

Alice Ravizza¹, Oriana Ciani², Antonio Ferrari³
Sabrina Grigolo⁴, Francesco Perrone⁵, Elisabetta Ravot⁶
Eugenio Santoro⁷, Federico Sternini¹

Come valutare le terapie digitali da un punto di vista clinico

1. Valutazione del beneficio clinico

I *software* che erogano terapie digitali sono dispositivi medici e pertanto devono rispettare, in Europa, i requisiti del Regolamento 2017/745.

Una differenza rilevante tra la MDD (*Medical Device Directive*) ormai verso la obsolescenza e MDR (*Medical Device Regulation*), la cui piena applicazione è prevista nel maggio 2021, è costituita dalla richiesta di valutazione dei dati di *performance* e di beneficio clinico, che viene esplicitata e descritta in dettaglio. La valutazione dell'efficacia del *software* deve essere basata su dati clinici e su evidenze cliniche. Non è quindi più sufficiente dimostrare che il *software* "funziona", ma è anche necessario documentare che esso ottiene, in modo statisticamente significativo, il beneficio clinico per cui è stato progettato. In molti casi, non sarà possibile raccogliere sufficienti evidenze cliniche in letteratura, e sarà quindi necessario eseguire degli studi clinici adeguati.

I progettisti hanno a disposizione, come riferimento, lo *standard* ISO 14155, che fornisce metodi per la pianificazione ed il monitoraggio dello studio (*Clinical investigation of medical devices for human subjects - Good Clinical Practice*).

¹UseMeD, Torino

²CERGAS, SDA Bocconi School of Management, Milano

³IQVIA Ltd, Durham, NC, United States

⁴Accademia del Paziente Esperto EUPATI Onlus

⁵Unità Sperimentazioni Cliniche, IRCCS Istituto Nazionale Tumori Fondazione Pascale, Napoli

⁶Healthware Group, Milano

⁷Laboratorio di Informatica Medica, Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Milano

Inoltre, è messa a disposizione di fabbricanti ed autorità regolatorie la linea guida MDCG 2020-1 *Guidance on Clinical Evaluation (MDR)/Performance Evaluation (IVDR) of Medical Device Software of March 2020* che fornisce indicazioni dettagliate sugli aspetti di pianificazione della raccolta delle evidenze cliniche e della loro rendicontazione ai fini regolatori.

La linea guida è particolarmente adeguata ad essere applicata a *software* di diagnosi e monitoraggio, il cui beneficio atteso è la messa a disposizione di informazioni accurate. Ciononostante, vi sono utili informazioni anche per la valutazione clinica di DTx, seppure un approfondimento da parte di MDCG sulle modalità specifiche di valutazione clinica delle terapie digitali appaia a nostro avviso auspicabile.

Innanzitutto, viene confermata la definizione di beneficio clinico (*Article 2 (53)*) come l'impatto positivo di un dispositivo sulla salute di un individuo, espresso in termini di esiti clinici significativi misurabili e rilevanti per il paziente, inclusi gli esiti relativi alla diagnosi, o un impatto positivo sulla gestione della malattia o sulla salute pubblica.

Lo stesso documento chiarisce la definizione di evidenze cliniche (*"I dati clinici e la valutazione clinica (MDR)/valutazione delle prestazioni (IVDR) si riferiscono a un dispositivo in quantità e qualità sufficienti per consentire una valutazione qualificata rispetto alla sicurezza del dispositivo e al fatto che lo stesso dispositivo raggiunga i vantaggi clinici previsti quando usato secondo le indicazioni del produttore"*).

Un altro concetto cardine richiamato dalla linea guida MDCG 1 è quello di "valida associazione clinica". Con questo si intende in che misura si possano associare gli *output* del *software* con gli effetti su uno stato fisiologico mirato o su una condizione clinica. Questa associazione deve dimostrare che vi sono robuste relazioni tra i principi scientifici alla base dell'uso del *software* ed il beneficio atteso.

Inoltre viene inserita una importante definizione di "generalizzabilità" (*"La generalizzabilità si riferisce alla capacità di un Medical Device Software/MDSW di estendere le prestazioni previste, testate su una serie specifica di dati, alla popolazione di riferimento più ampia"*). Il concetto di generalizzabilità dovrebbe essere preso in considerazione sia quando viene creato un *database* per la validazione tecnica e pre-clinica del *software*, sia quando vengono definiti i criteri di inclusione ed esclusione degli studi clinici, e più in generale quando si discute l'implementazione di un dispositivo al di fuori del contesto di studio clinico in cui è stato valutato.

Dal punto di vista della pianificazione, nella maggior parte dei casi lo svi-

luppo clinico di una terapia digitale (*Digital Therapeutics -DTx*) include una fase iniziale esplorativa (studi pilota) seguita da una fase di sperimentazione volta a produrre le evidenze di efficacia e sicurezza necessarie per l'approvazione nella specifica indicazione terapeutica (studi *pivotal* di certificazione CE).

Le caratteristiche degli studi pilota e *pivotal* delle DTx tendono a riprodurre gli schemi tipicamente adottati nella sperimentazione di interventi terapeutici tradizionali (es. terapia farmacologica, o terapia comportamentale); tuttavia, la natura stessa dell'intervento DTx può richiedere una revisione o un adattamento rispetto agli schemi *standard*, in particolare per quanto riguarda la scelta degli *endpoint*, gli indicatori di efficacia utilizzati, l'intervento di controllo, e senza escludere la possibilità della creazione/validazione di indicatori di *outcome* (clinici, di *Health Technology Assessment -HTA* etc) specificamente adatti a queste tipologie di prodotto. Fra le principali caratteristiche delle DTx vi è la partecipazione attiva del paziente/*caregiver* nel percorso di cura. Anche sulla scorta di indicazioni che ne hanno documentato il beneficio, la pianificazione degli studi orientati alla produzione di evidenze per le DTx dovrebbe quindi attentamente considerare le dimensioni riguardanti la selezione e la motivazione dei pazienti e dei *caregiver*, e auspicabilmente il coinvolgimento di loro rappresentanti già dalla fase di definizione dei progetti di studio^(1,2).

2. Panoramica degli studi clinici esistenti

Questa sezione intende fornire una panoramica degli studi disegnati per lo sviluppo di alcuni dispositivi medici identificati come *Digital Therapeutics*.

Al fine di avere una analisi coerente sono stati estrapolati diversi aspetti quali la presenza di soluzioni di confronto *gold standard*, la popolazione arruolata, la tipologia di intervento e controllo, e la definizione degli *outcome* primari.

2.1 Prodotti in fase avanzata

In questa sezione sono analizzati gli studi che hanno fornito evidenze cliniche di efficacia per alcuni prodotti in uno stato avanzato di sviluppo e/o già commercializzati.

Di seguito sono elencati gli studi i cui risultati sono utilizzati dai produttori per dimostrare l'efficacia della soluzione, e per i quali sono disponibili pubblicazioni scientifiche su riviste *peer-reviewed* o la registrazione sul sito clinicaltrials.gov.

Tabella 1 - Rappresentazione schematica di studi per la dimostrazione di efficacia di diverse DTx

Prodotto	Esperimento	Pazienti valutati	Condizione o malattia	Intervento	Controllo	Outcome primari
reSET ⁽³⁾	RCT	507	Abuso di sostanze	Terapia <i>standard</i> ridotta	Terapia <i>standard</i>	Riduzione uso sostanze (PRO e urine)
				<i>Therapeutic Education System</i> (terapia comportamentale <i>online</i>)		Tempo prima di cessazione terapia
reSET-O ⁽⁴⁾	RCT	206	Dipendenza da oppiacei	CRA digitale	Suboxone	Astinenza
				Suboxone	CM	
AKL-T01 ⁽⁵⁾	RCT	348	ADHD	AKL-T01	AKL-T09	Cambiamento <i>Attention Performance Index</i>
						Cambiamento score <i>Test of Variables of Attention</i>
Noom ⁽⁶⁾	RCT	202	Pre-diabete	Noom Coach DPP	Terapia <i>standard</i>	Cambiamento peso
Clickotine ⁽⁷⁾	Interventistico singolo braccio	416	Dipendenza da fumo	Soluzione <i>mobile</i> per la cessazione del fumo	\	Numero di partecipanti che rimangono utenti attivi
Bluestar ⁽⁸⁾	RCT	163	Diabete di tipo 2	<i>Coach</i>	Terapia <i>standard</i>	Diminuzione di emoglobina glicata
				<i>Coach-PCP Portal</i>		
				<i>Coach-PCP Portal with Decision Support</i>		
Diabeo ⁽⁹⁾	RCT	667	Diabete	Diabeo	Terapia <i>standard</i>	Diminuzione di emoglobina glicata
				Diabeo + Telemonitoraggio		
Sleepio ⁽¹⁰⁾	RCT	120	Insonnia cronica	Terapia cognitivo comportamentale	<i>Placebo</i>	Efficienza del sonno
					Terapia <i>standard</i>	
Kaia COPD ⁽¹¹⁾	Osservazionale	56	COPD	Riabilitazione polmonare digitale	/	<i>Test</i> di valutazione COPD

Kaia back pain ⁽¹²⁾	Retrospettivo	180	Dolore cronico alla schiena	Applicazione	/	Livello di dolore
Propeller Health ⁽¹³⁾	RCT	495	Asma	Monitoraggio e <i>feedback</i>	Monitoraggio	Riduzione dell'utilizzo di beta-agonisti
ProAir Digihaler ⁽¹⁴⁾	RCT	330	Asma	Albuterolo con integrato Module Digital System elettronico (eMDPI DS)	Albuterolo	Miglioramento Asma <i>Control Test</i>
myCOPD ⁽¹⁵⁾	RCT	60	COPD	Applicazione	Terapia <i>standard</i>	Questionario CAT
Omada health ⁽¹⁶⁾	RCT	600	Pre-diabete	Programma di prevenzione diabete <i>online</i>	Terapia <i>standard</i>	Riduzione emoglobina glicata
Happify health ⁽¹⁷⁾	RCT	4485	Depressione e ansia	Happify	Psicoeducazione	Questionario PHQ-9, GAD-7
ATENTIVmynd ⁽¹⁸⁾	RCT	46	ADHD	<i>Feed Forward Modeling</i>	Terapia <i>standard</i> non farmacologica	ADHD-RS, scala CGI, <i>Quotient</i> ADHD, PERMP, WJ-III
SmartQuit ⁽¹⁹⁾	RCT	2503	Fumo	Smartquit	Applicazione parte di terapia <i>standard</i>	Prevalenza di astinenza di 30 giorni
Natural Cycles ⁽²⁰⁾	Prospettico	16331	Controllo delle nascite	Natural Cycles	/	Numero di gravidanze indesiderate
Somryst ⁽²¹⁾	RCT	303	Insonnia cronica	<i>Sleep Healthy Using the Internet</i>	Educazione paziente <i>online</i>	Indice di severità dell'insonnia
CureApp ⁽²²⁾	RCT	584	Dipendenza da nicotina	CureApp	Dispositivo di controllo	Tasso di astinenza continuo
DigCog ⁽²³⁾	RCT	60	Sclerosi multipla	Tablet Game	Tablet Game 2	Cambiamento della velocità di esecuzione
iCanQuit ⁽²⁴⁾	RCT	2415	Dipendenza da fumo	iCanQuit (<i>Mobile Smoking Cessation Solution</i> basata su terapia cognitivo comportamentale)	QuitGuide (<i>Mobile Smoking Cessations</i> basata su indicazioni del National Cancer Institute)	Astinenza da fumo

Balance ⁽²⁵⁾	RCT	244	Potus pericoloso o dannoso	Online single session screening procedure + online multi-session follow-up program (Balance)	Online single session screening procedure + online booklet about the effects of alcohol	Riduzione consumo alcol
One Drop App With an Activity Tracker ⁽²⁶⁾	RCT	95	Diabete di tipo 1	One Drop App + Activity Tracker	One Drop	Riduzione di emoglobina glicata
Dynamicare ⁽²⁷⁾	Studio di coorte	108	Disturbo da uso di sostanze	Applicazione <i>smartphone</i>	Terapia <i>standard</i>	Presenza agli appuntamenti - Utilizzo di sostanze - Durata del trattamento
Constant Therapy ⁽²⁸⁾	Retrospektivo	3686	Afasia post-ictus	Constant therapy	/	Numero di giorni per miglioramento
MindScience ⁽²⁹⁾	RCT	65	Disordini d'ansia	Unwinding Anxiety Phone App e Terapia <i>standard</i>	Terapia <i>standard</i>	Unwinding Anxiety program management - Penny State Worry Questionnaire

La tabella rappresenta lo stato dell'arte al 1 ottobre 2020. RCT: *Randomized controlled trial* (trial clinico controllato e randomizzato); TAU: *Therapy As Usual* (terapia *standard*); CRA: *Community Reinforcement Approach* (nome di terapia comportamentale); CM: *Contingency Management* (nome di terapia comportamentale); ADHD: *Attention deficit hyperactivity disorder* (disturbo da deficit di attenzione/iperattività); PRO: *Patient Reported Outcome*; COPD: *Chronic obstructive pulmonary disease* (broncopneumopatia cronica ostruttiva).

Si può notare che per quanto riguarda i prodotti in fase avanzata di sviluppo, molti di essi hanno completato *Randomized Controlled Trials* (RCT), con almeno 200 pazienti valutati. Per quanto riguarda gli *outcome* primari, sono stati considerati parametri fisiologici, comportamenti dei pazienti o risultati di questionari validati. Per quanto concerne il gruppo di controllo, si riscontra la preferenza ad utilizzare la terapia abituale oppure *placebo* digitali considerati “analoghi” del *placebo* negli RCT. In particola-

re, con “*placebo* digitale” ci si riferisce ad una soluzione digitale comparabile rispetto alla DTx, in termini di contenuti (contiene tutte le informazioni fornite dalla DTx, ma tramite interfacce statiche e non interattive) o di presentazione grafica (presentando stesse interfacce e introducendo la stessa *routine* di utilizzo dell’applicazione, ma senza elementi responsabili del beneficio clinico).

2.2 Prodotti in fase di sviluppo

Numerose DTx sono attualmente in fase di sviluppo. Di seguito sono elencati studi completati o in corso, registrati su clinicaltrials.gov.

Tabella 2 - Rappresentazione schematica di studi per la dimostrazione di efficacia di diverse DTx

Prodotto	Esperimento	Partecipanti	Condizione o malattia	Intervento	Controllo	Outcome primari	Stato
MyDiPP ⁽³⁰⁾	RCT	200	Pre-diabete	MyDiPP	Terapia <i>standard</i>	Riduzione peso	Reclutamento
One Drop ⁽³¹⁾	RCT	500	Pre-diabete	One Drop	Terapia <i>standard</i>	Variazione emoglobina glicata	In corso
			Diabete tipo 1				
			Diabete tipo 2				
MyMee ⁽³²⁾	RCT	50	Lupus eritematoso	MyMee	Nessun intervento	Brief Pain Inventory-Short Form	In corso
						Functional Assessment of Chronic Illness Therapy-Fatigue	
						LupusQOL	
Ibis ⁽³³⁾	RCT	240	COPD	Ibis	Nessun intervento	Riduzione cure intensive	Ignoto
Therapeutic Education System-Native Version (TES-NAV) ⁽³⁴⁾	RCT	80	Uso di sostanze	TES-NAV	Terapia <i>standard</i>	Numero di settimane di astinenza	Reclutamento
				Terapia <i>standard</i>			
No name declared ⁽³⁵⁾	RCT	195	Dipendenza dal fumo	Trattamento A mobile application	Trattamento A mobile application	Astinenza dal fumo	Reclutamento
FareWell Program ⁽³⁶⁾	Interventistico a braccio singolo	118	Diabete tipo 2	FareWell Program	\	Livello di <i>engagement</i>	Completato

PEAR-004 ⁽³⁷⁾	RCT	113	Schizofrenia	PEAR-004	Sham	Cambiamento nella <i>Positive and Negative Syndrome Scale</i>	Completato
						Percentuale di dropout	
NightWare Therapeutic System ⁽³⁸⁾	RCT	400	Sindrome da <i>stress</i> post traumatico	NightWare Therapeutic System	NightWare Therapeutic System in <i>sham mode</i>	Variazione <i>Pittsburgh Sleep Quality Index</i>	Reclutamento
NightWare Therapeutic System ⁽³⁹⁾	Interventistico a braccio singolo	400	Sindrome da <i>stress</i> post traumatico	NightWare Therapeutic System	\	Variazione <i>Pittsburgh Sleep Quality Index</i>	Reclutamento
NightWare Therapeutic System ⁽⁴⁰⁾	Interventistico a braccio singolo	15	Sindrome da <i>stress</i> post traumatico	NightWare Therapeutic System	\	Variazione <i>Pittsburgh Sleep Quality Index</i>	In corso
NightWare Therapeutic System ⁽⁴¹⁾	RCT	270	Sindrome da <i>stress</i> post traumatico	NightWare Therapeutic System	NightWare Therapeutic System in <i>sham mode</i>	Variazione <i>Pittsburgh Sleep Quality Index</i>	Reclutamento
Luminopia ⁽⁴²⁾	Interventistico a braccio singolo	84	Ambliopia	Luminopia	\	Acutezza visiva occhio ambliopico	Completato
Project:EVO ⁽⁴³⁾	Interventistico a braccio singolo	21	Sclerosi multipla	Project: EVO	\	<i>Brief International Cognitive Assessment for Multiple Sclerosis</i>	Completato
Limbox Spark ⁽⁴⁴⁾	RCT	410	Depressione	Limbox Spark	Psicoeducazione	Variazione nei sintomi depressivi	Reclutamento
Continuing Care ⁽⁴⁵⁾	RCT	100	Disturbo da uso di sostanze	Continuing Care app	Terapia <i>standard</i>	Settimane di trattamento	Reclutamento non iniziato
iSage ⁽⁴⁶⁾	RCT	40	Diabete Mellito di tipo 2 - Complicanze diabete	iSage App con glucometro collegato	Terapia <i>standard</i>	Variazione della emoglobina glicata	Reclutamento
ApricityRx ⁽⁴⁷⁾	Studio di coorte	100	Cancro	ApricityRx mobile application	/	Utilizzo dell' <i>App</i> per vedere video educazionali e dati riportati da pazienti	Reclutamento

Leva Pelvic Digital Health System ⁽⁴⁸⁾	RCT	350	Incontinenza urinaria da <i>stress</i>	Leva Pelvic Digital Health System	Esercizi di Kegel	<i>Urogenital Distress Inventory (UDI-6)</i> - Diario della vescica	Reclutamento
Vitadio ⁽⁴⁹⁾	RCT	100	Insulinoreistenza - Pre-diabete - Diabete di tipo 2 - Obesità	Vitadio Health	Terapia <i>standard</i>	Variazioni di peso	Reclutamento
Cognoa ⁽⁵⁰⁾	Interventistico a braccio singolo	30	Disturbi nello spettro dell'autismo	Cognoa	/	Usabilità	Reclutamento
PEAR-008 ⁽⁵¹⁾	RCT	130	Dipendenza da oppiacei	PEAR-008	reSET-O	<i>Engagement</i> del paziente	Reclutamento non iniziato
DTx for pain software ⁽⁵²⁾	RCT (3 bracci)	100	Lombalgia cronica	DTx	<i>Sham</i> - Terapia standard	<i>Self-report of disability, Oswestry Disability Index - Objective report of disability</i>	In corso
Gotcha! Therapy App ⁽⁵³⁾	RCT	34	Demenza o Alzheimer	<i>App-based therapy</i> + app mantenimento	<i>App-based therapy</i> senza mantenimento	Variazione nel tempo del <i>Gotcha Outcome Measure (GOM)</i>	In corso
iTALKbetter ⁽⁵⁴⁾	RCT	34	<i>Stroke</i> /afasia	iTALKbetter avanzata (<i>reactive</i>) <i>App-based therapy</i> per persone con anomia dopo <i>stroke</i>	iTALKbetter base (<i>deterministic</i>)	Variazione nell'accuratezza della <i>performance</i> in un <i>Word Retrieval Test</i> specificamente disegnato	In corso

La tabella rappresenta lo stato dell'arte al 1 ottobre 2020.

RCT: *Randomized controlled trial* (trial clinico controllato e randomizzato); QOL: *Quality of life*
 COPD: *Chronic obstructive pulmonary disease* (broncopneumopatia cronica ostruttiva).

3. Caratteristiche degli studi pilota

Uno studio “pilota” o studio di fattibilità è uno studio generalmente su piccola scala, utile per verificare se il progetto è adeguato, stabilirne la fattibilità o ricavare informazioni che permettano di determinare la grandezza del campione dello studio allargato (*pivotal*)⁽⁵⁵⁾.

Uno studio pilota, nel caso di *software* medicali, può essere rilevante in caso di:

- dimostrazione dell’usabilità dell’interfaccia o dei dispositivi *hardware* associati (per esempio, sensori indossabili);
- creazione di un *database* di *training* per intelligenza artificiale;
- definizione, per studi successivi, della popolazione da valutare (inclusa la determinazione del campione), di elementi chiave del disegno dello studio come reclutamento, modalità di randomizzazione e *blinding*, nonché dell’obiettivo primario e degli *endpoint* secondari e/o esplorativi;
- dimostrazione e quantificazione della “valida associazione clinica” tra gli *output* erogati dal *software* e il beneficio clinico.

Quasi tutti gli studi pilota sono non randomizzati e senza gruppo di controllo. Inoltre, non sempre includono *endpoint* di efficacia, mentre spesso si concentrano su *endpoint* di sicurezza.

Di seguito sono illustrati due casi esemplificativi, con ampiezza e caratteristiche diverse.

3.1. Intervento per la cessazione dell’abitudine al fumo (Clickotine)

Caso esemplificativo di studio iniziale condotto su una popolazione ampia di >400 soggetti, in cui sono stati valutati sia outcome di usabilità, sia outcome di efficacia e sicurezza⁽⁷⁾.

Clickotine è un’*App* fruibile su *smartphone*, disegnata e sviluppata per rilasciare gli elementi essenziali del programma statunitense per la cessazione del fumo USCPG (*US Clinical Practice Guideline*), ad esempio avvisi e incoraggiamenti per smettere di fumare, valutazione della volontà di smettere, stimoli motivazionali, supporto alla pianificazione, connessione con altri interventi (consulenza, trattamento farmacologico, supporto sociale, linea telefonica dedicata), *follow-up*.

Nel 2017 sono stati pubblicati i risultati di uno studio iniziale, a singolo braccio, della durata di 8 settimane. Gli *outcome* misurati erano il coinvolgimento del soggetto (numero di aperture dell’*App*, numero di interazioni con il programma, settimane attive nel programma), l’efficacia sulla cessazione del fumo (astinenza dal fumo riferita dal paziente a 7 e 30

giorni) e la sicurezza (eventi avversi spontaneamente riferiti e attivamente indagati tramite questionario al termine delle 8 settimane del programma). I dati ottenuti sono stati analizzati mediante opportuni metodi statistici; in aggiunta, sono anche state condotte analisi *post hoc*, motivate dalla distribuzione non normale della maggior parte delle variabili predittive o delle variabili dicotomiche di *outcome*. Nel complesso, lo studio ha prodotto dati incoraggianti relativamente al coinvolgimento: in media ogni partecipante ha aperto più di 100 volte l'*App* durante le 8 settimane (media: 110.6; mediana: 69); i partecipanti hanno avuto numerose interazioni con il programma (media: 214.4; mediana 178) e sono rimasti ingaggiati per una media

di 5.3 settimane (mediana: 5 settimane). Inoltre, dopo le 8 settimane, il 45.2% della popolazione *intention-to-treat* ha riferito un'astinenza dal fumo di 7 giorni e il 26.2% ha riportato un periodo di astinenza di 30 giorni. I pochi eventi avversi segnalati erano in linea con i sintomi da sospensione di nicotina, mentre non sono stati segnalati eventi relativi alla *safety* riconducibili in maniera specifica al prodotto.

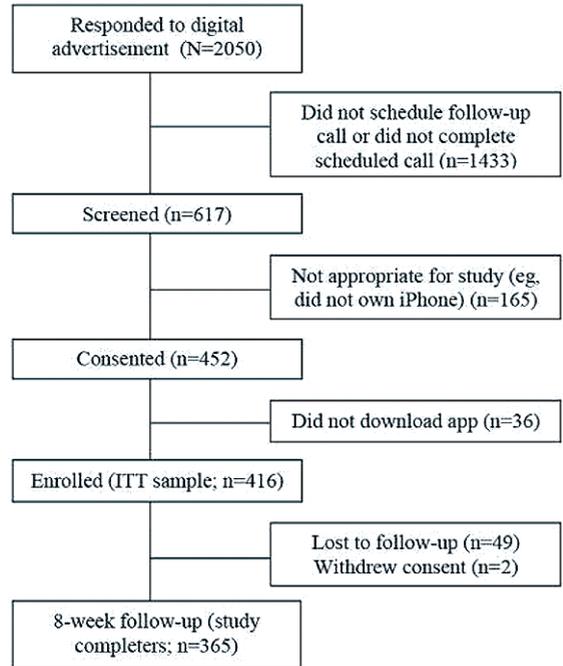
Nella *figura 1* è riportata la rappresentazione schematica della *flow-chart* dello studio (studio registrato in [clinicaltrials.gov: NCT02656745](https://clinicaltrials.gov/ct2/show/study/NCT02656745)).

3.2. Intervento per la gestione del disturbo d'ansia (*Unwinding Anxiety*)

Caso esemplificativo di studio pilota randomizzato e controllato verso terapia abituale, condotto su una popolazione di circa 60 soggetti (numero comune in studi pilota), in cui sono stati valutati sia outcome di ingaggio/accettabilità, sia outcome di efficacia.

Unwinding Anxiety è un programma di intervento comportamentale,

Figura 1 - Patient disposition per lo studio NCT02656745 condotto con il prodotto Clickotine



fruibile su *smartphone* o *tablet*, che si propone di aiutare l'utente nella gestione dell'ansia, inclusi il disturbo d'ansia generalizzata, il disturbo da attacchi di panico, l'ansia sociale. Il programma include video-lezioni, esercizi, contatto settimanale con esperti, stesura di diari (*journaling*), una *community* moderata da esperti, momenti quotidiani di verifica.

Unwinding Anxiety ha completato uno studio pilota, randomizzato, controllato verso la sola terapia abituale, in soggetti con disturbo d'ansia generalizzata, con obiettivo di valutare l'ingaggio degli utenti e l'accettabilità, misurare gli effetti dell'intervento in associazione a trattamento abituale *vs* il solo trattamento abituale, e valutare preliminarmente il meccanismo d'azione. Lo studio ha coinvolto 65 pazienti adulti (di cui 57 sono risultati utili per la valutazione). I criteri di inclusione prevedevano il possesso di uno *smartphone* e uno score GAD-7 (*General Anxiety Disorder-7*) pari o superiore a 10.

Le misure di *outcome* primario erano il numero di moduli completati (per valutare l'ingaggio) e la variazione del grado di preoccupazione secondo il *Penn State Worry Questionnaire* (PSWQ). Altre misure di *outcome* includevano l'utilizzo di questionari validati per la valutazione della *mindfulness* (FFMQ), dell'interocezione (MAIA), del livello di ansia (GAD-7) e dell'accettabilità del programma da parte degli utenti (NPS). L'aspettativa degli sperimentatori è di utilizzare i risultati di questa fase per disegnare un più ampio studio di fase 2 (Studio registrato in [clinicaltrials.gov: NCT03683472](https://clinicaltrials.gov/ct2/show/study/NCT03683472)).

È interessante osservare che, in uno studio indipendente condotto da sperimentatori australiani, *Unwinding Anxiety* è risultata essere una delle uniche 2 *App* (su 348 *App* di salute e benessere valutate) ad aver ottenuto un punteggio di 4 su 5 (massimo punteggio ottenuto). Lo studio ha preso in considerazione parametri multipli incentrati sulla funzionalità e sul potenziale di modificare il comportamento⁽⁵⁶⁾.

Inoltre, un recente studio pilota ha investigato gli effetti di *Unwinding Anxiety* in un campione di medici statunitensi affetti da ansia (N=34), dimostrando una riduzione dell'ansia (GAD-7) pari al 48% dopo 28 moduli (1 mese), e del 57% dopo 3 mesi ($p < 0.001$ in entrambi i casi)⁽⁵⁷⁾.

4. Caratteristiche degli studi *pivotal*

Uno studio "*pivotal*" è uno studio specificamente progettato per arricchire i dati raccolti durante lo sviluppo e quindi fornire adeguate evidenze cliniche di efficacia e *safety* per la valutazione da parte di

un'agenzia regolatoria. In tutti i casi in cui le evidenze in letteratura non sono sufficienti per la certificazione CE, il fabbricante deve eseguire uno studio *pivotal*.

In generale, le prove di efficacia per le terapie digitali dovrebbero provenire da studi clinici di elevata qualità, robusti e condotti secondo gli *standard* promossi dall'*evidence based medicine* e dalla migliore letteratura metodologica disponibile. Un utile riferimento è rappresentato dall'estensione dei *Consolidated Standards of Reporting Trials* (CONSORT) a interventi *web-based* e *m-health*⁽⁵⁸⁾. Inoltre si dovrebbe dimostrare che il *software*, durante lo studio, è stato testato per tutti gli usi previsti, rispettando la popolazione *target*, le condizioni d'uso, il funzionamento e l'ambiente d'uso, e con tutti i gruppi di utenti previsti.

Il disegno dello studio deve consentire di raccogliere evidenze statisticamente significative, il che comporta la necessità di fare riferimento a casistiche di adeguata ampiezza, la cui dimensione sia definita a priori sulla base di una ipotesi statistica. Deve anche consentire di documentare quale sia l'eventuale *place in therapy* del dispositivo, definendo in base a questo quale sia l'appropriato gruppo di controllo. Le caratteristiche del *software* medico rendono spesso molto complicati i disegni di studio in cieco, per cui la scelta degli *endpoint* e delle metriche risulta essere essenziale ai fini dell'affidabilità e consistenza metodologica degli studi.

Si consiglia, in particolare per gli studi *pivotal*, la realizzazione di studi multicentrici. Oltre a permettere la raccolta di dati da più ospedali/zone geografiche e offrire quindi evidenze maggiormente rappresentative del potenziale *target* delle DTx, questo approccio avrebbe anche il vantaggio di ridurre le tempistiche di arruolamento per lo studio, una esigenza di particolare importanza per affrontare il rischio di obsolescenza tecnologica dello strumento proposto.

A livello europeo le **indicazioni fornite da NICE** (*The National Institute for Health and Care Excellence*, UK) nell'ambito del documento sugli *standard* di evidenza per le tecnologie di *digital health* (marzo 2019) possono assumere la funzione di guidare gli sperimentatori.

I requisiti dettati dal NICE per l'evidenza clinica variano a seconda del tipo di intervento che compie il dispositivo, distinguendo se il *software* è destinato alla prevenzione o se il *software* produce esso stesso un ef-

fetto terapeutico. Nel secondo caso i requisiti sono, come prevedibile, più stringenti.

I *software* vengono classificati secondo lo schema presente nel documento “*Evidence standards for digital health technologies*”⁽⁵⁹⁾. Le DTx in particolare ricadono nel livello di evidenza 3, che si suddivide a sua volta in 3a e 3b, e secondo il seguente schema:

- se il *software* viene utilizzato per modificare il comportamento del soggetto, come nel caso della prevenzione e aspetti corollari al trattamento della condizione clinica, viene inserito nella classe 3a;
- se il *software* viene utilizzato per il trattamento viene inserito nella classe 3b.

Ogni classe deve soddisfare specifici requisiti di evidenza (*tabelle 3 e 4*), ma è anche necessario che questi rispettino tutte le condizioni dettate per le classi più basse (i requisiti sono quindi cumulativi).

Tabella 3 - Evidenze richieste a dispositivi di classe 3a

Categoria	Evidenza minima	Evidenza consigliata
Dimostrazione di efficacia	<p>Studi osservazionali o interventistici non randomizzati di alta qualità dimostranti risultati significativi e che presentino dati comparativi.</p> <p>Il braccio di controllo può includere:</p> <ul style="list-style-type: none"> • <i>Outcome</i> raccolti <i>ex novo</i> in un gruppo di controllo • Un braccio di controllo “storico” (confronto retrospettivo) • Dati raccolti abitualmente nella pratica clinica (e.g. <i>database</i> amministrativi, <i>electronic health records</i>). <p>Questi risultati possono essere:</p> <ul style="list-style-type: none"> • Evidenza di cambiamenti positivi di comportamento • Soddisfazione utente. 	<p>Studi interventistici di alta qualità con gruppi di controllo, dimostranti miglioramenti in risultati rilevanti.</p> <p>I risultati possono essere:</p> <ul style="list-style-type: none"> • <i>Patient-related outcomes</i> (PRO) • Misure cliniche di gravità o disabilità • Comportamenti salutari • Misure fisiologiche • Soddisfazione e <i>engagement</i> utente • Uso di risorse sanitarie o di assistenza sociale. <p>Il controllo deve essere rappresentativo dell’attuale percorso di cura.</p>
Utilizzo di tecniche appropriate per i cambiamenti comportamentali (se rilevante):	<p>Dimostrare che le tecniche usate sono:</p> <ul style="list-style-type: none"> • Coerenti con una teoria riconosciuta di influenza comportamentale (in linea con linee guida NICE o di organizzazioni professionali) • Appropriate per la popolazione <i>target</i>. 	<p>Pubblicazioni di evidenze qualitative e quantitative:</p> <ul style="list-style-type: none"> • Basate su tecniche pubblicate e riconosciute • Allineate con la pratica clinica raccomandata • Appropriate per la popolazione <i>target</i>.

Tabella 4 - Evidenze richieste a dispositivi di classe 3b

Categoria	Evidenza minima	Evidenza consigliata
Dimostrazione di efficacia	<p>Studi interventistici di alta qualità con gruppi di controllo. I risultati possono essere:</p> <ul style="list-style-type: none"> • Accuratezza diagnostica • <i>Patient-related outcomes</i> (PRO) (con strumenti validati) • Misure cliniche di gravità o disabilità • Comportamenti salutari • Misure fisiologiche • Soddisfazione ed <i>engagement</i> utente. <p>Il controllo deve essere rappresentativo dell'attuale percorso di cura.</p>	<p>Studio controllato randomizzato di alta qualità o studi condotti in un ambiente rilevante per il sistema sanitario e di assistenza sociale del Paese <i>target</i> della sperimentazione, confrontando DTx con un comparatore pertinente e dimostrando un beneficio sugli esiti clinici nella popolazione <i>target</i>, utilizzando metriche specifiche. In alternativa, se ci sono abbastanza studi disponibili sulla DTx, una meta-analisi ben condotta di studi controllati randomizzati.</p>

Per questo è importante analizzare anche i requisiti che devono essere rispettati dai dispositivi di prevenzione. I livelli di evidenza indicati dal NICE includono dei livelli minimi di evidenza scientifica e dei livelli consigliati, definiti come “*best practices*”.

È possibile notare come gli studi di prodotti in stato di sviluppo avanzato (*tabella 1*) ricadano tutti nella categoria dell'evidenza consigliata per i dispositivi 3b, mentre gli studi pilota possono avere una maggiore flessibilità nella progettazione. Di seguito sono illustrati tre casi esemplificativi di studi *pivotal* relativi a DTx in fase avanzata di sviluppo. Tutti gli studi sotto riportati hanno dimostrato un beneficio statisticamente significativo della DTx (*vs* il trattamento di controllo) rispetto agli *endpoint* primari, rappresentati, nei diversi casi, da *endpoint* di efficacia clinica o da *endpoint* associati alla prevenzione rispetto alla patologia di interesse. Da ricordare che tanto la finalità terapeutica quanto quella preventiva sono coerenti con le comuni definizioni di DTx, per esempio quella di *Digital Medicine Society/Digital Therapeutic Alliance*.

4.1. DTx per deficit di attenzione e iperattività (ADHD) in età pediatrica (Akili, AKL-T01)

Caso esemplificativo di studio pivotal per DTx utilizzata in monoterapia, controllato verso intervento digitale attivo.

AKL-T01 è un trattamento digitale basato su una interfaccia *videogame*, sviluppato per il trattamento del disturbo da deficit dell'attenzione

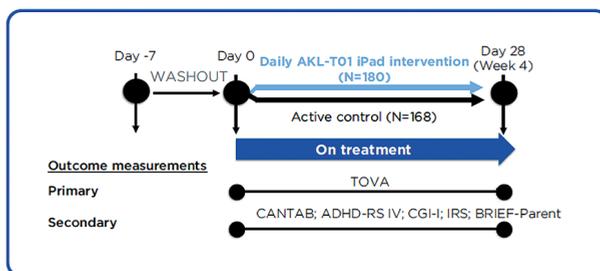
ed iperattività (ADHD) pediatrico. Il *software* agisce inviando stimoli motori e sensori in grado di raggiungere e attivare la corteccia pre-frontale (area cerebrale coinvolta nelle funzioni cognitive). Lo sviluppo clinico si è articolato in vari studi, sino al completamento di uno studio *pivotal* multicentrico (20 centri), randomizzato, in doppio cieco, controllato verso intervento attivo. Lo studio ha coinvolto 348 bambini di 8-12 anni con diagnosi confermata di ADHD. La DTx era utilizzata in monoterapia (i bambini non assumevano terapia farmacologica, o erano in grado di sospendere un'eventuale terapia farmacologica precedente), l'intervento di controllo era rappresentato da un diverso *videogame* (AKL-T09); l'*endpoint* primario era rappresentato dal miglioramento dell'indicatore TOVA-API (*Test of Variable Attention Performance Index*), misurato al basale e al termine di 4 settimane di studio (terapia circa 25 minuti al giorno, 5 giorni la settimana)⁽⁵⁾.

Questo studio è rilevante in quanto ha rappresentato uno dei primi e più rigorosi casi in cui una DTx è stata valutata in un ampio RCT prospettico, con approccio analogo a quanto usualmente effettuato per tera-

Figura 2 - Flow-chart dello studio NCT02674633, per valutare gli effetti del prodotto AKL-T01 nel deficit di attenzione e iperattività (ADHD) nei bambini

Pediatric ADHD - Registration Trial

Description	<ul style="list-style-type: none"> Multi-center, double-blind, randomized placebo-controlled At-home treatment for 4 weeks (28 days), 30 minutes per day, 5 times per week
Control	<ul style="list-style-type: none"> 1:1 randomization versus validated active control digital intervention (with similar expectation of effect)
Sample	<ul style="list-style-type: none"> Children aged 8 to 12 years (N=348)
Stratification	<ul style="list-style-type: none"> Not currently on medication; Switching off medication



Key inclusion criteria

- Confirmed diagnosis of ADHD per MINI-Kid as defined by DSM-V
- Currently off of stimulants OR poor controlled on stimulants and willing to wash-out
- ADHD-RS ≥ 28 without medication
- TOVA-API ≤ -1.8

Key exclusion criteria

- Comorbid psychiatric diagnosis
- Motor condition, intellectual disability, or color blindness

CGI-I, Clinical Global Impressions-Improvement; DSM-V, Diagnostic and Statistical Manual of Mental Disorders v5; IRS, Impairment Rating Scale; MINI-KID, Mini-International Neuropsychiatric Interview for Children and Adolescents; PedsQL, Pediatric Quality of Life Inventory

pie farmacologiche. In *figura 2* è riportata una schematizzazione del disegno dello studio (studio registrato in clinicaltrials.gov: NCT02674633).

4.2. DTx per il trattamento dell'abuso di sostanze (reSET)

Caso esemplificativo di studio pivotal per DTx utilizzata in associazione a terapia abituale modificata, controllato verso la sola terapia abituale.

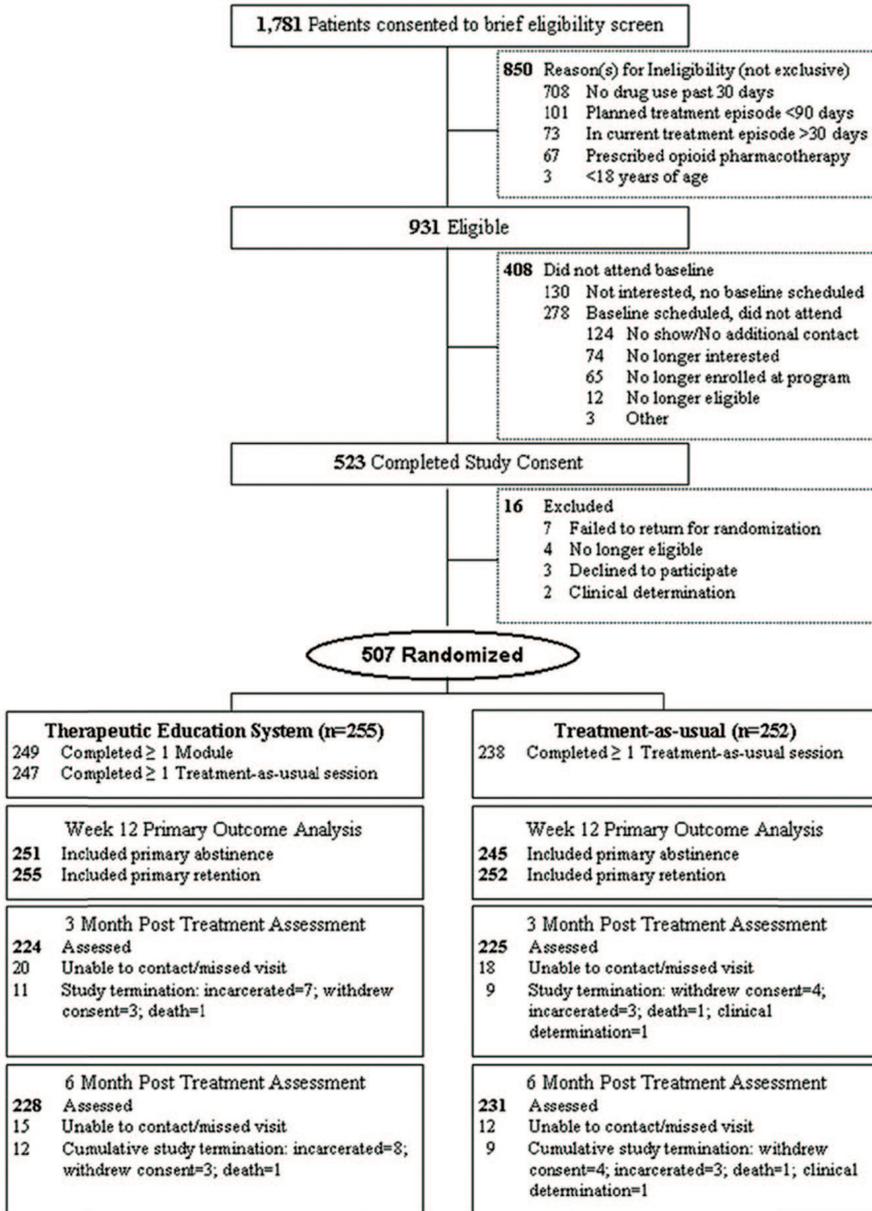
reSET è un intervento digitale per il trattamento dell'abuso di sostanze. Consiste in un programma di terapia cognitivo-comportamentale (CBT) della durata di 90 giorni, da somministrarsi come intervento parzialmente sostitutivo della terapia abituale (una parte del trattamento *vis-à-vis* viene sostituito da reSET) e aggiuntivo alla normale gestione delle contingenze. La soluzione consente di integrare una *dashboard* per medici e altri operatori sanitari, con informazioni sull'uso di reSET da parte dei pazienti, uso di sostanze riferito dai pazienti e rilevato da dati di laboratorio, e altre informazioni fornite dai pazienti.

Lo studio *pivotal*, randomizzato, controllato, multicentrico (10 centri) ha coinvolto inizialmente 1.781 soggetti adulti, 507 dei quali sono stati randomizzati a ricevere la sola terapia abituale o reSET in associazione a un programma di terapia abituale modificata (ridotta). La componente reSET poteva essere fruita dal partecipante su *computer* presenti nei centri o autonomamente al di fuori del centro. Gli *outcome* primari pre-specificati erano l'astinenza da droghe pesanti o alcool nelle ultime 4 settimane (auto-riferita o basata su analisi delle urine)⁽³⁾.

Lo studio ha evidenziato la significatività statistica dell'aumento della probabilità di astinenza per i soggetti che fanno uso di stimolanti. In particolare, i pazienti trattati con reSET hanno avuto una minore riduzione del *drop-out* dal trattamento rispetto al trattamento abituale (*Hazard Ratio*=0.72 [IC 95% CI, 0.57-0.92], $p=0.010$), oltre ad un incremento del periodo di astinenza (*Odds Ratio*=1.62 [IC 95%: 1.12-2.35], $p=0.010$); l'effetto era più marcato nei pazienti che, al momento dello *screening* iniziale per l'ingresso nello studio, presentavano un campione di urina positivo per la presenza di sostanze e/o positività al *breath-test* per l'alcool ($n=228$) (*Odds Ratio*=2.18 [IC 95%: 1.30-3.68], $p=0.003$).

Di seguito il diagramma di flusso relativo all'arruolamento e disegno dello studio (studio registrato in clinicaltrials.gov: NCT01104805):

Figura 3 - Patient disposition dello studio NCT01104805, condotto con il prodotto reSET



4.3. Intervento digitale per la prevenzione del diabete (*Noom Coach app*)

Caso esemplificativo di RCT per un programma digitale di coaching con finalità preventiva.

Noom Coach app è un intervento di *coaching* virtuale che eroga un programma di prevenzione del diabete (DPP) attraverso un'interfaccia interattiva. L'App include funzioni di messaggistica tra *coach* e partecipante, messaggi di gruppo, supporto alle modifiche comportamentali attraverso l'impostazione di sfide quotidiane, articoli educazionali sul DPP, *input* relativo al consumo di alimenti e *feedback* automatico rispetto alle scelte alimentari. Dopo un'ampia fase pilota, è stato recentemente completato un RCT che ha coinvolto 202 partecipanti adulti e il cui obiettivo primario era valutare l'efficacia della piattaforma *mobile* Noom Coach DPP rispetto all'intervento medico abituale (i partecipanti erano liberi di cercare assistenza o supporto medico durante lo studio, oltre ad avere a disposizione una versione cartacea del DPP)⁽⁶⁾. Obiettivo primario era la modifica del peso corporeo misurato al basale, dopo 6 mesi e al termine dello studio (durata complessiva 52 settimane); ulteriori obiettivi includevano le variazioni nei livelli di emoglobina glicosilata (HbA1c), la variazione della produttività lavorativa (attraverso questionario validato), la variazione della qualità di vita (attraverso questionario CDC HRQOL-4).

5. Valutazione del profilo di sicurezza

Accanto alla valutazione del beneficio, gli studi *pivotal* (in alternativa "le sperimentazioni cliniche randomizzate") forniscono in genere informazioni anche sui possibili effetti collaterali associati al (nuovo) trattamento in fase di valutazione. Questa misura, insieme a quella del beneficio, serve a determinarne la tollerabilità e quindi a valutare il rapporto rischio/beneficio del nuovo trattamento.

Anche gli interventi di DTx possono provocare effetti indesiderati. Sono in genere meno gravi rispetto a quelli causati dai farmaci tradizionali, spesso più gestibili, ma possono presentarsi in misura maggiore rispetto ai rispettivi bracci di controllo e non devono comunque in ogni caso essere sottovalutati.

Per esempio, uno studio che ha dimostrato l'efficacia dell'intervento digitale per la cessazione da fumo ha evidenziato in due pazienti la comparsa di effetti collaterali (cambiamento dell'umore, incubo) probabil-

mente associati al trattamento e, in pochi altri casi, affaticamento (non associato al trattamento)⁽⁷⁾.

Di recente, nell'illustrare i benefici di Endeavor (precedentemente noto con il codice sperimentale AKL-T01), la terapia digitale basata sui videogiochi per il trattamento di bambini con diagnosi di Disturbo da Deficit Attentivo con Iperattività e che ha ricevuto nel giugno 2020 l'autorizzazione alla commercializzazione da parte di FDA, non sono stati segnalati eventi avversi gravi associati al trattamento digitale. Tuttavia, nel 9.3% dei casi, i pazienti trattati negli studi hanno presentato eventi avversi non gravi (associati al trattamento digitale), tra cui senso di frustrazione, mal di testa, vertigini, reazioni emotive, nausea o aggressività⁽⁶⁰⁾.

Gli effetti avversi associabili alle terapie digitali vanno comunque misurati e valutati. Non sempre tuttavia sono evidenziati eventi imputabili all'intervento digitale. Per esempio, in uno studio randomizzato per misurare l'efficacia dell'intervento digitale nella gestione di pazienti diabetici si è osservato che eventi ipoglicemici, ricoveri e visite al Pronto Soccorso avevano una bassa percentuale, del tutto sovrapponibile a quella riscontrata nel gruppo di controllo, e che tali eventi avversi non erano correlati allo studio⁽⁸⁾.

Come accade per le terapie tradizionali, un'attenta valutazione degli effetti indesiderati associabili all'intervento digitale potrà avvenire con l'attivazione degli studi *post-marketing* (gli studi di fase IV), una volta che la terapia digitale sia stata immessa sul mercato, così da ottenere informazioni su popolazioni più ampie ed eterogenee, e sulla sicurezza nel lungo termine.

Altrettanto interessante e forse più rilevante è il tema dei *drop-out*, di coloro cioè che interrompono il trattamento (digitale) prima della conclusione della sperimentazione clinica. In uno studio per la valutazione dell'efficacia di una terapia digitale per il trattamento del dolore in pazienti con lombalgia, si è assistito a un progressivo abbandono dell'intervento da parte dei pazienti prima della conclusione dello studio⁽¹²⁾. Dopo 8 settimane di trattamento con la terapia digitale i pazienti rimasti aderenti alla terapia erano il 32% della popolazione in studio, mentre al termine dello studio (12 settimane) erano calati al 18%.

Una particolare attenzione al *design* dell'intervento digitale, un maggiore coinvolgimento dei pazienti nel suo sviluppo, una aumentata attenzione ai *feedback* degli utenti, un maggiore loro *engagement* grazie a promemoria, *e-mail*, SMS e messaggi, sono aspetti che dovrebbero essere presi in considerazione per ridurre al massimo questo fenomeno.

6. Il ruolo dei Comitati Etici

Per favorire un sistema che possa consentire all'Italia di contribuire in maniera significativa allo sviluppo di DTx, i Comitati Etici devono disporre di adeguato *know-how* e di specifiche linee guida, in considerazione del fatto che la sperimentazione di una DTx può differire significativamente sia da quella di un farmaco, sia da quella di un dispositivo medico "tradizionale" (ad esempio una protesi articolare) che i Comitati Etici sono abituati a valutare.

Il disegno degli studi con DTx deve garantire la protezione dei dati sensibili e permettere di gestire la rapida evoluzione tecnologica alla quale sono sottoposte.

In tale prospettiva, le principali criticità riguardano il controllo del cambiamento del prodotto nel corso dello studio e il salvataggio dei dati relativi allo studio stesso. A causa della rapidità di evoluzione del dispositivo *software stand-alone*, sono probabili frequenti aggiornamenti di versione del dispositivo nel corso dello svolgimento dello studio.

Il fabbricante pertanto deve prevedere per il protocollo di ricerca stesso e per il prodotto una serie di modifiche attese ("*change control*") che non richiedano l'approvazione del Comitato Etico e dell'Autorità Competente, perché relative a "manutenzione perfetta" (cosiddetto "*bugfix*") oppure perché sono modifiche relative ad aspetti di usabilità e non ad aspetti clinici o ancora ad aspetti tecnici inderogabili (per esempio, aggiornamento della piattaforma Android/iOs di una *App* medica).

7. Sorveglianza *post-marketing*

Già dalla prima emissione del certificato CE, il fabbricante della DTx è tenuto a pianificare una di raccolta dei dati clinici *post-marketing*. Per definizione, questi studi devono essere eseguiti rispettando le indicazioni d'uso e sulla popolazione di pazienti per cui il dispositivo è già certificato: non possono essere quindi utilizzati per testare nuove applicazioni cliniche del *software* ed hanno come scopo quello di confermare il rapporto rischi/benefici e di raccogliere informazioni relative all'utilizzo su larga scala. In tale prospettiva, gli studi di sorveglianza *post-marketing* offrono una ottima possibilità di raccogliere importanti informazioni in termini di aderenza e "*user experience*".

Il Regolamento MDR incoraggia l'uso dei risultati degli studi di *follow-*

up per fornire nuove informazioni per migliorare e aggiornare il *software*. Risulta quindi particolarmente virtuoso, per esempio, l'utilizzo di studi *post-marketing*, con adeguato controllo, per aumentare il *database* di *training* di una intelligenza artificiale medica.

La natura stessa della DTx permette una raccolta dati massiva nella fase post-commercializzazione del dispositivo. I produttori dovrebbero quindi mettere in pratica misure di raccolta dati clinici *real-world*, rispettando la *privacy* dei pazienti.

Un esempio virtuoso di questa prassi di raccolta dati è *Natural Cycles*, una DTx per il controllo delle nascite. Tramite l'utilizzo del dispositivo sono stati raccolti dati in diversi studi clinici osservazionali e retrospettivi per stimare l'efficacia del dispositivo sia in utilizzo sperimentale che in utilizzo nel mondo reale^(20, 61-63). Inoltre, dai dati ottenuti, il fabbricante ha avuto la possibilità di ottenere ulteriori indicazioni che esulano dalla dimostrazione di efficacia del dispositivo, ma legate ad attività direttamente connesse con l'uso del dispositivo, permettendo di condividere importanti informazioni di interesse scientifico^(64, 65).

Il sistema di raccolta dati del fabbricante dovrebbe inoltre tenere conto degli effetti sulla salute del paziente nelle fasi successive all'interruzione del trattamento.

8. Massime modifiche accettabili alla DTx che non richiedano inizio *ex-novo* della validazione clinica

Come precedentemente sottolineato, una peculiarità di prodotti come le DTx, rispetto per esempio al farmaco o a dispositivi medici di altra natura, è la possibilità di andare incontro a rapide evoluzioni tecnologiche che possono intervenire durante le fasi di validazione clinica. Si tratta di un argomento di particolare delicatezza e per il quale sarebbe opportuno prevedere linee di indirizzo comportamentale sufficientemente chiare e definite.

In tal senso, una proposta che si ritiene sostanzialmente condivisibile è quella avanzata dai “*Leaders in mHealth research, industry and health care systems from around the globe*” e riportata nell'articolo di Torous et al. “*Towards a consensus around standards for smartphone apps and digital mental health*”⁽⁶⁶⁾. Nello specifico, i principali contenuti della proposta sono:

“Changes in technology may mean that app updates need to be re-evaluated for their efficacy. Small cosmetic changes, platform changes and aspect

changes do not likely require a retest of an intervention, as long as the therapeutic principle that has been evaluated remains intact. However, significant changes, such as adding a new therapeutic principle or substantial changes to that principle, must demonstrate efficacy through the same evaluation pathways as novel therapeutics.

Our recommendations are:

a. *newly adapted therapeutic principles, which should be identified and defined, must undergo controlled clinical trials to determine their efficacy and effectiveness;*

b. *small changes to an app with an evidence base need not undergo another clinical trial, but any major change requires a re-evaluation of app effectiveness”.*

Sulla base di questi principi generali è peraltro fondamentale comprendere quali possano essere le modifiche “*significant*” o “*small*” che impongano o meno una rivalutazione degli effetti del prodotto.

Il fabbricante quindi deve definire delle politiche che includano la definizione dell’impatto delle modifiche sul dispositivo basate sull’analisi di rischio del dispositivo. Di seguito è presentato un esempio applicabile a molte DTx:

a. Modifica di tipo “*Major*”:

- qualsiasi variazione che abbia potenzialmente un impatto in termini di sicurezza, efficacia, caratteristiche di rischio del sistema *software* (esempio, viene creato un nuovo *database* di *training* con diversi indicatori clinici);

- qualsiasi variazione nell’uso previsto, per esempio:

- utilizzare il sistema per una diversa popolazione di pazienti

- il sistema viene integrato con un modulo che fornisce un mezzo per la segnalazione di emergenza clinica

- il sistema viene integrato con un modulo per gestire la interazione con dispositivi indossabili

- cambia il metodo di interazione con il paziente, introducendo l’interazione vocale

- il sistema viene integrato con un nuovo parametro misurabile aggiuntivo utilizzato come *input* per l’algoritmo;

- qualsiasi variazione nella struttura architettonica principale, intesa come una modifica dell’organizzazione dei moduli e degli elementi *software* che vada a modificare il flusso dei dati necessari per l’uso medico previsto, oppure che vada a modificare la segmentazione (per esempio il passag-

gio da un sistema basato su *App* scaricabile in locale a un sistema basato principalmente su *cloud*);

- qualsiasi variazione che potrebbe causare incompatibilità tra i moduli oppure con i sistemi di interfaccia (*database* o sistema operativo).

b. Modifica di tipo “Minor”

- qualsiasi modifica non inclusa nella definizione *Major*, per esempio
- miglioramento delle prestazioni, senza modifiche all’uso previsto o nuovi tipi di *input*

- aggiornamento del *database* di formazione o di quello di convalida, con dati comparabili, in termini di qualità e contenuti informativi, rispetto a quelli utilizzati nella versione precedente;

- modifiche relative agli *input* di accesso, logistici o gestionali, senza variazioni dell’uso previsto, per esempio

- nuovi mezzi per identificare l’utente (esempio biometria anziché *password*)

- inserimento dati tramite comando vocale da parte dell’utente professionale;

- aggiornamento dell’interfaccia utente, nuove icone

- traduzioni

- aggiornamento del manuale dell’utente

- modifiche o integrazioni a SOUP (Software Of Unknown Provenance)

non critiche

- Bot per assistenza tecnica a pazienti/utenti professionali, ad esempio per l’assistenza al cambio *password*.

Ogni modifica così classificata deve poi essere interpretata e analizzata quando proposta all’interno di una sperimentazione clinica:

- nel caso di una modifica “*Major*”, essa dovrebbe essere valutata nel dettaglio, per identificare se la modifica sia completamente non accettabile oppure se le informazioni raccolte con il prodotto nella versione precedente sono ancora parzialmente significative;

- le modifiche “*Minor*”, che come detto, dovrebbero essere accettabili e non comportare inizio *ex-novo* del processo di validazione, possono essere comunicate agli Investigatori e formalizzate tramite aggiornamento della *Investigator’s Brochure*, ed eventualmente agli utenti (per esempio nel caso di modifiche minori dell’interfaccia, e in tal caso potrebbero essere indicate eventuali attività di formazione per garantire che l’utente sia in grado di continuare ad utilizzare il prodotto);

- i *bugfix* sono da considerare accettabili e dovrebbero essere comunicati agli Investigatori e agli utenti con notifiche automatiche (esempio: notifiche in-App).

What is known

- I *software* che erogano terapie digitali (DTx) sono dispositivi medici e pertanto devono rispettare, in Europa, i requisiti del Regolamento 2017/745
- Il Regolamento stabilisce sia necessario dimostrare un beneficio clinico, attraverso una indagine clinica, senza però fornire dettagli circa le caratteristiche di questa indagine
- Esistono ad oggi alcuni esempi di DTx approvate, tante altre sono in fase di sviluppo. Il livello di evidenza disponibile o in via di generazione è variabile
- La valutazione del beneficio clinico, per le DTx, deve basarsi su evidenze cliniche dimostrabili e statisticamente significative nel migliorare uno specifico obiettivo di salute che sia misurabile. La novità ed innovatività di queste tecnologie non devono portare a facili entusiasmi, poiché non tutto ciò che è tecnologicamente avanzato è automaticamente utile, efficace e costo-efficace, almeno quanto se non più di quanto offerto dai metodi tradizionali
- I metodi tradizionali di ricerca delle evidenze possono non essere perfetti né applicabili *tout-court* alle DTx, ma deve essere chiaro l'obiettivo di raggiungere una esaustiva descrizione del rapporto rischio-beneficio e una buona conoscenza di ciò che può essere ragionevolmente previsto in termini di uso clinico.

What is uncertain

- In assenza di prescrizioni precise nel Regolamento, non è chiaro quali caratteristiche avranno gli studi clinici a supporto della certificazione di DTx in Italia e in Europa
- Le DTx presentano peculiarità che richiedono alcune accortezze in fase di disegno di studio, se si vuole raggiungere una esaustiva caratterizzazione del profilo rischio-beneficio: la scelta del controllo, le metriche per misurare gli esiti in termini di efficacia e sicurezza devono essere spesso “personalizzate” caso per caso.

What we recommend

- Avviare un approfondimento di tipo regolatorio a livello europeo relativamente alle indagini cliniche per le DTx, così da garantire, per un dispositivo a finalità terapeutica, *standard* di efficacia e sicurezza adeguati e uniformi
- Per consentire agli sviluppatori di raccogliere chiare evidenze sul beneficio, e alle Autorità regolatorie di poter valutare correttamente queste evidenze, si auspica:
 - che i benefici clinici possano essere espressi con metriche adeguate

- non solo all'indicazione d'uso prevista, ma anche alla tecnologia;
- che la scelta di un eventuale controllo in uno studio sia giustificata non solo in termini clinici ma anche tecnici (ad esempio, nel caso di una DTx per terapia cognitivo-comportamentale, il controllo potrebbe essere un farmaco, una terapia cognitivo-comportamentale in presenza o in telemedicina, oppure un normale videogioco);
 - che nel protocollo clinico sia possibile definire la “*pipeline*” dei possibili cambiamenti tecnici, di interfaccia e di architettura *software* attesi, e per ciascuno di essi sia definita una politica di impatto sull'andamento dello studio stesso;
 - che il numero di pazienti reclutati sia adeguato per mostrare i possibili vantaggi di DTx rispetto al braccio di controllo;
 - che siano attivate strategie di coinvolgimento dei pazienti (*engagement*), per ottimizzare la pianificazione e conduzione dello studio (es. limitare il numero di possibili abbandoni dagli studi/*drop out*) e selezionarne gli obiettivi maggiormente in grado di rispondere alle esigenze dei pazienti;
 - che siano proposti dei piani di validazione clinica che tengano conto degli effetti collaterali a lungo termine
- Si raccomanda inoltre che gli studi clinici riguardanti DTx possano essere valutati con il contributo delle diverse Autorità Regolatorie, con ad esempio l'intervento di *European Medicines Agency* (EMA) se la DTx viene confrontata con un farmaco, oppure se ha una destinazione d'uso simile o identica a un farmaco.

Riferimenti bibliografici

1. O'Connor S, Hanlon P, O'Donnell CA, et al. Barriers and facilitators to patient and public engagement and recruitment to digital health interventions: protocol of a systematic review of qualitative studies. *BMJ Open* 2016; 6: e010895.

2. Putrino D, Zanders H, Hamilton T, et al. Patient engagement is related to impairment reduction during digital game-based therapy in stroke. *Games Health J* 2017; 6: 295-302.

3. Campbell AN, Nunes EV, Matthews AG, et al. Internet-delivered treatment for substance abuse: a multi-site randomized controlled clinical trial. *Am J Psychiatry* 2014; 171: 683-90.

4. Christensen DR, Landes RD, Jackson L, et al. Adding an Internet-delivered treatment to an efficacious treatment package for opioid dependence. *J Consult Clin Psychol* 2014; 82: 964-72.

5. Kollins SH, Bower J, Findling RL, et al. 2.40 A multicenter, randomized, active-control registration trial of software treatment for actively reducing severity of ADHD (Stars-Adhd) to assess the efficacy and safety of a novel, home-based, digital treatment for pediatric ADHD. *J ACAAP* 2018; 57: S172.

6. Prevention of Diabetes Using Mobile-enabled, Virtual Delivery of the National Diabetes Prevention Program. <https://clinicaltrials.gov/ct2/show/NCT03865342>. I risultati di precedente studio pilota sono pubblicati come: Michaelides A, Major J, Pienkosz E Jr, Wood M, Kim Y, Toro-Ramos T. Usefulness of a novel mobile diabetes prevention program delivery platform with human coaching: 65-week observational follow-up. *JMIR Mhealth Uhealth* 2018; 6: e93.

7. Iacoviello BM, Steinerman JR, Klein DB, et al. Clickotine, a personalized smartphone App for smoking cessation: initial evaluation. *JMIR Mhealth Uhealth*. 2017; 5: e56.

8. Quinn CC, Shardell MD, Terrin ML, et al. Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. *Diabetes Care* 2011; 34: 1934-42.

9. Jeandidier N, Chaillous L, Franc S, et al. DIABEO App software and telemedicine versus usual follow-up in the treatment of diabetic patients: protocol for the TELESAGE randomized controlled trial. *JMIR Res Protoc* 2018; 7: e66.

10. Espie CA, Kyle SD, Williams C, et al. A randomized, placebo-controlled trial of online cognitive behavioral therapy for chronic insomnia disorder delivered via an automated media-rich web application. *Sleep* 2012; 35: 769-81.

11. Rassouli F, Boutellier D, Duss J, et al. Digitalizing multidisciplinary pulmonary rehabilitation in COPD with a smartphone application: an international observational pilot study. *Int J Chron Obstruct Pulmon Dis* 2018; 13: 3831-6.

12. Huber S, Priebe JA, Baumann KM, et al. Treatment of low back pain with a digital multidisciplinary pain treatment App: short-term results. *JMIR Rehabil Assist Technol* 2017; 4: e11.

13. Merchant RK, Inamdar R, Quade RC. Effectiveness of population health management using the Propeller Health Asthma Platform: a randomized clinical trial. *J Allergy Clin Immunol Pract* 2016; 4: 455-63.

14. A 12-week treatment study to evaluate the effectiveness of albuterol multidose dry powder inhaler with integrated electronic Module Digital System (eMDPI DS) in patients 13 years or older with asthma. <https://clinicaltrials.gov/ct2/show/NCT03890666>.

15. Clinical efficacy and cost effectiveness of MYCOPD in patients with mild and moderate newly diagnosed COPD. <https://clinicaltrials.gov/ct2/show/NCT03620630>.

16. Preventing diabetes with digital health and coaching. <https://clinicaltrials.gov/ct2/show/NCT03312764>.

17. Parks AC, Williams AL, Tugade MM, et al. Testing a scalable web and smartphone based intervention to improve depression, anxiety, and resilien-

ce: a randomized controlled trial. *Int J Wellbeing* 2018; 2: 8.

18. McDermott AF, Rose M, Norris T, Gordon EA. Novel feed-forward modeling system leads to sustained improvements in attention and academic performance. *J Atten Disord* 2016; 1087054715623044. doi:10.1177/1087054715623044.

19. Trial of an innovative smartphone intervention for smoking cessation. <https://clinicaltrials.gov/ct2/show/NCT02724462>.

20. Bull J, Rowland S, Lundberg O, et al. Typical use effectiveness of Natural Cycles: postmarket surveillance study investigating the impact of previous contraceptive choice on the risk of unintended pregnancy. *BMJ Open* 2019; 9: e026474.

21. Ritterband LM, Thorndike FP, Ingersoll KS, et al. Effect of a web-based cognitive behavior therapy for insomnia intervention with 1-year follow-up: a randomized clinical trial. *JAMA Psychiatry* 2017; 74: 68-75.

22. Masaki K, Tateno H, Nomura A, et al. A randomized controlled trial of a smoking cessation smartphone application with a carbon monoxide checker. *NPJ Digit Med* 2020; 3: 35.

23. Bove R, Rowles W, Zhao C, et al. A novel in-home digital treatment to improve processing speed in people with multiple sclerosis: a pilot study. *Mult Scler* 2020; 1352458520930371. doi: 10.1177/1352458520930371.

24. Bricker JB, Watson NL, Mull KE, et al. Efficacy of smartphone applications for smoking cessation: a randomized clinical trial. *JAMA Intern Med* 2020; e204055. doi: 10.1001/jamainternmed.2020.4055.

25. Brendryen H, Lund IO, Johansen AB, et al. Balance--a pragmatic randomized controlled trial of an online intensive self-help alcohol intervention. *Addiction* 2014; 109: 218-26.

26. Osborn CY, Hirsch A, Sears LE, et al. One Drop App With an Activity Tracker for adults with type 1 diabetes: randomized controlled trial. *JMIR Mhealth Uhealth* 2020; 8: e16745.

27. Contingency management using smartphone App in patients with SUD. <https://clinicaltrials.gov/ct2/show/NCT04162132>.

28. Braley M. A parallel, virtual, randomized trial of PCT for speech, language, and cognitive intervention in stroke patients. <https://clinicaltrials.gov/ct2/show/study/NCT04488029> (2020).

29. Developing a novel digital therapeutic for the treatment of generalized anxiety disorder. <https://clinicaltrials.gov/ct2/show/NCT03683472>.

30. A digital therapy for diabetes prevention among overweight adults in Terengganu, Malaysia. <https://clinicaltrials.gov/ct2/show/NCT03997656>.

31. Fit-One: A trial evaluating the effect of One Drop and Fitbit on diabetes and pre-diabetes outcomes. <https://clinicaltrials.gov/ct2/show/NCT03459573>.

32. FLARE Lupus Research Study Systemic Lupus Erythematosus. <https://clinicaltrials.gov/ct2/show/NCT03426384>.

33. Impact of Ibis on patients with advanced COPD. <https://clinicaltrials.gov/ct2/show/study/NCT03426384>.

als.gov/ct2/show/NCT03131622

34. Web-based addiction treatment: cultural adaptation with American Indians. <https://clinicaltrials.gov/ct2/show/NCT03363256>.

35. Innovative digital therapeutic for smoking cessation. <https://clinicaltrials.gov/ct2/show/NCT03694327>.

36. Effectiveness of a digital therapeutic on type 2 diabetes. <https://clinicaltrials.gov/ct2/show/NCT03369626>.

37. Study of efficacy of PEAR-004 in schizophrenia. <https://clinicaltrials.gov/ct2/show/NCT03751280>.

38. Remote study of NightWare for PTSD with nightmares. <https://clinicaltrials.gov/ct2/show/NCT03934658>.

39. NightWare Open Enrollment Study. <https://clinicaltrials.gov/ct2/show/NCT03828656>.

40. Traumatic nightmares treated by NightWare (to arouse not awaken). <https://clinicaltrials.gov/ct2/show/NCT03795987>.

41. Traumatic nightmares treated by NightWare (to arouse not awaken). <https://clinicaltrials.gov/ct2/show/NCT04040387>.

42. Luminopia VR Device Pilot Study. <https://clinicaltrials.gov/ct2/show/NCT02782117>.

43. Bove RM, Rush G, Zhao C, et al. A videogame-based digital therapeutic to improve processing speed in people with multiple sclerosis: a feasibility study. *Neurol Ther* 2019; 8: 135-45.

44. Limbix Health, Inc. Limbix Spark: A CBT-based mobile intervention as first line treatment for adolescent depression during COVID-19. <https://clinicaltrials.gov/ct2/show/study/NCT04524598> (2020)

45. Continuing Care App for probationers and parolees. <https://clinicaltrials.gov/ct2/show/NCT04574193>.

46. mHealth titration and management. <https://clinicaltrials.gov/ct2/show/NCT03908762>.

47. Apricity Health, LLC. A study of ApricityRx™ digital therapeutic for management of immune-related adverse events in patients on immuno-oncology therapy. <https://clinicaltrials.gov/ct2/show/study/NCT04571398> (2020).

48. Pelvic floor muscle training with a digital therapeutic device to standard exercises for stress urinary incontinence. <https://clinicaltrials.gov/ct2/show/NCT04508153>.

49. Impact of digital therapeutic on metabolic parameters. <https://clinicaltrials.gov/ct2/show/NCT04573296>.

50. Cognoa, Inc. Cognoa ASD digital therapeutic engagement and usability study. <https://clinicaltrials.gov/ct2/show/study/NCT04326231> (2020).

51. A remote, decentralized opioid use disorder study to evaluate patient engagement with a game-based digital therapeutic. <https://clinicaltrials.gov/>

ct2/show/NCT04542642.

52. Orion Corporation, Orion Pharma. DTx for pain: behavioural revalidation in augmented and virtual reality for chronic pain; and exploratory pilot study of a virtual reality software. <https://clinicaltrials.gov/ct2/show/NCT04225884> (2020).

53. University College, London. Digital interventions in neuro-rehabilitation: a small scale randomised clinical trial to test a web-based therapy application for people with naming difficulties with mild- moderate dementia (Gotcha!). <https://clinicaltrials.gov/ct2/show/NCT04570215> (2020)

54. University College, London. Digital interventions in neuro-rehabilitation (DINR): a digital neuro intervention (DNI) for word retrieval (anomia). <https://clinicaltrials.gov/ct2/show/NCT04566081> (2020).

55. Eldridge SM, Lancaster GA, Campbell MJ, et al. Defining feasibility and pilot studies in preparation for randomised controlled trials: development of a conceptual framework. *PLoS ONE* 2016; 11: e0150205.

56. <https://www.prnewswire.com/news-releases/new-study-ranks-mindsciences-unwinding-anxiety-among-top-1-of-health-wellness-apps-300789441.html>

57. Roy A, Druker S, Hoge EA, Brewer JA. Physician anxiety and burn-out: symptom correlates and a prospective pilot study of App-delivered mindfulness training. *JMIR Mhealth Uhealth* 2020; 8: e15608.

58. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH Improving and standardizing evaluation reports of web-based and mobile-health interventions. *J Med Internet Res* 2011; 13: e126.

59. Evidence standards framework for digital health technologies | Our programmes | What we do | About. NICE <https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies>.

60. Endeavour TM. Instruction for use. Akili Interactive Labs. 2020.

61. Berglund Scherwitzl E, Gemzell Danielsson K, Sellberg JA, Scherwitzl R. Fertility awareness-based mobile application for contraception. *Eur J Contracept Reprod Health Care* 2016; 21: 234-41.

62. Berglund Scherwitzl E, Lundberg O, Kopp Kallner H, et al. Perfect-use and typical-use Pearl Index of a contraceptive mobile app. *Contraception* 2017; 96: 420-5.

63. Kleinschmidt TK, Bull JR, Lavorini V, et al. Advantages of determining the fertile window with the individualised Natural Cycles algorithm over calendar-based methods. *Eur J Contracept Reprod Health Care* 2019; 24: 457-63.

64. Bull JR, Rowland SP, Berglund Scherwitzl E, et al. Real-world menstrual cycle characteristics of more than 600,000 menstrual cycles. *NPJ Dig Med* 2019; 2: 1-8.

65. Berglund Scherwitzl E, Lundberg O, Kopp Kallner H, et al. Short- and long-term effect of contraceptive methods on fecundity. *Eur J Contracept Reprod Health Care* 2019; 24: 260-5.

66. Torous J, Andersson G, Bertagnoli A, et al. Towards a consensus around standards for smartphone apps and digital mental health. *World Psychiatry* 2019; 18: 97-8.