

STELLUNGNAHME Deutsches Netzwerk Versorgungsforschung zum Referentenentwurf der 2. DiGAV ÄndV vom 13.12.2024 erarbeitet durch die DNVF AG Digital Health

Die **AG Digital Health** im Deutschen Netzwerk Versorgungsforschung (DNVF) hat für das DNVF, vertreten durch folgende Mitglieder, den Referentenentwurf zur 2. DiGAV ÄndV vom 13.12.2024 in der folgenden Stellungnahme kommentiert:

Theresa Sophie Busse, Jennifer Dell, Nikolas Dietzel, Mareike Geisler, Ria Heinrich, Sabrina Hölzer, Ekkehart Jenetzky, Ursula Kramer, Melanie Mäder, Jutta Richter, Madlen Scheibe, Madlen Sixtensson, Susanne Singer, Patrick Timpel, Michael Zeiler

Ziel der Versorgungsforscher:innen ist es, Impulse für die Ausgestaltung der anwendungsbegleitenden Erfolgsmessung gemäß **§ 139e SGB V**, im folgenden **abEM DiGA** genannt, zu liefern, um neu gewonnene Daten aus der Versorgungsrealität mit digitalen Gesundheitsanwendung zukünftig auch als Forschungs-Ressourcen nutzen zu können und die Implementierung von digitalen Therapien zu fördern.

1. Zusammenfassung

Die im Gesetz (2. DiGAV ÄndV) definierten Ziele, einen Qualitätswettbewerb zwischen den **dauerhaft gelisteten DiGAs** zu ermöglichen und die Grundlage zu schaffen für eine datenbasierte, erfolgsorientierte Vergütung, ist mit dem methodischen Vorgehen **nicht erreichbar**. Therapeutischer Erfolg und medizinischer Nutzen digitaler Versorgungsangebote werden mit den so erhobenen **Daten für die GKV weder mess- noch sichtbar**. Die Daten erlauben allenfalls ein **orientierendes Benchmarking** im Hinblick auf die **Nutzung, nicht jedoch auf den Nutzungserfolg bzw. den Therapieerfolg**. Insofern wird die **Transparenz über die tatsächliche Nutzung und Qualität** digitaler Versorgungsangebote weder für die GKV-Versicherten noch für die Leistungserbringenden mit diesem Ansatz **erhöht**.

Die AG Digital Health im DNVF präferiert ein methodisches Vorgehen, das den Aufwand der Erhebung neuer Daten rechtfertigt und einen **echten Qualitätswettbewerb** erlaubt. Dazu braucht es unverzerrte, unabhängig erhobene, aussagekräftige Daten aus der Versorgungsrealität, mit denen sich qualitätsbestimmende Faktoren einer DiGA identifizieren lassen, sowohl an der Schnittstelle zu Nutzer:innen als auch zu Therapeut:innen. Die Frage, was den **Therapieerfolg** einer DiGA und deren Stellenwert als ergänzenden Therapiebaustein im Rahmen einer evidenzbasierten Patient:innenversorgung beeinflusst, wird durch die Methodik der vorgeschlagenen abEM DiGA nicht beantwortet.

2. Begründung

2.1 Erfolgsmessung

Im Referentenentwurf fehlt eine klare Definition, was unter dem Erfolg zu verstehen ist, der mit dem skizzierten methodischen Vorgehen gemessen werden soll. Die derzeitige Ausgestaltung fokussiert ausschließlich auf **Nutzungserfolg**.

Erfolg im Sinne eines therapeutischen Nutzens kann über die erhobenen Daten der anwendungsbegleitenden Erfolgsmessung nicht abgebildet werden, da **selektierte Patient:innengruppen ohne Komparator** betrachtet werden. Die Gegenüberstellung der einzelnen DiGAs auf Basis der abEM-Daten kommt einem naiven indirekten Vergleich nahe, der **hoch verzerrt** ist und dessen **Aussagesicherheit sehr gering** ist. Aus Sicht der Versorgungsforschung raten wir dringend davon ab. Die Messung von Nutzungsdaten als Grundlage für (Therapie-)Erfolg zu interpretieren, kann auf Fehlannahmen beruhen und somit zu Fehlschlüssen führen: Nicht-(Mehr-)Nutzung kann auch ein Zeichen gesteigerter Selbstwirksamkeit oder verbesserter Selbstmanagementkompetenzen darstellen, kann aber auch auf eine fehlende Wirksamkeit der DiGA oder subjektiv ausreichende Wirksamkeit nach verkürzter Nutzungszeit hinweisen, was sich mit dem methodischen Vorgehen nicht voneinander differenzieren lässt.

2.2 Preisgestaltung

Die Daten der abEM DiGA sollen die Basis liefern für eine erfolgsabhängige Preisgestaltung, indem sie vergleichende Ergebnisse generieren

1. zu Nutzungsumfang, -häufigkeit und -abbruch (Stufe I)
2. zum patient:innenberichteten Gesundheitszustand (indikationsübergreifend) und zur Patient:innenzufriedenheit in Bezug auf die Qualität der DiGA (Stufe II)
3. zum patient:innenberichteten Gesundheitszustand (indikationsbezogen), Stufe III

Im vorliegenden Referentenentwurf finden sich keinerlei Hinweise, wann und wie die in drei Stufen generierten Ergebnisse der abEM DiGA zukünftig in die erfolgsabhängige Preisbildung einfließen sollen.

2.3 Qualitätswettbewerb

Ein Qualitätswettbewerb der DiGAs untereinander, der sehr zu begrüßen ist, lässt sich mit dem vorgeschlagenen Vorgehen nicht erreichen. **Der Vergleich der Nutzung der DiGAs untereinander** erlaubt lediglich ein grobes Benchmarking und keine vergleichende Differenzierung im Hinblick auf Erfolg und (Therapie)nutzen der DiGAs untereinander. Den etablierten Methoden der evidenzbasierten Medizin hält der vorgeschlagene methodische Ansatz **nicht stand**.

1. Es gibt **keine Vergleichsgruppe bzw. keinen Vergleich** zu einer anderen (wirksamen) Therapie. Daher lassen sich die Veränderungen des Gesundheitszustandes nicht

abgrenzen von Veränderungen, die auf den natürlichen Krankheitsverlauf zurückgehen, oder anderen externen, relevanten Einflussfaktoren geschuldet sind, wie z. B. zusätzliche Therapien/Verhaltensänderungen.

2. Es kommt zu einer **Selektion der Personen, da diese freiwillig** Angaben zum Gesundheitszustand machen, d. h. nicht für alle Nutzenden der DiGA werden Daten zum Gesundheitszustand generiert. Auch Abbrecher, die die Nutzung aufgrund eines Therapieerfolges eingestellt haben, liefern keine Ergebnisse zum Gesundheitszustand. Diese Selektion macht es unmöglich, generalisierbare Schlüsse für die Zielpopulation der Personen zu ziehen, die einen Freischaltcode eingelöst haben.

3. **Der Gesundheitszustand wird** über den PGI-C Fragebogen (1) retrospektiv gemessen. Der damit verbundene Recall Bias verzerrt die Daten. Wir präferieren eine **prospektive Erfassung des Gesundheitszustandes** mit dem **PGI-S Fragebogen** (2), und einen **prä-post-Vergleich**, um durch eine Nutzendenbefragung zu Beginn und am Ende des Nutzungszeitraumes **Veränderungen erfassen** zu können.

2.4 Reporting der Daten

Reporting durch die Hersteller

Neben der Veröffentlichung der von den Herstellern bereitgestellten Ergebnisse aus der **abEM DiGA** muss es eine **objektive Bewertung und Kontextualisierung** geben, bestenfalls **durch das BfArM**. Ein Bericht, der ausschließlich die durch die Hersteller gelieferten Daten umfasst, ist wenig hilfreich.

Es besteht grundsätzlich die Gefahr, dass die durch die **Hersteller gesammelten Daten** dem Ziel eines Qualitätswettbewerbs der DiGA entgegenstehen und **nicht die gewünschte, unabhängige Entscheidungsbasis liefern**, um die Wahl der DiGA für Patient:innen und Verschreibende durch mehr Transparenz zu unterstützen.

Vorhandene Daten besser nutzen

Bevor zusätzliche, neue Daten aus der Versorgung erhoben werden, empfehlen wir die **bereits vorhandenen Daten** zur klinischen Evidenz aus Listungsstudien besser zu nutzen und Unterschiede in deren **methodischer Qualität** durch unabhängige Organisationen, wie dem BfArM, transparenter darzustellen und dabei etablierte **Reporting- und Bewertungsstandards der evidenzbasierten Medizin** zu nutzen (3).

Auch die Daten zur empfohlenen **Nutzungshäufigkeit** und **Anwendungsdauer** aus Listungsstudien **sollten für eine passfähige Nutzung der DiGA** im DiGA-Verzeichnis verfügbar sein und transparent dargestellt werden.

Die Herausforderung, die "Minimaldosis" einer DiGA im Rahmen der CE-Kennzeichnung zu definieren und als Minimaldosis in den Zulassungsstudien zugrunde zu legen, ist bekannt.

Definition der „Nutzer:innen“, die in die Auswertung der abEM DiGA einfließen

Die Definition der Nutzer:innen, die in die Erhebung und geplante Auswertung einbezogen werden, ist von großer Bedeutung. Als Nutzende sollten in diesem Kontext **alle Personen** gelten, die den **Freischaltcode zulasten der GKV eingelöst** haben, unabhängig davon, ob sie die DiGA dann auch tatsächlich nach der Einlösung nutzen. Es ist zu vermeiden, lediglich jene Personen als Nutzende zu erfassen, die die DiGA auch wirklich aktiv anwenden, denn so werden Nutzungshäufigkeit und -dauer überschätzt.

Differenzierung der Gründe für einen Nutzungsabbruch

Ein Nutzungsabbruch kann viele Gründe haben und lässt nicht zwingend auf den Misserfolg einer DiGA schließen (s.o.). Abhängig von der Indikation und der Schwere von Symptomen kann sich der Therapieerfolg früher oder später einstellen, sodass eine DiGA-Nutzung ggfs. nicht über den gesamten Zeitraum (z. B. 90 Tage) erforderlich sein könnte. Will man Abbrecher:innen in der abEM qualifizieren, müssen daher die **Gründe des Nutzungsabbruchs** erhoben werden, um dies dann – positiv oder negativ - in die "erfolgsabhängige" Preisgestaltung einfließen lassen zu können. Mit dem im Referentenentwurf vorgeschlagenen methodischen Vorgehen ist das nicht möglich.

Hersteller werden vor die Herausforderung gestellt werden, Wege zu finden, um Abbruchgründe zu erfassen und sich damit von vergleichbaren DiGA abgrenzen zu können.

2.5 Patient:innenbefragung

Neben den Nutzungsdaten, die vom Hersteller in aggregierten und anonymisierten Reports geliefert werden, sieht die **abEM DiGA** in Stufe 2 Nutzendenbefragungen vor, die neben der **Patientenzufriedenheit bezogen auf die Qualität der DiGA** auch den **patientenberichteten Gesundheitszustand** erheben.

Zeitpunkt der Nutzer:innenbefragung

Die vorgeschlagene Erhebung „**während**“ **der Nutzung** ist vermutlich der Tatsache geschuldet, dass eine Befragung der Nutzer:innen innerhalb des Nutzungszeitraums über die DiGA praktisch einfacher realisierbar ist. Nach Abschluss der Intervention oder gar in einem "Follow-up" sind die Personen, die eine DiGA nicht mehr aktiv nutzen, schwer erreichbar.

„Während“ meint in diesem Kontext **7 Tage vor Abschluss des Anwendungszeitraums** von 90 Tagen und ist damit de facto „nach der Intervention“. Hat die DiGA eine empfohlene Anwendungsdauer von 90 Tagen, sollten sich die Veränderungen des Gesundheitszustandes **innerhalb des 90 Tage-Fensters** erfassen lassen, anders ist das bei DiGAs mit einer deutlich längeren Nutzungsdauer von 6 oder 12 Monaten.

Datenerhebung abEM DiGA für alle DiGAs auf max. 90 Tage begrenzen – eine inakzeptable Ungleichbehandlung

Es ist von Interesse, wie viele Personen eine DiGA entsprechend ihrer Anwendungsdauer vollständig anwenden, um diese Daten später in einem Qualitätswettbewerb nutzen und in ein erfolgsabhängiges Preismodell einfließen lassen zu können.

Mehr als drei Viertel aller dauerhaft gelisteten DiGAs deklarieren eine vorgesehene Mindestdauer von 3 Monaten. Einige wenige dauerhaft gelistete DiGAs (4,5) haben eine vorgesehene Mindestnutzungsdauer von 6 bzw. 12 Monaten, einige wenige werden kürzer als 90 Tage genutzt.

Der vorliegende Referentenentwurf sieht vor, die Daten aus der abEM für alle DiGAs auf max. 90 Tage zu beschränken. Dies stellt eine **inakzeptable Ungleichbehandlung dar**, da DiGA mit einer Anwendungsdauer von mehr als 90 Tagen nur einen Teil ihrer relevanten Nutzung belegen müssen und das Interventionsende bei einer Nutzungsdauer von 6 bzw. 12 Monaten nach 90 Tagen noch lange nicht erreicht ist.

Vollständigkeit der Daten

In der Stufe 2 der **abEM DiGA** soll der „Patientenberichtete Gesundheitszustand“ mittels PGI-C-Fragebogen am Ende des Anwendungszeitraums von 90 Tagen retrospektiv erhoben werden. Ohne eine Erhebung zu Beginn der Intervention sind **keine vollständigen Daten der Zielpopulation** zu erwarten. Es wird eine Selektion bzgl. der erhobenen Daten geben, die keine Ableitung von allgemeinen Aussagen zur Zielpopulation ermöglicht (s. 2.3).

Die Wahrscheinlichkeit ist groß, dass **Nicht-Nutzende ein geringeres Interesse haben**, sich als Versicherte an einer Datenerhebung zu beteiligen, die die Patient:innenzufriedenheit in Bezug auf die Qualität der digitalen Gesundheitsanwendung sowie den patient:innenberichteten Gesundheitszustand erfassen will (Stufe 2), insbesondere dann, wenn der **Zugang zur Befragung über die DiGA** selbst erfolgt.

Genutzte Fragebogen

Der Fragebogen PGI-C zur Erfassung des patient:innenberichteten Gesundheitszustandes in Stufe 2 ist als validiertes Messinstrument etabliert. **Gleichwohl genügt es nicht**, den patient:innenberichteten Gesundheitszustand bzw. die Veränderung des Gesundheitszustands **durch eine einzelne Frage und nur an einem Messzeitpunkt adäquat zu erfassen**. Ob Änderungen durch die Anwendung der DiGA bedingt sind, lässt sich grundsätzlich ohne eine Vergleichsgruppe und zufällige (d. h. randomisierte) Allokation zu den beiden Gruppen nicht erfassen (s. 2.3).

Prä- und Post-Erhebung mit PGI-S: Der als Messinstrument vorgeschlagene **PGI-C** Fragebogen liefert aufgrund der retrospektiven Erfassung durch den zu erwartenden **Recall Bias verzerrte Informationen**. Um dieses Problem zu umgehen, schlagen wir vor, den **PGI-S Fragebogen** zu nutzen und durch zwei Erfassungszeitpunkte in einer **Prä- und Post-Erhebung** die Veränderungen zu berechnen. Dieses Vorgehen hat außerdem den Vorteil, dass man insgesamt mehr Daten zur Beschreibung des Gesundheitszustands der Personen erfassen kann, die an der abEM teilnehmen. Denn: Zu Beginn der Anwendung sind **die Responseraten** vermutlich höher, als am Ende der Anwendungszeit. Auf diese Weise hätte man auch von den Personen, die keine Antworten zur Posterhebung liefern, zumindest eine Beschreibung bzgl. ihres Gesundheitszustands aus der Anfangserhebung.

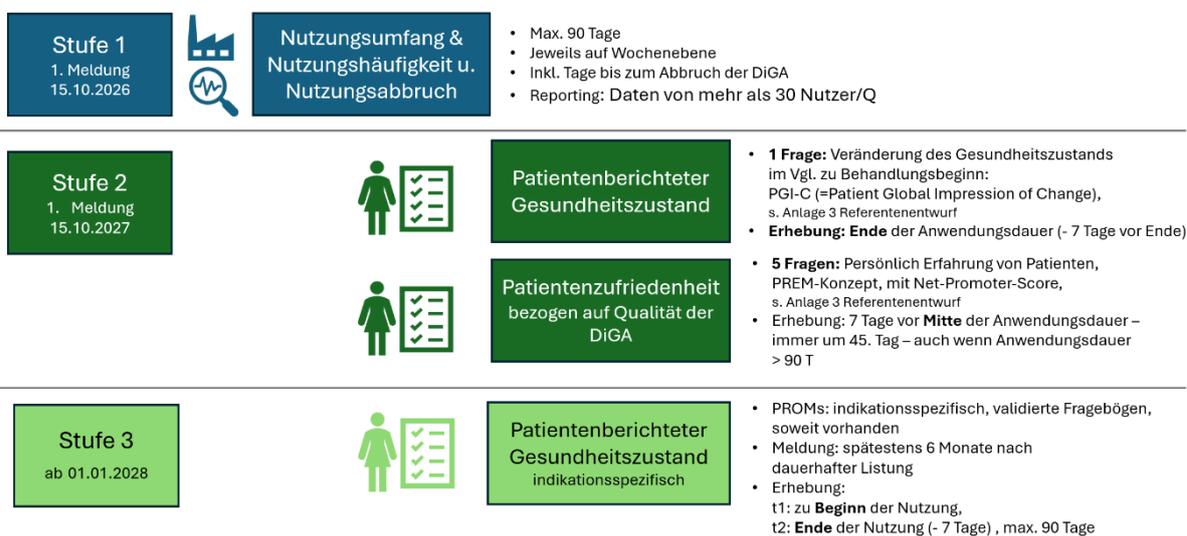
Fragebogen – Datenauswertung: Zur Nutzung sollten weitere Lage- und Streuungsparameter berichtet werden. Es wird mit hoher Wahrscheinlichkeit eine schiefe Verteilung der Nutzungsdaten auftreten, die besser mit Median, Perzentilen (5., 25., 75., 95.) und Minimum und Maximum beschrieben werden kann.

3. Ausblick

Die interdisziplinäre AG Digital Health im DNVF entwickelt derzeit ein Positionspapier, das konkretisiert, wie verfügbare Datenressourcen besser genutzt und neue Real World Daten zu DiGA erhoben werden können, um einen **echten Qualitätswettbewerb unter den DiGA** zu fördern und **valide Erkenntnisse zu Einflussfaktoren** zu generieren, die den Therapieerfolg bzw. Nutzen digitaler Therapien auf Basis der ebM-Prinzipien weiter festigen. Dabei soll auch die Rolle von neuen Registern beleuchtet und auf Basis erster, vorliegender Ansätze (6) diskutiert werden.

abEM DiGA: 3-stufiges Verfahren

Kommentierung DNVF AG Digital Health bis 28.01.2025, Referentenentwurf vom 13.12.2024



Quellen:

1. Nicht Krankheitsspezifischer PROM Fragebogen. PGI-C The patient global impression of change (PGI-C) scale. „Since the start of the intervention, my overall status is...“
https://www.researchgate.net/figure/The-patient-global-impression-of-severity-PGI-S-and-change-PGI-C-scales_fig1_339980664
2. Nicht Krankheitsspezifischer PROM Fragebogen. PGI-S. The patient global impression of severity (PGI-S) scale. „Please rate the severity of your x right now“.
https://www.researchgate.net/figure/The-patient-global-impression-of-severity-PGI-S-and-change-PGI-C-scales_fig1_339980664
3. CONSORT 2010: Aktualisierte Leitlinie für Berichte randomisierter Studien im Parallelgruppen-Design. [Thieme E-Journals - DMW - Deutsche Medizinische Wochenschrift / Abstract](#)
4. Kalmeda. Vorgesehene Mindestdauer der Anwendung der DiGA: 12 Monate
<https://diga.bfarm.de/de/verzeichnis/00350>
5. Zanadio. Als Mindestdauer der Nutzung werden 6 Monate angesetzt, wobei eine Nutzung für 12 Monate dringend empfohlen wird. <https://diga.bfarm.de/de/verzeichnis/00294>
6. Albrecht et al. (2025). Rheumatol Ther. [Real-World-Evidence of Digital Health Applications \(DiGAs\) in Rheumatology: Insights from the DiGAReal Registry](#)