructed multiple pairwise comparisons to evaluate the effectiveness of each type of underwriting evidence in identifying a key impairment using the DigitalOwl tool.

Exhibit 1A: Evidence types studied

Life insurance applications	ExamOne LabPiQture®
Milliman IntelliScript IRIX medical data (Dx)	Electronic health records (EHR)
Attending physician statements (APS)	Insurance labs

Excluded: Non-medical documents such as motor vehicle records and financial statements

Exhibit 1B: Sample size and page count

Evidence type	Case count	Page count
Life insurance applications	1,869	35,503
Medical data (Dx)	337	5,802
ExamOne LabPiQture®	387	3,527
Electronic health records (EHR)	220	6,520
Attending physician statements (APS)	1,446	200,524
Insurance labs	1,454	5,498

Limitations and considerations

- Impairments are extracted using AI tagging logic and are not adjudicated by underwriter review.
- Results do not reflect clinical accuracy, severity, or time sensitivity.
- APS and EHR documents may overlap in content, influencing comparative output.
- LabPiQture® tags reflect physician ordering behavior, not result interpretation.
- APS/EHR cohorts likely include more complex applicants, which may introduce sampling bias.
- Detection of impairments does not imply underwriters will have sufficient information to assess severity.

Analysis and findings

Exhibit 2 presents a pairwise comparison of key impairments identified across two separate runs. The first run included applications only, while the second run included applications and Dx. The analysis focused on the cases having data from both runs. Key impairments may come from either the application disclosures or from data identified in medical billing histories.

The first column in Exhibit 2 shows the impairment prevalence based solely on application disclosures. The second column shows the total prevalence from the second run - capturing impairments identified through both the application and Dx data. The third column is the difference between the first two, which is the prevalence identified from Dx data only.

Exhibit 2: Application vs. application + Dx

Key impairment	Application	Application + DX	DX only
Cardiovascular disease	17%	42%	25%
Malignant neoplasm	6%	29%	23%
Diabetes mellitus	7%	9%	2%
Tobacco usage	5%	8%	3%
Psychiatric and mood disorders	22%	31%	9%
Kidney disease	3%	11%	8%
Dyslipidemia	12%	21%	9%
Overweight	12%	18%	6%
Marijuana usage	2%	2%	0%
Alcohol usage	12%	16%	4%

Cohort = 337 cases with available application and Dx data

For example, 17% (57 out of 337) disclosed information related to CVD in their applications. When using application and medical billing data, DigitalOwl identified 42% (141 out of 337) as having CVD. Comparing those two numbers, we see that 25% (84 out of 337) did not disclose CVD in their applications, yet their medical billing histories contained CVD-related information. Note: The DigitalOwl tool defined a key impairment broadly, allowing for all three mortality severity levels (e.g., CVD includes hypertension and hyperlipidemia).

Key insight

Dx history contributed additional structured tags for several impairments not disclosed on applications, especially chronic and behavioral conditions.

Note: Results reflect billing code extraction, not diagnosis verification.

Exhibit 3: Application vs. application + LabPiQture®

Key impairment	Application	Application + LabPiQture®	DX only
Cardiovascular disease	13%	17%	4%
Malignant neoplasm	4%	18%	14%
Diabetes mellitus	9%	12%	3%
Tobacco usage	7%	7%	0%
Psychiatric and mood disorders	19%	21%	2%
Kidney disease	2%	6%	4%
Dyslipidemia	7%	20%	13%
Overweight	5%	10%	5%
Marijuana usage	3%	3%	0%
Alcohol usage	12%	13%	1%

Cohort = 387 cases with available application and LabPiQture® data

Key insight

LabPiQture® offered structured tags for metabolic impairments based on test order codes.

Note: The DigitalOwl tool focused on the presence of specific lab tests as an indication of key impairments rather than analyzing the abnormal range of the results. The test order does not confirm diagnosis, but it may reflect physician concern.

Exhibit 4: Application vs. application + EHR

Key impairment	Application	Application + EHR	DX only
Cardiovascular disease	13%	36%	23%
Malignant neoplasm	3%	23%	20%
Diabetes mellitus	6%	8%	2%
Tobacco usage	4%	10%	6%
Psychiatric and mood disorders	23%	34%	11%
Kidney disease	2%	11%	9%
Dyslipidemia	2%	14%	12%
Overweight	1%	15%	14%
Marijuana usage	1%	3%	2%
Alcohol usage	10%	16%	6%

Cohort = 176 cases with available application and EHR data

Key insight

The DigitalOwl tool found EHR data contributed additional mentions for several key impairments beyond the application-only findings. Notable increases were observed in CVD (+23%), overweight (+14%), and dyslipidemia (+12%), highlighting the tool's ability to expand visibility into both clinical and lifestyle-related risks. However, it is important to remember that while the tool identifies patterns and recognizes key impairments based on what is written in the EHRs, different doctors and systems may document findings differently. Hence, while our results show DigitalOwl can help surface more information, they do not necessarily suggest that EHR can capture more key impairments than the application or that findings are clinically verified.

Note: These results reflect DigitalOwl's ability to extract key impairment mentions based on EHR documentation. They also illustrate the potential of AI-based tools to complement traditionally available evidence. Mentions may vary based on the provider's documentation. Importantly, these results do not imply diagnostic accuracy or source superiority.

Exhibit 5: APS vs. EHR

Key impairment	APS	EHR	Overlap
Cardiovascular disease	64%	29%	25%
Malignant neoplasm	49%	24%	17%
Diabetes mellitus	16%	4%	4%
Tobacco usage	32%	5%	4%
Psychiatric and mood disorders	64%	25%	20%
Kidney disease	24%	9%	8%
Dyslipidemia	36%	12%	9%
Overweight	30%	11%	8%
Marijuana usage	9%	2%	1%
Alcohol usage	58%	8%	6%

Cohort = 186 cases with available APS and EHR data

Key insight

APSs produced more total key impairment tags, though EHRs contributed independent mentions for several behavioral and lifestyle impairments. Exhibit 5 reflects what the DigitalOwl tool was able to extract from EHR and APS inputs, not the intrinsic value of EHR data. Cases with both an APS and an EHR likely represent more complex medical histories. An APS may have been ordered because the EHR was incomplete, inconclusive, or unavailable.

Note: During the development of the NLP solution, some EHR information may also appear in APS files, complicating attribution.

Exhibit 6: Application, Dx, LabPiQture®, or EHR vs. APS or insurance lab

Key impairment	Application, Dx, LabPiQture®, or EHR	APS or insurance lab	Overlap
Cardiovascular disease	23%	50%	18%
Malignant neoplasm	15%	46%	12%
Diabetes mellitus	10%	19%	7%
Tobacco usage	8%	25%	5%
Psychiatric and mood disorders	25%	51%	20%
Kidney disease	6%	20%	4%
Dyslipidemia	10%	33%	7%
Overweight	6%	26%	5%
Marijuana usage	3%	5%	1%
Alcohol usage	25%	45%	16%

Cohort = 1,843 cases with at least one digital and one traditional data source

Key insight

Traditional sources (i.e., APS and insurance lab results) identify more impairments overall, but digital sources offer early, scalable detection and high overlap in key categories.

Note: This supports having a layered approach that uses both traditional and digital sources.

Conclusion

This study demonstrates that GenAI, and in this case, DigitalOwl specifically, can play a meaningful role in extracting and structuring health-related information from diverse forms of underwriting evidence. However, the DigitalOwl tool is not intended to replace clinical judgment or underwriting expertise; rather, it serves as an example of how GenAI can support scalable document review and impairment identification across large case volumes. By producing consistent, structured outputs, the tool may enable more efficient evidence workflows in life insurance underwriting. Potential applications include:

- Enhancing the design of triage and intake models
- Informing the prioritization and sequencing of evidence review
- Supporting the development of digital rules engines
- Facilitating consistent, high-volume screening of applicant files

While these results are exploratory, they suggest a foundation for broader use of AI-assisted extraction tools in underwriting research.

Next steps

As insurers continue to adopt digital-first approaches, AI tools promise to help standardize and scale evidence interpretation. Ongoing collaboration among underwriters, actuaries, and data scientists will be critical in guiding the responsible evolution of these capabilities. If your team is exploring digital evidence strategies, impairment visibility tools, or underwriting automation, we welcome the opportunity to collaborate. Contact RGA.

Meet the Authors

Laiping Wong-Stewart

Laiping Wong-Steward, FSA, MAAA is the Vice President and Actuary of the Strategic Underwriting Initiatives unit within RGA's US Individual Life division. She joined RGA in 2010 and has over 20 years of experience in the life insurance industry including Pricing, Valuation, and Experience Studies. Her current role involves performing data analytics on digital underwriting evidence, especially in the rapidly changing AI-assisted underwriting environment and utilizing the acquired insights to estimate the potential mortality impact.

Guizhou Hu

Guizhou Hu, M.D., PhD, is Vice President and Head of Risk Analytics in RGA's Global Underwriting team. Guizhou leverages his medical training and long experience in epidemiological research and predictive analytics in healthcare and insurance medicine to support RGA underwriting initiatives. He provides direct project support and review, and also produces internal and external thought leadership materials.

Hezhong (Mark) Ma

Hezhong (Mark) Ma is Vice President and Managing Actuary in the Strategic Underwriting Initiatives unit within RGA's US Individual Life division. He leads the actuarial solutions for US Mortality Market's Strategic Underwriting Initiatives. Prior to his current role, Mark took on a variety of responsibilities including head of US experience and data analytics, to actuarial lead of enterprise valuation transformation project, and to the valuation of long term care business.