Health Sciences Informatics Leads and Empowers the Digital Health Transformation
J. Mantas et al. (Eds.)

doi:10.3233/SHTI260794

The Anatomy of a Failed DiGA: Taxonomic Insights from Withdrawn Reimbursable Digital Therapeutics

Benjamin KINAST[a,1], Björn SCHREIWEIS[a] and Sascha Noel WEIMAR[b]
[a] *Institute for Medical Informatics and Artificial Intelligence, Kiel University and University Hospital Schleswig-Holstein, Campus Kiel, Germany*
[b] *Institute for Entrepreneurship, Technology Management and Innovation (EnTechnon), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany*
ORCiD ID: Benjamin Kinast https://orcid.org/0000-0003-2554-4381, Björn Schreiweis https://orcid.org/0000-0002-1748-1563, Sascha Noel Weimar https://orcid.org/0009-0004-7912-5849

Abstract. Germany's Digital Health Application (DiGA) pathway offers a fast-track route to reimbursement for digital therapeutics, yet a share of initially listed applications are subsequently withdrawn from the directory. **Introduction:** This study examines which software features are associated with sustained listing versus delisting, addressing a gap in the literature on unsuccessful DiGA trajectories. **Methods:** We compared 35 permanently listed and 17 delisted DiGAs (n = 52) using qualitative feature coding following the Gioia methodology and the framework of Weimar et al. (2025). Feature prevalence at the 2nd Order theme and Aggregate Dimension levels was compared using Fisher's Exact Test. **Results:** Overall feature breadth was identical across groups. Two themes were significantly more prevalent among delisted DiGAs: Reminders and Notifications and Reports and User Data Export. No significant differences were found at the Aggregate Dimension level or across indication groups. **Discussion:** The findings suggest that continued reimbursement is associated with feature composition rather than quantity. Delisted DiGAs more often incorporated engagement- and infrastructure-oriented features, whereas listed DiGAs more often included components closer to the therapeutic mechanism. **Conclusion:** Feature composition, rather than breadth, may differentiate sustained from withdrawn DiGAs, offering preliminary orientation for developers navigating the pathway and motivation for a more granular taxonomic framework in future research.

Keywords. Digital therapeutics, Digital Health Application, mHealth

1. Introduction

Digital therapeutics (DTx) are anticipated to influence health care delivery by providing evidence-based interventions for chronic and difficult-to-treat conditions [1]. To realize this potential, DTx must demonstrate clinical effectiveness while also being accessible to patients. Germany has sought to advance this dual objective through the digital health

[1] Corresponding Author: Benjamin Kinast, benjamin.kinast@uksh.de.

application pathway (DiGA, German: Digitale Gesundheitsanwendung), a fast-track reimbursement scheme managed by the Federal Institute for Drugs and Medical Devices (BfArM) [2]. To achieve permanent listing as DiGAs, solutions must demonstrate positive healthcare effects and meet requirements, including data security standards [3]. Once listed, they are reimbursable by statutory health insurers, providing developers with a structured path from market entry to sustained coverage.

However, not all listed DiGAs achieve permanent status. Some are removed from the directory after initial listing, most commonly due to insufficient evidence of a positive healthcare effect or at the manufacturer's request [2]. Although the DiGA model has attracted substantial interest regarding evidence generation, pricing, and regulatory pathway design, the characteristics associated with withdrawal remain underexplored. Existing studies tend to focus on general DTx components [4], the evolution of the DiGA pathway [2], the robustness of positive health care effects [5], or features of permanently listed DiGAs [6], leaving the question of what differentiates successful from unsuccessful DiGA trajectories largely unaddressed. Similar access frameworks have emerged in other countries, such as the PECAN in France, while several other countries reimburse digital solutions as part of broader pathways [7], though systematic comparison of withdrawal patterns across these systems remains absent.

This study addresses that gap through a systematic comparison of listed and delisted DiGAs. Drawing on a qualitative feature coding of all applications in the BfArM (as of March 2026), we ask: Which software features are associated with the withdrawal of digital health applications from the DiGA directory?

2. Methodology

We adopted a comparative research design using the BfArM's DiGA directory [8] as the empirical basis. At the time of the study, the sample comprised 35 permanently listed and 17 delisted DiGAs, enabling systematic comparison between applications that achieved sustained reimbursement status and those that did not. Data sources included official instructions for use and app store materials from the BfArM directory [8].

To examine the role of software components in listing outcomes, we applied the feature framework of Weimar et al. as the basis for qualitative coding [6]. Materials were coded independently by two authors until consensus was reached, with coding disagreements resolved through iterative discussion rounds and a subset of items double-coded to verify agreement prior to final coding, following the Gioia methodology [9], resulting in a three-level coding hierarchy: 1st Order Concepts, 2nd Order Themes, and Aggregated Dimensions. Beyond software features, each DiGA was assigned to one of four International Classification of Diseases (ICD) chapter groups to assess whether withdrawal patterns vary by indication area.

All analyses were conducted at the level of unique DiGA rather than individual coding rows, with feature prevalence calculated as the proportion of DiGAs in each group exhibiting a given category. The 1st Order Concept level was treated descriptively only due to high coding granularity (n > near-unique categories). Inferential testing was applied at the 2nd Order Theme and Aggregate Dimension levels, as well as across ICD groups. For each category, Fisher's Exact Test was applied to a 2x2 contingency table (feature present/absent x listed/delisted), appropriate for the small sample size. Significance thresholds were set at $p < 0.05$, $p < 0.01$, and $p < 0.001$. Effect sizes are reported as odds ratios (OR) with 95% confidence intervals. No correction for multiple

comparisons was applied, given the exploratory nature. Mean feature counts per DiGA were additionally calculated to assess overall feature breadth independent of specific content. All analyses were implemented in Python (pandas, scipy, numpy).

3. Results

The sample comprised n = 52 DiGAs, of which 35 were permanently listed, and 17 had been delisted. The mean total number of coded features per DiGA was identical across groups (listed: 15.7; delisted: 15.7). Permanently listed DiGAs showed higher counts at the 1st Order Concept level (mean 6.6 vs 4.6), whereas delisted DiGAs showed higher counts at the 2nd Order Theme level (mean 6.9 vs. 5.6) and Aggregate Dimension level (mean 4.2 vs. 3.5). No significant differences in then listing status were found across the four ICD chapter groups (**Table** ; all $p > 0.07$). Mental & Behavioral Disorders were more prevalent among permanently listed DiGAs (51.4% vs. 23.5%, Δ +27.9%).

Table 1. Disease group distribution among listed and delisted digital health applications

Disease Group	% Listed	% Delisted	Δ % (L – D)	p-Value	OR
Mental & Behavioral Disorders	51.4%	23.5%	+27.9	0,076	3,44
Musculoskeletal, Neurological & Sensory Disorders	22.9%	29.4%	-6.5%	0,735	0,71
Metabolic, Nutritional & Digestive Disorders	14.3%	23.5%	-9.2%	0,451	0,54
Cardiovascular, Oncological & Urogenital Disorders	11.4%	23.5%	-12.1%	0,413	0,42

At the Aggregate Dimension level (Table 2), no significant differences were observed between groups (all $p > 0.14$). *Health Interventions* and *Data Collection and Monitoring* were slightly more prevalent in listed DiGAs, but differences did not reach statistical significance (Δ +6.1%, $p = 0.589$; Δ +3.3%, $p = 1.000$). *User Preferences and Data Security* showed the largest numerical difference (Δ -21.0%, p = 0.145), and *Communication and Information Exchange* showed the second largest (Δ -19.5%, $p = 0.208$), both more prevalent among delisted DiGAs, though neither reached statistical significance.

Table 2. Distribution of features on the aggregate dimension level for listed and delisted digital health applications.

No.	Aggregate Dimension	% Listed	% Delisted	Δ % (L – D)	p-Value	OR
1	Health Interventions	94.3	88.2	+6.1	0,589	2,2
2	Data Collection and Monitoring	85.7	82.4	+3.3	1,000	1,29
3	Resource Management	17.1	17.6	-0.5	1,000	0,97
4	User Engagement and Incentives	77.1	88.2	-11.1	0,467	0,45
5	Communication and Information Exchange	62.9	82.4	-19.5	0,208	0,36
6	User Preferences and Data Security	14.3	35.3	-21.0	0,145	0,31

At the 2nd Order Theme level (Table 3), two features were significantly more prevalent among delisted DiGAs. *Reminders and Notifications* were present in 70.6% of delisted DiGAs compared to 22.9% of listed DiGAs (Δ -47.7%, $p = 0.002$), and *Reports and User Data Export* appeared in 58.8% of delisted versus 20.0% of listed applications (Δ -38.8%, $p = 0.010$). Additional non-significant trends toward higher prevalence in delisted DiGAs included *Sensor-Driven Structured Data Tracking* (Δ -21.2%, $p = 0.181$), *Progress Overview and Data Visualization* (Δ -19.2%, $p = 0.240$), and *Goal*

Management and Rewards (Δ -18.5%, $p = 0.224$). In the opposite direction, *Structured Exercise Programs* (Δ +24.1%, $p = 0.089$) were more common among listed DiGA, though none were statistically significant.

Table 3. Distribution of features on the 2nd-order theme level among listed and delisted digital health applications.

No.	2nd Order Themes	% Listed	% Delisted	Δ % (L – D)	p-Value	OR
1	Structured Exercise Programs	82.9	58.8	+24.1	0,089	3,38
2	User-Reported Unstructured Data Tracking	40.0	17.6	+22.4	0,128	3,11
5	Communication with Health Care Providers	40.0	17.6	+22.4	0,128	3,11
4	Activity Feeds and Exercise Overviews	31.4	17.6	+13.8	0,341	2,14
1	Interactive Health Interventions	28.6	17.6	+11.0	0,506	1,87
3	Non-Customizable Content Library	11.4	5.9	+5.5	1,000	2,06
1	Health Education Elements	60.0	58.8	+1.2	1,000	1,05
3	Filterable Content Library	8.6	11.8	-3.2	1,000	0,7
5	Community Building and Social Engagement	5.7	11.8	-6.1	0,589	0,45
5	Communication with Customer Support	11.4	17.6	-6.2	0,670	0,6
4	Usage Instructions and Tips	11.4	17.6	-6.2	0,670	0,6
6	Settings and Profile	8.6	17.6	-9.0	0,379	0,44
2	User-Reported Structured Data Tracking	65.7	76.5	-10.8	0,532	0,59
6	Data Security Measures	8.6	23.5	-14.9	0,198	0,3
4	Goal Management and Rewards	28.6	47.1	-18.5	0,224	0,45
4	Progress Overview and Data Visualization	51.4	70.6	-19.2	0,240	0,44
2	Sensor-Driven Structured Data Tracking	20.0	41.2	-21.2	0,181	0,36
5	Reports and User Data Export	20.0	58.8	-38.8	0,010*	0,17
4	Reminders and Notifications	22.9	70.6	-47.7	0,002**	0,12

4. Discussion

Our findings suggest that listing and delisting in the DiGA pathway differ by feature composition rather than feature quantity. Listed and delisted applications had nearly the same mean number of coded features, but differed in feature type: delisted DiGAs more often included engagement, notification, and data management features, whereas permanently listed DiGAs more often offered structured therapeutic programs and direct communication with healthcare providers. Effect sizes were large, with delisted DiGAs showing approximately eight-fold higher odds for Reminders and Notifications and six-fold higher odds for Reports and User Data Export. This pattern was consistent across coding levels and indication groups, suggesting a cross-indication phenomenon and a possible distinction between features tied to therapeutic mechanisms and those serving infrastructural or engagement functions. However, these findings are exploratory and correlational. Whether feature composition is causally related to listing outcomes, or instead reflects factors such as clinical evidence, regulatory requirements, financial constraints, or voluntary withdrawal, cannot be determined.

Mental and behavioral disorders were numerically more common among permanently listed DiGAs, though not significantly. Interpretation is further limited by the small sample, lack of multiple-comparison correction, reliance on public materials, and potential coder bias.

5. Conclusion

This study examined software feature differences between permanently listed and delisted DiGAs in the German reimbursement pathway. The findings suggest that feature composition, rather than quantity, distinguishes the two groups: delisted DiGAs more often included engagement- and infrastructure-oriented features, whereas listed DiGAs more often featured components closer to therapeutic action.

These results contribute empirically to understanding failed DiGAs by showing, at the feature level, what distinguishes applications that sustain reimbursement from those that do not. For developers, the findings offer preliminary orientation, suggesting that compositional design choices may matter beyond the number of features offered. The results and taxonomy applied may furthermore serve as a structured input for health technology assessment bodies. Future research should replicate and strengthen these findings as the DiGA directory grows.

Acknowledgements

The authors used the generative artificial intelligence tools Claude (Sonnet 4.6, Anthropic) and ChatGPT (GPT-5, OpenAI) to support corrections in grammar, style, and spelling.

References

[1] Dang A, Arora D, Rane P. Role of digital therapeutics and the changing future of healthcare. Journal of Family Medicine and Primary Care 2020;9(5):2207, doi: 10.4103/jfmpc.jfmpc_105_20

[2] Schmidt L, Pawlitzki M, Renard BY, Meuth SG, Masanneck L. The three-year evolution of Germany's Digital Therapeutics reimbursement program and its path forward. npj Digital Medicine. 2024;7:139. doi:10.1038/s41746-024-01137-1.

[3] Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM). Das Fast-Track-Verfahren für digitale Gesundheitsanwendungen (DiGA) nach § 139e SGB V: Ein Leitfaden für Hersteller, Leistungserbringer und Anwender. Version 3.6. 2025 Dec 10. Available from: https://www.bfarm.de/SharedDocs/Downloads/DE/Medizinprodukte/diga_leitfaden.pdf?__blob=publicationFile. Accessed 2026 Mar 25.

[4] Recchia G, Gussoni G. Digital therapeutics: scientific, technological, and regulatory challenges. In: Cesario A, D'Oria M, Auffray C, Scambia G, editors. Personalized Medicine Meets Artificial Intelligence. Cham: Springer International Publishing; 2023. p. 37-50. doi: 10.1007/978-3-031-32614-1_4.

[5] Sippli K, Deckert S, Schmitt J, Scheibe M. Healthcare effects and evidence robustness of reimbursable digital health applications in Germany: a systematic review. npj Digital Medicine. 2025;8:495. doi:10.1038/s41746-025-01879-6.

[6] Weimar SN, Minaschek H, Martjan RS, Terzidis O. Decoding Digital Therapeutics: A Qualitative Analysis of Software Features in German Digital Health Applications. In: Andrikopoulou E, et al., editors. Intelligent Health Systems – From Technology to Data and Knowledge. 2025, doi: 10.3233/SHTI250507

[7] Van Kessel R, Srivastava D, Kyriopoulos I, Monti G, Novillo-Ortiz D, Milman R, Zhang-Czabanowski WW, Nasi G, Stern AD, Wharton G, Mossialos E. Digital Health Reimbursement Strategies of 8 European Countries and Israel: Scoping Review and Policy Mapping. JMIR Mhealth Uhealth. 2023;11:e49003. doi:10.2196/49003.

[8] Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM). DiGA-Verzeichnis. Available from: https://diga.bfarm.de/de/verzeichnis. Accessed 2026 Feb.

[9] Gioia DA, Corley KG, Hamilton AL. Seeking qualitative rigor in inductive research: notes on the Gioia methodology. Organizational Research Methods. 2013 Jan;16(1):15-31. doi:10.1177/1094428112452151.