

Telemedizinische Nachbetreuung

Evidenzbericht zur S3-Leitlinie Schwangerschaftsabbruch
im ersten Trimenon

A decorative horizontal bar spanning the width of the page, composed of 18 squares of varying shades of blue and grey. The word 'EVIDENZBERICHT' is centered in white text on a dark blue background.

EVIDENZBERICHT

Projekt: V21-12E

Version: 1.1

Stand: 27.01.2023

IQWiG-Berichte – Nr. 1500

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema

Telemedizinische Nachbetreuung: Evidenzbericht zur S3-Leitlinie Schwangerschaftsabbruch im ersten Trimenon

Auftraggeber

Bundesministerium für Gesundheit

Datum des Auftrags

01.10.2021

Interne Projektnummer

V21-12E

Anschrift des Herausgebers

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8
50670 Köln

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

E-Mail: berichte@iqwig.de

Internet: www.iqwig.de

ISSN: 1864-2500

Dieser Bericht wurde ohne die Beteiligung externer Sachverständiger erstellt.

Für die Inhalte des Berichts ist allein das IQWiG verantwortlich.

Mitarbeiterinnen und Mitarbeiter des IQWiG

- Julia Gowik
- Tatjana Hermanns
- Claudia Mischke
- Kevin Rudolf
- Andrea Tasar

Schlagwörter

Telemedizin, Abort – Eingeleiteter, Systematische Übersicht

Keywords

Telemedicine, Abortion – Induced, Systematic Review

Seite

- iii -

5.4.2	Ergebnisse zum kritischen Endpunkt Befolgung der Nachsorgestrategie	25
5.4.3	Evidenzprofil für den wichtigen Endpunkt ungeplante Besuche oder Telefonanrufe in einer medizinischen Einrichtung	26
5.4.4	Evidenzprofil für den wichtigen Endpunkt unvollständiger Schwangerschaftsabbruch mit der Notwendigkeit eines gynäkologisch- chirurgischen Eingriffs	28
6	Anmerkungen zu den Ergebnissen	29
7	Literatur	31
8	Studienlisten	34
8.1	Liste der identifizierten systematischen Übersichten.....	34
8.2	Liste der ausgeschlossenen Publikationen mit Ausschlussgründen	34
Anhang A	Bewertung der Qualität der Informationsbeschaffung der Basis-SÜ aus der fokussierten Recherche	35
Anhang B	Bewertung der methodischen Qualität der berücksichtigten systematischen Übersicht aus der fokussierten Recherche	36
Anhang C	Gegenüberstellung der Ergebnisse.....	37
Anhang D	Bewertung der Studienlimitationen.....	38
Anhang E	Suchstrategien	40
E.1	Bibliografische Datenbanken	40
E.2	Studienregister.....	43

Tabellenverzeichnis

	Seite
Tabelle 1: Übersicht über die Kriterien für den Studieneinschluss.....	5
Tabelle 2: Berücksichtigte systematische Übersicht.....	15
Tabelle 3: Studienpool der Evidenzdarstellung.....	17
Tabelle 4: Charakteristika der in die Evidenzdarstellung eingeschlossenen Studien	18
Tabelle 5: Charakterisierung der Interventionen in den eingeschlossenen Studien	19
Tabelle 6: Ein- / Ausschlusskriterien für Teilnehmerinnen in den Studien.....	21
Tabelle 7: Charakterisierung der Studienpopulationen.....	21
Tabelle 8: Matrix der Endpunkte.....	22
Tabelle 9: Evidenzprofil für den kritischen Endpunkt Zufriedenheit der Frauen mit der Behandlung.....	23
Tabelle 10: Evidenzprofil für den kritischen Endpunkt Befolgung der Nachsorgestrategie	25
Tabelle 11: Evidenzprofil für den wichtigen Endpunkt ungeplante Besuche oder Telefonanrufe in einer medizinischen Einrichtung.....	26
Tabelle 12: Evidenzprofil für den wichtigen Endpunkt unvollständiger Schwangerschaftsabbruch mit der Notwendigkeit eines gynäkologisch-chirurgischen Eingriffs	28
Tabelle 13: Bewertung der Informationsbeschaffung zur systematischen Übersicht NICE 2019.....	35
Tabelle 14: Bewertung der methodischen Qualität der Basis-SÜ NICE 2019	36
Tabelle 15: Gegenüberstellung der Effektschätzungen von der Basis-SÜ mit Endler 2022	37
Tabelle 16: Studienlimitationen zum Endpunkt Zufriedenheit der Frauen mit der Behandlung.....	38
Tabelle 17: Studienlimitationen zum Endpunkt Befolgung der Nachsorgestrategie durch die Frauen	38
Tabelle 18: Studienlimitationen zum Endpunkt ungeplante Besuche oder Telefonanrufe in einer medizinischen Einrichtung	39
Tabelle 19: Studienlimitationen zum Endpunkt unvollständiger Schwangerschaftsabbruch mit der Notwendigkeit eines gynäkologisch-chirurgischen Eingriffs	39

Abbildungsverzeichnis

	Seite
Abbildung 1: Ergebnis der Informationsbeschaffung	16

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften
BMG	Bundesministerium für Gesundheit
DGGG	Deutsche Gesellschaft für Gynäkologie und Geburtshilfe e. V.
DVG	Digitale-Versorgung-Gesetz
G-BA	Gemeinsamer Bundesausschuss
GRADE	Grading of Recommendations Assessment, Development and Evaluation (Bewertung von Empfehlungen, Entwicklung und Auswertung)
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
KI	Konfidenzintervall
MWD	Mittelwertdifferenz
OR	Odds Ratio
PICO	Participants Intervention Comparison Outcomes (Akronym für Studienteilnehmer, Interventionsgruppe, Vergleichsgruppe, Endpunkte)
p. m.	post menstruationem (nach der letzten Menstruation)
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
RD	absolute Risikodifferenz
RR	relatives Risiko
SÜ	systematische Übersicht

1 Hintergrund

Auf Basis des Digitale-Versorgung-Gesetzes (DVG) kann die Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) dem Bundesministerium für Gesundheit (BMG) Leitlinien vorschlagen, bei denen das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) gemäß SGB V (§§ 139a Abs. 3 Nr. 3, 139b Abs. 6) deren Entwicklung oder Aktualisierung mit Evidenzrecherchen unterstützen soll [1].

Der vorliegende Auftrag umfasst die Beantwortung von Fragestellungen zur Weiterentwicklung der interdisziplinären S2k-Leitlinie „Schwangerschaftsabbruch im ersten Trimenon“ (Registernummer der AWMF: 015-094) [2] zu einer S3-Leitlinie.

2 Fragestellung

Ziel des Evidenzberichts ist die Darstellung der Evidenz bezüglich der Effekte einer telemedizinischen Nachbetreuung im Vergleich zur persönlichen ärztlichen Nachuntersuchung oder einer anderen telemedizinischen Nachbetreuung bei Frauen, bei denen ein medikamentöser Schwangerschaftsabbruch bis zur Schwangerschaftswoche 10 + 0 nach Eintritt der letzten Menstruation (p. m.) mit Blutungsinduktion zu Hause durchgeführt wird.

3 Projektverlauf

3.1 Zeitlicher Verlauf des Projekts

Das IQWiG wurde am 01.10.2021 vom BMG beauftragt, die Deutsche Gesellschaft für Gynäkologie und Geburtshilfe e. V. (DGGG) bei der Weiterentwicklung einer interdisziplinären S2k-Leitlinie „Schwangerschaftsabbruch im ersten Trimenon“ [2] zu einer S3-Leitlinie zu unterstützen.

Insgesamt wurden 8 Population-Intervention-Comparison-Outcome(PICO)-Fragen von den Leitlinienkoordinatorinnen und -koordinatoren in Abstimmung mit Patientenvertreterinnen und -vertretern und mit Beratung durch die AWMF formuliert, zu denen das IQWiG jeweils auf Basis einer internen Projektskizze einen Evidenzbericht erstellt. Die Projektbearbeitung im IQWiG begann mit dem Kick-off-Treffen, an dem 1 Leitlinienkoordinator und Ansprechpartnerinnen und -partner des IQWiG teilnahmen. Das Kick-off-Treffen hat am 05.07.2022 stattgefunden. Die Evidenzberichte wurden an die Koordinatorinnen und Koordinatoren der Leitlinienerstellung sowie an die zuständige Ansprechpartnerin für die Leitlinie bei der AWMF übermittelt.

Nach Abschluss aller Evidenzberichte für diesen Auftrag wurden diese zusammen an das BMG übermittelt und 4 Wochen später auf der Website des IQWiG veröffentlicht.

4 Methoden

4.1 Spezifizierungen und Änderungen im Projektverlauf

Es wurde eine aktuelle Basis-SÜ identifiziert (NICE 2019 [3]) und folgende Spezifizierungen im Projektverlauf vorgenommen:

- 1) Spezifizierung in Abschnitt 4.2.3: Neben der qualitativ hochwertigen Informationsbeschaffung wies die Basis-SÜ NICE 2019 eine ausreichende methodische Qualität auf (Wert von 6 Punkten im Oxman-Guyatt-Index [4], siehe Anhang B) und es wurden Evidenzprofile nach den methodischen Vorgaben von Grading of Recommendations Assessment, Development and Evaluation (GRADE) erstellt. Deshalb wurde ein Abgleich der Effektschätzungen aus der Basis-SÜ mit den Effektschätzungen aus den neu identifizierten Primärstudien durchgeführt (siehe Abschnitt 6 und Anhang C). Da die Ergebnisse sich augenscheinlich nicht widersprachen, wurden im Evidenzbericht nur die neu identifizierten Studien dargestellt.
- 2) Spezifizierung in Abschnitt 4.2.2.1: Die eingeschlossenen Studien aus der Basis-SÜ NICE 2019 wurden nicht in den Studienpool des Evidenzberichts aufgenommen. Es wurden lediglich die Effektschätzungen aus der Basis-SÜ extrahiert, um sie mit den Ergebnissen aus den neu identifizierten Studien abzugleichen.
- 3) Spezifizierung in Abschnitt 4.2.2.3: Die Suche wurde auf das Publikationsdatum ab 2018 eingeschränkt.

Redaktionelle Änderung in der Version 1.1: Der Titel des Berichts wurde korrigiert.

4.2 Methodik gemäß Projektskizze

Die folgenden Abschnitte geben den Wortlaut der Berichtsmethodik aus der Projektskizze wieder. Über diese Methodik hinausgehende Spezifizierungen oder Änderungen der Methoden im Projektverlauf werden in Abschnitt 4.1 erläutert. Im folgenden Text wird an den entsprechenden Stellen auf diesen Abschnitt verwiesen.

Die PICO-Fragestellung wird durch die Darstellung von Evidenzprofilen beantwortet. Die Erstellung der Evidenzprofile erfolgt auf Grundlage der methodischen Vorgaben von GRADE [5] und ist konform zu den Allgemeinen Methoden [6]. Das PICO-Schema wurde von der Leitliniengruppe zur Weiterentwicklung der interdisziplinären S2k-Leitlinie „Schwangerschaftsabbruch im ersten Trimenon“ (Registernummer der AWMF: 015-094) [2] zu einer S3-Leitlinie festgelegt.

4.2.1 Tabellarische Darstellung der Kriterien für den Studieneinschluss

In der folgenden Tabelle 1 sind die Kriterien aufgelistet, die Studien erfüllen müssen, um in die Evidenzdarstellung eingeschlossen zu werden.

Tabelle 1: Übersicht über die Kriterien für den Studieneinschluss

Einschlusskriterien	
E1	Population: Frauen, bei denen ein medikamentöser Schwangerschaftsabbruch bis zur Schwangerschaftswoche 10 + 0 p. m. mit Blutungsinduktion zu Hause durchgeführt wird
E2	Prüfintervention: telemedizinische Nachbetreuung
E3	Vergleichsintervention: persönliche ärztliche Nachuntersuchung oder eine andere telemedizinische Nachbetreuung
E4	kritische Endpunkte: nicht diagnostizierte fortbestehende Schwangerschaft, Zufriedenheit der Frauen mit der Behandlung wichtige Endpunkte: Befolgung der Nachsorgestrategie durch die Frauen, ungeplante Besuche oder Telefonanrufe in einer medizinischen Einrichtung, unvollständiger Schwangerschaftsabbruch mit der Notwendigkeit eines gynäkologisch-chirurgischen Eingriffs
E5	Studientypen: RCTs
E6	Publikationssprache: Deutsch oder Englisch
E7	Vollpublikation verfügbar ^a
a. Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [7] oder ein Bericht über die Studie, der den Kriterien des CONSORT-Statements [8] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zur Studienmethodik und zu den Studienergebnissen nicht vertraulich sind. CONSORT: Consolidated Standards of Reporting Trials; ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; p.m.: nach der letzten Menstruation; RCT: randomisierte kontrollierte Studie	

4.2.1.1 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen

Für die Einschlusskriterien E1 (Population), E2 (Prüfintervention, bezogen auf die Interventionsgruppe der Studie) und E3 (Vergleichsintervention, bezogen auf die Vergleichsgruppe der Studie) reicht es aus, wenn bei mindestens ca. 80 % der eingeschlossenen Personen diese Kriterien erfüllt sind. Liegen für solche Studien Subgruppenanalysen für Personen vor, die die Einschlusskriterien erfüllen, wird auf diese Analysen zurückgegriffen. Studien, bei denen die Einschlusskriterien E1, E2 und E3 bei weniger als ca. 80 % erfüllt sind, können ggf. unberücksichtigt bleiben.

4.2.2 Informationsbeschaffung

4.2.2.1 Fokussierte Informationsbeschaffung von systematischen Übersichten

Parallel zur Erstellung der Projektskizze erfolgte eine Recherche nach systematischen Übersichten in der Datenbank MEDLINE (umfasst auch die Cochrane Database of Systematic Reviews), HTA Database sowie auf den Websites des National Institute for Health and Care Excellence (NICE) und der Agency for Healthcare Research and Quality (AHRQ).

Die Suche fand am 20.07.2022 statt. Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Anhang E.

Die Selektion erfolgte durch 1 Person und wurde anschließend von einer 2. Person überprüft. Diskrepanzen wurden durch Diskussion zwischen beiden aufgelöst.

Es wird geprüft, ob mindestens 1 hochwertige und aktuelle systematische Übersicht infrage kommt, deren Informationsbeschaffung als Grundlage für die Evidenzdarstellung verwendet werden kann (im Folgenden: Basis-SÜ). Dafür erfolgt eine Bewertung der Qualität der Informationsbeschaffung dieser systematischen Übersicht(en). Kann mindestens 1 diesbezüglich hochwertige und aktuelle Basis-SÜ identifiziert werden, werden die zugrunde liegenden Studien beziehungsweise Dokumente von 1 Person auf deren Relevanz für die vorliegende Evidenzdarstellung geprüft und das Ergebnis von einer 2. Person überprüft. Bewertungen der Studienlimitationen bzw. der Studienqualität der eingeschlossenen Studien oder die Datenextraktion aus den systematischen Übersichten werden nicht übernommen.

Zu diesem Vorgehen gab es eine Spezifizierung im Projektverlauf, siehe Abschnitt 4.1.

Die finale Entscheidung, welche systematische(n) Übersicht(en) als Basis-SÜ für die Evidenzdarstellung herangezogen werden, erfolgt nach Erstellung der Projektskizze anhand der darin festgelegten Kriterien. In jedem Fall werden die Referenzlisten der identifizierten systematischen Übersichten hinsichtlich relevanter Primärstudien gesichtet (siehe Abschnitt 4.2.2.2).

4.2.2.2 Fokussierte Informationsbeschaffung von Studien

Für den Fall, dass mindestens 1 systematische Übersicht als Basis-SÜ für die Informationsbeschaffung des Evidenzberichts verwendet werden kann (siehe Abschnitt 4.2.2.1), wird diese für die Informationsbeschaffung von Studien für den von der Übersicht abgedeckten Zeitraum herangezogen. Dieser Teil der Informationsbeschaffung wird ergänzt um eine systematische Recherche nach relevanten Studien beziehungsweise Dokumenten für den nicht von der Übersicht abgedeckten Zeitraum. Gegebenenfalls wird auf eine ergänzende fokussierte Informationsbeschaffung von Studien ganz verzichtet.

Für den Fall, dass keine Basis-SÜ identifiziert werden kann, findet eine systematische Recherche für den gesamten relevanten Zeitraum statt.

Folgende primäre und weitere Informationsquellen sowie Suchtechniken werden dabei berücksichtigt:

Primäre Informationsquellen

- bibliografische Datenbanken
 - MEDLINE
 - Cochrane Central Register of Controlled Trials

- Studienregister (es erfolgt eine Einschränkung auf Studienergebnisse)
 - U.S. National Institutes of Health. ClinicalTrials.gov
 - World Health Organization. International Clinical Trials Registry Platform Search Portal

Weitere Informationsquellen und Suchtechniken

- Anwendung weiterer Suchtechniken
 - Sichten von Referenzlisten identifizierter systematischer Übersichten
 - Identifizieren von Studienregistereinträgen zu eingeschlossenen Studien
- Autorenanfrage(n)

4.2.2.3 Anwendung von Limitierungen auf Datenbankebene

Fokussierte Informationsbeschaffung von systematischen Übersichten

Hinsichtlich des Publikationszeitraums gab es keine Einschränkung. Die MEDLINE-Suchstrategie enthielt Limitierungen auf deutsch- und englischsprachige Publikationen [6] sowie auf Humanstudien.

Fokussierte Informationsbeschaffung von Studien

Hinsichtlich des Publikationszeitraums gibt es keine Einschränkung.

Zu diesem Vorgehen hat es eine Spezifizierung im Projektverlauf, siehe Abschnitt 4.1.

Mit der MEDLINE-Suchstrategie werden folgende Publikationstypen ausgeschlossen: Kommentare und Editorials, da diese i. d. R. keine Studien enthalten [9]. Außerdem enthalten die Suchstrategien Limitierungen auf deutsch- und englischsprachige Publikationen [6] sowie auf Humanstudien (MEDLINE). In der Cochrane Central Register of Controlled Trials Suche werden Einträge aus Studienregistern ausgeschlossen.

4.2.2.4 Selektion relevanter Studien

Selektion relevanter Studien bzw. Dokumente aus den Ergebnissen der bibliografischen Recherche

Duplikate werden mit Hilfe von EndNote entfernt. Die in bibliografischen Datenbanken identifizierten Treffer werden in einem 1. Schritt anhand ihres Titels und, sofern vorhanden, Abstracts in Bezug auf ihre potenzielle Relevanz bezüglich der Einschlusskriterien bewertet. Als potenziell relevant erachtete Dokumente werden in einem 2. Schritt anhand ihres Volltextes auf Relevanz geprüft. Beide Schritte erfolgen durch 2 Personen unabhängig voneinander. Diskrepanzen werden durch Diskussion zwischen beiden aufgelöst.

Selektion relevanter Studien bzw. Dokumente aus weiteren Informationsquellen

Rechercheergebnisse aus den folgenden Informationsquellen werden von 1 Person auf Studien gesichtet:

- Studienregister
- Referenzlisten identifizierter systematischer Übersichten

Die identifizierten Studien werden dann auf ihre Relevanz geprüft. Der gesamte Prozess wird anschließend von einer 2. Person überprüft. Diskrepanzen werden durch Diskussion zwischen beiden aufgelöst.

4.2.3 Informationsdarstellung und Synthese

Zu diesem Vorgehen gab es eine Spezifizierung, siehe Abschnitt 4.1.

4.2.3.1 Darstellung der Einzelstudien

Alle für den Evidenzbericht notwendigen Informationen werden aus den Unterlagen zu den eingeschlossenen Studien in standardisierte Tabellen extrahiert. Ergeben sich im Abgleich der Informationen aus unterschiedlichen Dokumenten zu einer Studie (aber auch aus multiplen Angaben zu einem Aspekt innerhalb eines Dokumentes selbst) Diskrepanzen, die auf die Interpretation der Ergebnisse erheblichen Einfluss haben könnten, wird dies an den entsprechenden Stellen des Berichts dargestellt.

Die Extraktion erfolgt durch eine Person und wird von einer 2. Person auf Grundlage der Studien kontrolliert.

Die Ergebnisse zu den in den Studien berichteten kritischen und wichtigen Endpunkten werden im Bericht vergleichend beschrieben.

Das Vorgehen zu der Bewertung des Einflusses der Studienlimitationen auf die berichtsrelevanten Ergebnisse wird in Abschnitt 4.2.3.2 endpunktspezifisch pro Studie beschrieben. Nach dieser Bewertung werden die Informationen zu den ausgewählten Endpunkten in Evidenzprofilen zusammengeführt und analysiert [10]. Wenn möglich werden über die Gegenüberstellung der Ergebnisse der Einzelstudien hinaus die in Abschnitt 4.2.3.3 beschriebenen Verfahren eingesetzt.

Für binäre Daten wird als relative Effektschätzung primär das Odds Ratio (OR) herangezogen. Dabei wird im Fall von 0 Ereignissen in einem Studienarm bei der Berechnung von Effekt und Konfidenzintervall (KI) der Korrekturfaktor 0,5 in beiden Studienarmen verwendet. Für absolute Effekte wird vorrangig die absolute Risikodifferenz (RD) verwendet. Das Basisrisiko wird in der Regel durch den Median des Risikos der Kontrollgruppe in den eingeschlossenen Einzelstudien bestimmt. Auf Grundlage dieses angenommenen Basisrisikos werden mithilfe

der relativen Gesamtschätzung der Metaanalyse das absolute Risiko in der Interventionsgruppe und die Risikodifferenz errechnet. Zur Abschätzung der quantitativen Ergebnisunsicherheit werden ergänzend die oberen und unteren KI-Grenzen der RD basierend auf den KI-Grenzen der relativen Gesamtschätzung der Metaanalyse berechnet. Wird keine Metaanalyse durchgeführt, erfolgt eine Berechnung der RD pro Studie durch die Risiken in den beiden Behandlungsgruppen; als zugehöriges KI wird standardmäßig dasjenige nach der Wilson-Score-Methode [11] angegeben. Sollte das Ergebnis des CSZ-Tests [12] qualitativ nicht zum Ergebnis des KI nach der Wilson-Score-Methode, sondern zum Wald-KI passen, wird dieses angegeben.

Für stetige Daten wird primär die Mittelwertdifferenz (MWD) herangezogen. Falls notwendig (z. B. wenn verschiedene Skalen gepoolt werden sollen oder dies zur Bewertung der Qualität der Evidenz benötigt wird), wird außerdem eine standardisierte Mittelwertdifferenz (SMD, z. B. Hedges' g) angegeben.

Ergebnisse können ggf. im Evidenzbericht unberücksichtigt bleiben, wenn ein großer Anteil der in die Auswertung eigentlich einzuschließenden Personen nicht in der Auswertung berücksichtigt worden ist. Für die Entscheidung hierüber wird sich an einem Anteil von ca. 70 % orientiert, die in der Auswertung mindestens berücksichtigt sein sollten.

Die Ergebnisse können ggf. auch dann unberücksichtigt bleiben, wenn der Unterschied der Anteile nicht berücksichtigter Personen zwischen den Gruppen sehr groß ist. Als Orientierung für die Entscheidung dient ein Unterschied von 15 Prozentpunkten.

Die festgelegten Grenzen von 70 % und 15 Prozentpunkte können in begründeten Einzelfällen projektspezifisch verändert werden.

4.2.3.2 Bewertung der Studienlimitationen (Verzerrungspotenzial)

Der Einfluss möglicher Studienlimitationen auf die Ergebnisse wird zunächst endpunktspezifisch pro Studie insbesondere anhand der unten aufgeführten Kriterien bewertet. Dazu erfolgt jeweils eine Bewertung mit „adäquat“, „unklar“ oder „nicht adäquat“. Die Voraussetzungen für eine Bewertung mit „adäquat“, sowie die wesentlichen Aspekte welche zu einer Bewertung mit „nicht adäquat“ führen können, werden im Folgenden für jedes Kriterium erläutert. Eine Bewertung mit „unklar“ erfolgt grundsätzlich dann, wenn keine bzw. keine ausreichenden Angaben zur adäquaten Bewertung zur Verfügung stehen. Im Evidenzprofil werden abschließend die Studienlimitationen studienübergreifend und endpunktspezifisch durch die Bewertungen „keine“, „schwerwiegend“, „sehr schwerwiegend“ dargestellt.

- adäquate Erzeugung der Randomisierungssequenz

Voraussetzung für eine Bewertung mit „adäquat“ ist, dass aus der Beschreibung der Studie hervorgeht, dass die Gruppenzuteilung rein zufällig erfolgte und die Erzeugung der Zuteilungssequenz beschrieben und geeignet ist. Die Antwortmöglichkeit „nicht adäquat“ besteht im Falle eines Evidenzberichts auf Basis von RCTs nicht, da die Studie in diesem Fall als nicht randomisierte Studie aufgefasst, ausgeschlossen und nicht weiter bewertet wird.

- Verdeckung der Gruppenzuteilung

Voraussetzung für eine Bewertung mit „adäquat“ ist, dass aus der Beschreibung der Studie hervorgeht, dass die Zuteilung der Patientinnen und Patienten in die verschiedenen Studiengruppen den Personen, die die Gruppenzuteilung veranlassen oder über den Einschluss der Patientinnen und Patienten entscheiden, nicht bekannt ist. Die Antwortmöglichkeit „nicht adäquat“ besteht im Falle eines Evidenzberichts auf Basis von RCTs nicht, da die Studie in diesem Fall als nicht randomisierte Studie aufgefasst, ausgeschlossen und nicht weiter bewertet wird.

- Verblindung

Voraussetzung für eine Bewertung mit „adäquat“ ist, dass die zu verblindene Personengruppe (Patientinnen und Patienten, behandelnde Personen und Endpunkterheberin und Endpunkterheber, Datenanalytistinnen und Datenanalysten, Auswerterinnen und Auswerter) explizit benannt und die Methode, die eine Verblindung gewährleistet, beschrieben und geeignet ist.

- Umsetzung des Intention-to-treat(ITT)-Prinzips

Voraussetzung für eine Bewertung mit „adäquat“ ist, dass 1) randomisierten Patientinnen und Patienten entsprechend ihrer Gruppenzuteilung ausgewertet und 2) der Anteil der fehlenden Werte gering ist oder sie in geeigneter Weise berücksichtigt wurden.

Regelhaft erfolgt eine Bewertung mit „adäquat“, wenn der Anteil der nicht in der Auswertung berücksichtigten Patientinnen und Patienten insgesamt kleiner als 10 % oder der Unterschied der Anteile zwischen den Gruppen kleiner als 5 Prozentpunkte ist. Wurden fehlende Werte in geeigneter Weise berücksichtigt, hängt die Bewertung des Ergebnisses von vielen Aspekten ab: u. a. der Häufigkeit fehlender Werte, der Art des Auswertungsverfahrens, der Gründe für das Auftreten fehlender Werte.

- ergebnisunabhängige Berichterstattung

Voraussetzung für eine Bewertung mit „adäquat“ ist, dass der Endpunkt gemäß einer prospektiven Planung (in der Regel im Studienregistereintrag) operationalisiert, erfasst und

ausgewertet wird oder relevante Abweichungen von den Studienautorinnen und Studienautoren plausibel erklärt werden.

4.2.3.3 Metaanalysen

Die geschätzten Effekte und KIs aus den Studien werden mittels Forest Plots zusammenfassend dargestellt. Anschließend wird die Heterogenität des Studienpools anhand des statistischen Tests auf Vorliegen von Heterogenität [13] untersucht. Ergibt der Heterogenitätstest ein statistisch nicht signifikantes Ergebnis ($p \geq 0,05$), wird davon ausgegangen, dass die Schätzung eines gemeinsamen (gepoolten) Effekts sinnvoll ist. Im Fall von mindestens 5 Studien erfolgt die Metaanalyse mithilfe des Modells mit zufälligen Effekten gemäß der Methode nach Knapp und Hartung unter Verwendung des Heterogenitätsschätzers nach Paule und Mandel [14]. Als Ergebnis wird der gemeinsame Effekt inklusive KI dargestellt. Weil die Heterogenität im Fall weniger Studien nicht verlässlich geschätzt werden kann, werden bei 4 oder weniger Studien gegebenenfalls Modelle mit festem Effekt verwendet. Sind in einer Metaanalyse nur 2 Studien enthalten, so wird aus pragmatischen Gründen regelhaft ein Modell mit festem Effekt gewählt, wenn keine deutlichen Gründe dagegensprechen. Bei 3 und 4 Studien erfolgt vorzugsweise eine Metaanalyse nach der Methode von Knapp und Hartung, sofern das KI informativ ist und kein Widerspruch zu einer qualitativen Zusammenfassung besteht. Andernfalls wird ein Modell mit festem Effekt verwendet, wenn keine deutlichen Gründe dagegensprechen. Ist ein Modell mit festem Effekt nicht vertretbar, bleibt es bei der Schätzung nach Knapp und Hartung.

Ergibt der Heterogenitätstest ein statistisch signifikantes Ergebnis ($p < 0,05$), wird untersucht, welche Faktoren diese Heterogenität möglicherweise verursachen. Dazu zählen methodische und klinische Faktoren, sogenannte Effektmodifikatoren. Kann die Heterogenität durch diese Faktoren erklärt werden, wird das Poolen der Effekte ggf. auf Subgruppen beschränkt [15]. Kann die Heterogenität nicht erklärt werden, so wird, sofern möglich, ebenfalls ein gemeinsamer (gepoolter) Effekt berechnet. Dabei ist zu beachten, dass ein gepoolter Effekt bei bedeutsamer Heterogenität möglicherweise nicht sinnvoll interpretierbar ist. Von der Berechnung eines gepoolten Effekts wird abgesehen, falls sich die KIs der eingehenden Studien nur gering oder gar nicht überlappen und gleichzeitig die Effektschätzungen der Studien in unterschiedliche Richtungen weisen.

4.2.3.4 Bewertung der Qualität der Evidenz

Alle für den Evidenzbericht relevanten Ergebnisse werden hinsichtlich der Beeinflussung durch Studienlimitationen, des Risikos von Publikationsbias, der Genauigkeit und Konsistenz der Effekte, der Übertragbarkeit und ggf. hinsichtlich möglicher Aspekte zur Aufwertung der Qualität der Evidenz überprüft. Für jeden für die Entscheidung kritischen und wichtigen Endpunkt wird eine endpunktbezogene studienübergreifende Aussage zur Qualität der Evidenz in 4 Abstufungen bezüglich des jeweiligen Vertrauens in die Effektschätzung getroffen

und diese in eine der Stufen hohe, moderate, niedrige und sehr niedrige Qualität der Evidenz entsprechend der GRADE-Guideline eingeteilt [16,17]:

- Das Vertrauen in die Effektschätzung ist hoch: Eine hohe Qualität der Evidenz bedeutet, dass der wahre Effekt sehr sicher nahe bei der Effektschätzung liegt.
- Das Vertrauen in die Effektschätzung ist mäßig: Eine moderate Qualität der Evidenz bedeutet, dass der wahre Effekt wahrscheinlich nahe bei der Effektschätzung liegt, aber die Möglichkeit besteht, dass er relevant verschieden ist.
- Das Vertrauen in die Effektschätzung ist begrenzt: Eine niedrige Qualität der Evidenz bedeutet, dass der wahre Effekt durchaus relevant verschieden zur Effektschätzung sein kann.
- Das Vertrauen in die Effektschätzung ist sehr gering: Eine sehr niedrige Qualität der Evidenz bedeutet, dass der wahre Effekt wahrscheinlich relevant verschieden von der Effektschätzung ist.

Die Bewertung erfolgt durch 2 Personen unabhängig voneinander. Diskrepanzen werden durch Diskussion zwischen beiden aufgelöst.

Ergebnissen aus RCTs wird im 1. Bewertungsschritt eine hohe, Ergebnissen aus Beobachtungsstudien eine niedrige Qualität der Evidenz attestiert [16]. Von dieser rein designbedingten Einschätzung ausgehend, wird bei Studienlimitationen, einem Risiko von Publikationsbias, fehlender Genauigkeit der Effekte, inkonsistenten Effekten oder eingeschränkter Übertragbarkeit die Qualität der Evidenz abgewertet sowie bei (sehr) großen Effekten, einer Dosis-Wirkungs-Beziehung oder, wenn die Berücksichtigung aller potenziellen Confounder zu einer Effekterhöhung / Verstärkung des beobachteten Ergebnisses führen würde, aufgewertet. Unterstützend für diese Bewertungsschritte können Sensitivitätsanalysen durchgeführt werden. Alle Ab- und Aufwertungen der Qualität der Evidenz werden durch Fußnoten gekennzeichnet.

4.2.3.4.1 Abwertung der Qualität der Evidenz

A: Studienlimitationen

Die Qualität der Evidenz wird aufgrund von schwerwiegenden oder sehr schwerwiegenden Studienlimitationen um 1 oder 2 Stufen abgewertet, wenn starke Limitierungen in einem oder mehreren der in Abschnitt 4.2.3.2 genannten Kriterien das Vertrauen in die Effektschätzung beeinträchtigen. Die endpunktbezogene studienübergreifende Bewertung der Studienlimitationen erfolgt unter Berücksichtigung des Einflusses der einzelnen Studien auf die Effektschätzung für jeden Endpunkt [18].

B: Inkonsistente (heterogene) Effekte

Die Qualität der Evidenz wird bei Inkonsistenz (unerklärter Heterogenität) zwischen Studienergebnissen um 1 oder 2 Stufen abgewertet. Die Einschätzung einer möglichen Heterogenität erfolgt anhand der in Abschnitt 4.2.3.3 beschriebenen Kriterien.

Falls für binäre Daten Metaanalysen für 2 Effektmaße durchgeführt wurden und eine der beiden einen statistisch signifikanten Gruppenunterschied anzeigt, wird die Inkonsistenz anhand dieser Metaanalyse bewertet. Ist der Gruppenunterschied für beide Effektmaße nicht statistisch signifikant, wird die Bewertung in der Regel anhand des relativen Effektmaßes vorgenommen.

Bei stetigen Daten wird die Bewertung in der Regel anhand der Metaanalyse der Mittelwertdifferenz vorgenommen. Falls die fehlende Genauigkeit anhand der standardisierten Mittelwertdifferenz bewertet wurde, wird die Inkonsistenz auch anhand dieser bewertet.

Eine Abwertung um 1 Stufe erfolgt, wenn der Heterogenitätstest ein signifikantes Ergebnis liefert. Um 2 Stufen kann z. B. abgewertet werden, wenn aufgrund der Heterogenität der Studienergebnisse keine gepoolte Effektschätzung berechnet wird.

C: Indirektheit

Die Qualität der Evidenz wird bei starker oder sehr starker Indirektheit um 1 oder 2 Stufen abgewertet. Indirektheit kann auf Abweichungen zwischen dem PICO und den Einschlusskriterien der Studien basieren [19].

D: Publikationsbias

Die Qualität der Evidenz wird um 1 Stufe abgewertet, wenn ein Publikationsbias wahrscheinlich oder sehr wahrscheinlich ist.

Kriterien für die Annahme eines Publikationsbias umfassen eine Evidenzgrundlage, welche vorrangig auf kleinen Studien und wenigen Ereignissen beruht, welche einen beobachteten Zusammenhang zwischen Größe der Effektschätzung und Studiengröße (oder Genauigkeit der Effekte) zeigt, welche Studienabbrüche auf der Grundlage von hohen Effektschätzungen aufweist oder welche Auffälligkeiten in der gesichteten Datenlage erkennen lässt [20].

E: Fehlende Genauigkeit der Effekte

Die Qualität der Evidenz wird wegen geringer oder sehr geringer Genauigkeit der Effekte um 1 oder 2 Stufen abgewertet. Maßgeblich hierfür ist neben der Lage und Breite des 95 %-KI auch die Wahl des Metaanalysemodells (zufällige Effekte oder fester Effekt) (siehe Abschnitt 4.2.3.3). Außerdem können sehr kleine Fallzahlen zu einer Abwertung wegen fehlender Genauigkeit führen [21]. Bei ausreichender Patientenzahl wird regelhaft nicht abgewertet,

wenn das KI eines Effekts (relativ oder absolut im Fall binärer Daten bzw. Mittelwertdifferenz oder standardisierte Mittelwertdifferenz im Fall stetiger Daten) einen signifikanten Unterschied anzeigt.

Andernfalls erfolgt eine Abwertung wegen fehlender Präzision, falls das KI (des relativen Effekts bei binären Daten bzw. Hedges' g bei stetigen Daten) einen mittelgroßen Effekt und gleichzeitig auch den Nulleffekt überdeckt. Bei binären Daten wird unter einem mittelgroßen Effekt eine Halbierung oder eine Verdopplung (0,5 oder 2) des relativen Effekts verstanden. Bei stetigen Daten wird von einem mittelgroßen Effekt ausgegangen, wenn das KI für eine standardisierte Mittelwertdifferenz die Grenze 0,5 oder -0,5 umfasst. In diesen Fällen, in denen ein mittelgroßer Effekt und gleichzeitig auch der Nulleffekt überdeckt wird, sollte eine Abwertung wegen fehlender Genauigkeit um 1 Stufe in Betracht gezogen werden. Eine Abwertung um 2 Stufen wird in Fällen in Betracht gezogen, in denen die zuvor genannten KI-Grenzen beide, d. h. z. B. im Fall binärer Daten sowohl 0,5 als auch 2 im KI enthalten sind.

4.2.3.4.2 Aufwertung der Qualität der Evidenz

A: Große Effekte

Die Qualität der Evidenz wird in methodisch guten, nicht randomisierten Interventionsstudien bei großen bzw. sehr großen Effekten um 1 oder 2 Stufen aufgewertet. Eine Aufwertung um 1 Stufe ist bei großen Effekten möglich, wenn beispielsweise das relative Risiko zwischen 2 und 5 bzw. zwischen 0,2 und 0,5 liegt, ohne dass plausible Confounder diesen Effekt verringern könnten. Eine Aufwertung um 1 weitere Stufe kann bei sehr großen Effekten erfolgen, wenn beispielsweise das relative Risiko über 5 bzw. unter 0,2 liegt, keine relevanten Verzerrungsrisiken vorliegen und die Effekte präzise (mit hinreichend schmalen KIs) geschätzt werden konnten [22].

B: Dosis-Wirkungs-Beziehung

Die Qualität der Evidenz wird bei einer Dosis-Wirkungs-Beziehung aufgewertet [22].

C: Confounding

Die Qualität der Evidenz wird bei einer möglichen Verstärkung des beobachteten Ergebnisses unter Berücksichtigung von potenziellen Confoundern oder Verzerrungsquellen aufgewertet [22].

5 Ergebnisse

5.1 Ergebnisse der Informationsbeschaffung

Abbildung 1 zeigt das Ergebnis der Informationsbeschaffung inklusive Studienselektion gemäß den Kriterien für den Studieneinschluss. Die Suchstrategien für die Suche in bibliografischen Datenbanken (letzte Suche am 06.09.2022) und die Suche in Studienregistern (letzte Suche am 19.07.2022) sind in Anhang E dargestellt.

Die Referenzen der als Volltexte geprüften, jedoch ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Abschnitt 8.2.

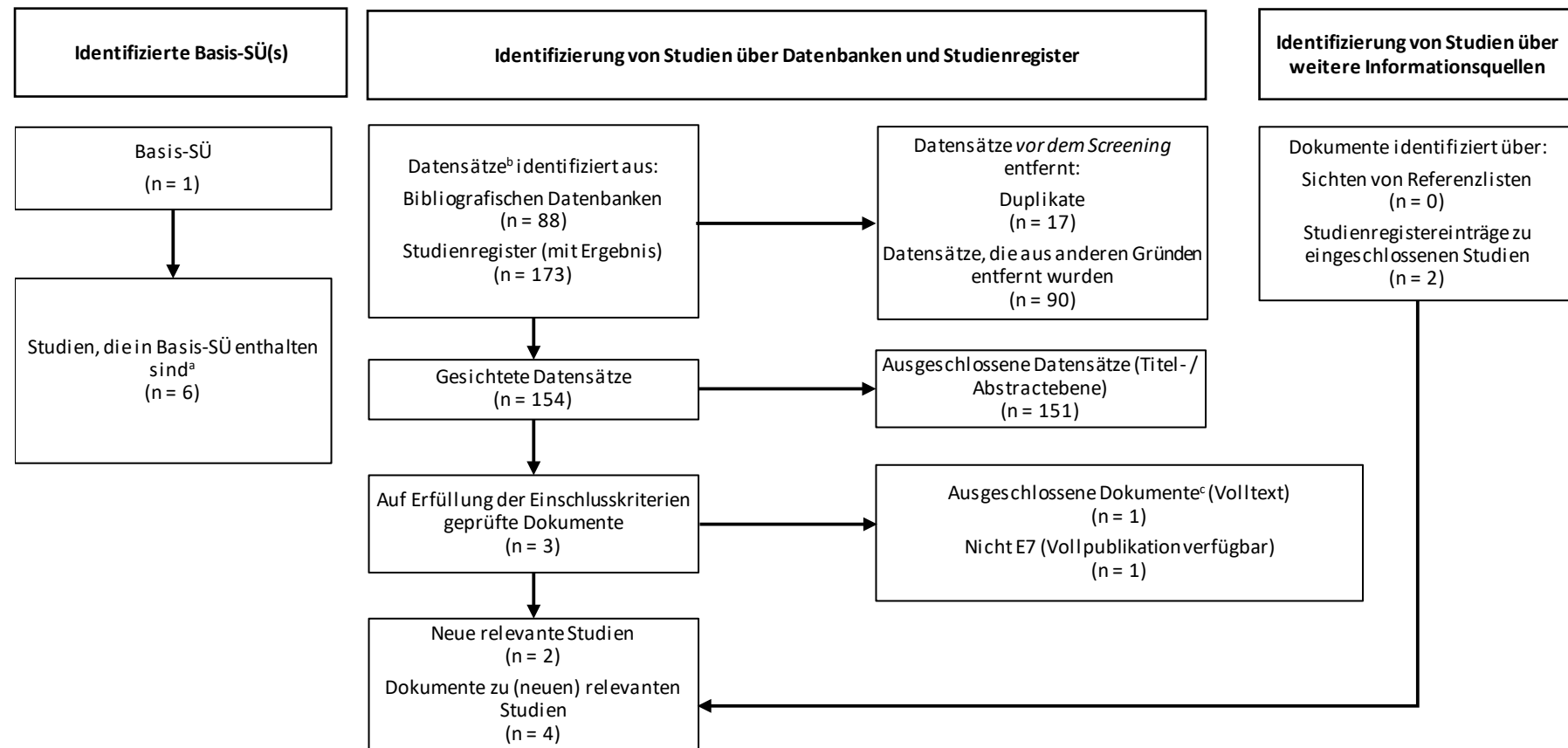
Autorenanfragen wurden nicht gestellt.

Von den 3 identifizierten systematischen Übersichten (zu 1 SÜ lagen 2 Publikationen vor, siehe Abschnitt 8.1) wurde 1 systematische Übersicht (siehe Tabelle 2) als Basis-SÜ berücksichtigt. Die Bewertung der Qualität der Informationsbeschaffung dieser Basis-SÜ findet sich in Anhang A, die Bewertung der methodischen Qualität ist in Anhang B dargestellt.

Tabelle 2: Berücksichtigte systematische Übersicht

Dokument	Vollpublikation
NICE 2019	ja [3,23]

Es erfolgte eine ergänzende fokussierte Suche nach Studien in bibliografischen Datenbanken für den Zeitraum (ab 2018), der nicht durch die berücksichtigte Basis-SÜ abgedeckt war. Für alle weiteren Informationsquellen wurde ohne zeitliche Beschränkung gesucht.



- a. Die Primärstudien aus der Basis-SÜ werden im Evidenzbericht nicht dargestellt und gehen daher nicht in den Studienpool ein.
b. Datensatz: Titel oder Abstract (oder beides) einer Studie, der in einer Datenbank (z. B. MEDLINE) oder auf einer Website indiziert ist.
c. nicht gelistete Ein- und Ausschlusskriterien (n = 0)

Abbildung 1: Ergebnis der Informationsbeschaffung

Resultierender Studienpool

Durch die verschiedenen Rechenschritte konnten insgesamt 2 neue relevante Studien identifiziert werden (siehe auch Tabelle 3).

Die Primärstudien der Basis-SÜ fließen nicht in den Studienpool ein, da beim Vergleich der Effektschätzungen der NICE 2019 mit den neu identifizierten Studien keine widersprüchlichen Ergebnisse identifiziert wurden (siehe Anhang C).

Tabelle 3: Studienpool der Evidenzdarstellung

Studie	Verfügbare Dokumente	
	Vollpublikation (in Fachzeitschriften)	Registereintrag / Ergebnisbericht aus Studienregistern
Telemedizinische Nachbetreuung vs. persönliche ärztliche Nachuntersuchung		
Endler 2022	ja [24]	ja [25] / nein
Telemedizinische Nachbetreuung vs. andere telemedizinische Nachbetreuung		
Fok 2021	ja [26]	ja [27] / nein

5.2 Charakteristika der in die Evidenzdarstellung eingeschlossenen Studien

Studiendesign und Studienpopulationen

In diesem Abschnitt werden die eingeschlossenen Studien einschließlich der wesentlichen Ein- und Ausschlusskriterien für die Teilnehmerinnen sowie die Studienpopulationen charakterisiert.

Dabei handelt es sich um 2 RCTs, die in Südafrika und den USA durchgeführt wurden (siehe Tabelle 4). Davon wurde in 1 Studie der Vergleich einer telemedizinischen Nachbetreuung mit der üblichen persönlichen Nachbehandlung untersucht [24]. Die telemedizinische Nachbetreuung wurde durch eine asynchrone medizinische Online-Beratung mit Instruktionen zur Durchführung des medikamentösen Aborts umgesetzt. Insgesamt wurden in dieser Studie 900 Frauen randomisiert.

Zum Vergleich einer telemedizinischen Nachbetreuung mit einer anderen telemedizinischen Nachbetreuung wurde 1 weitere RCT identifiziert [26]. In der Studie wurden 88 Frauen randomisiert.

Es wurde ein Schwangerschaftstest mit niedriger Sensitivität mit einem Multi-Level-Schwangerschaftstest verglichen. Die Ergebnisse der Schwangerschaftstests sowie weiterer Ereignisse (z. B. Blutungen und Schwangerschaftssymptome) wurden von den Teilnehmerinnen über einen Online-Fragebogen an das klinische Studienpersonal gesendet, das die Einträge auf behandlungspflichtige Besonderheiten prüfte. Somit unterscheidet sich der telemedizinische Anteil der Interventionen zwischen den beiden Gruppen nicht und

entspricht streng genommen nicht den formalen Einschlusskriterien des Berichts. Die Studie wurde dennoch berücksichtigt, da die Basis-SÜ NICE 2019 eine Primärstudie mit sehr vergleichbarem Design eingeschlossen hat (Primärstudie Blum 2016 [28]) und in diesem Evidenzbericht ein Update der NICE 2019 vorgenommen wurde.

Eine detaillierte Beschreibung der Interventionen wird in Tabelle 5 dargestellt.

Tabelle 4: Charakteristika der in die Evidenzdarstellung eingeschlossenen Studien

Studie	Studien- design	Personen- zahl (randomi- siert) N	Intervention	Vergleich	Ort und Zeitraum der Durchführung	Dauer der Nach- beobachtung
Telemedizinische Nachbetreuung vs. persönliche ärztliche Nachuntersuchung						
Endler 2022	RCT	I: 450 C: 450	asynchrone medizinische Online- Beratung mit Instruktionen für den Abort	TAU	Südafrika; 2020 bis 2021	6 Wochen
Telemedizinische Nachbetreuung vs. andere telemedizinische Nachbetreuung						
Fok 2021	RCT	I: 43 C: 45	Schwangerschaftstest mit niedriger Sensitivität und Evaluation über Online-Fragebogen	Multi-Level- Schwanger- schaftstest und Evaluation über Online- Fragebogen	USA; 2017 bis 2018	2 Wochen
C: Vergleichsgruppe; I: Interventionsgruppe; N: Anzahl randomisierter Personen; RCT: randomisierte kontrollierte Studie; TAU: übliche Behandlung						

Tabelle 5: Charakterisierung der Interventionen in den eingeschlossenen Studien
(mehrsseitige Tabelle)

Studie	Intervention	Vergleich	Beide Gruppen
Telemedizinische Nachbetreuung vs. persönliche ärztliche Nachuntersuchung			
Endler 2022	Telemedizin Telemedizinische Umsetzung: - asynchrone medizinische Beratung und Instruktionen für den medikamentösen Abort über Online-Kommunikationssystem - abdominale oder bimanuale Palpation des Uterus vor dem Abort durch eine Pflegekraft als Sicherheitsmaßnahme Ablauf: - Überprüfung der Eignung für einen medikamentösen Abort über Online-Fragebogen mit anschließender Prüfung der Inhalte durch Studienärztinnen und -ärzte - Einnahme von Mifepriston und Misoprostol zu Hause - Online-Messenger stand als Beratungsstelle während des gesamten Prozesses zur Verfügung - eigenständige Durchführung eines Schwangerschaftstests nach 3 bis 4 Wochen	TAU Ablauf: - persönliche Überprüfung der Eignung für einen medikamentösen Abort - persönliche medizinische Beratung und Instruktionen für den medikamentösen Abort - Ultraschalluntersuchung zur Bestimmung einer intrauterinen Schwangerschaft und der Schwangerschaftswoche - Einnahme von Mifepriston in der Klinik, Einnahme von Misoprostol zu Hause - Vereinbarung eines nachfolgenden Untersuchungstermins in der Klinik oder eigenständige Durchführung eines Schwangerschaftstests nach 6 Wochen	Medikamentöser Abort: Mifepriston (200 mg, orale Einnahme) und Misoprostol (200 µg, sublinguale Einnahme) nach 24 bis 48 Stunden

Tabelle 5: Charakterisierung der Interventionen in den eingeschlossenen Studien
(mehrsseitige Tabelle)

Studie	Intervention	Vergleich	Beide Gruppen
Telemedizinische Nachbetreuung vs. andere telemedizinische Nachbetreuung			
Fok 2021	Schwangerschaftstest mit niedriger Sensitivität Sensitivitätsgrenze (beta-hCG): 1000 mIU/mL Ablauf: - Durchführung des Tests am 7. Tag nach dem Abort - bei positivem Testergebnis: Durchführung des zweiten Tests am 14. Tag nach dem Abort	Multi-Level Schwangerschaftstest Sensitivitätsgrenzen (beta-hCG): 25–99 mIU/mL 100–499 mIU/mL 500–1999 mIU/mL 2000–9999 mIU/mL > 10 000 mIU/mL Ablauf: - Durchführung des Tests am 7. Tag nach dem Abort - wenn beta-hCG-Wert nicht um 1 Stufe gesunken ist: Durchführung des zweiten Tests am 14. Tag nach dem Abort	Medikamentöser Abort: - Mifepriston (200 mg, Einnahme in der Klinik) und Misoprostol (200 µg, bukkale Einnahme zu Hause) nach 6 bis 48 Stunden Telemedizinische Umsetzung: - Erhebung der Ergebnisse und bestehender Schwangerschafts-symptome über einen Online-Fragebogen - Prüfung der ausgefüllten Fragebögen durch eine Ärztin / einen Arzt - zusätzlicher Nachuntersuchungstermin bei starken Blutungen oder dem Verdacht auf eine fortbestehende Schwangerschaft Ablauf: - Frauen erhalten Bildanweisungen - Studienkoordinatorinnen und –koordinatoren - erklären Verwendung und Interpretation der Tests - führen zusammen mit den Frauen vor Beginn des Aborts den jeweiligen Test durch - Ausfüllen des Online-Fragebogens am 7. und 14. Tag nach dem Abort
hCG: humanes Choriongonadotropin; mIU/mL: milli internationale Einheiten / Milliliter; TAU: übliche Behandlung			

Ein- und Ausschlusskriterien für Teilnehmerinnen in den Studien

Die wesentlichen Ein- und Ausschlusskriterien der Studien sind in Tabelle 6 dargestellt.

Tabelle 6: Ein- / Ausschlusskriterien für Teilnehmerinnen in den Studien

Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Telemedizinische Nachbetreuung vs. persönliche ärztliche Nachuntersuchung		
Endler 2022	- Frauen im Alter von ≥ 18 Jahren - Anstreben eines medizinischen Aborts bis einschließlich zur 9. Schwangerschaftswoche	- Erhalt der üblichen Behandlung vor Studienbeginn - nicht im Besitz eines Smartphones oder nicht in der Lage, die App runterzuladen > 9. Schwangerschaftswoche - medizinische Kontraindikation für einen Abort zu Hause
Telemedizinische Nachbetreuung vs. andere telemedizinische Nachbetreuung		
Fok 2021	- Alter ≥ 18 Jahre - Einwilligung zum medikamentösen Abort - Einwilligung zur Kontaktaufnahme via Telefon oder E-Mail	Schwangerschaft > 63 Tage

Die Charakterisierung der Studienpopulation wird in Tabelle 7 dargestellt.

Tabelle 7: Charakterisierung der Studienpopulationen

Gruppe Studie	Intervention	Vergleich
Telemedizinische Nachbetreuung vs. persönliche ärztliche Nachuntersuchung		
Endler 2022		
N	450	450
Alter [Jahre] Median (IQR)	28 (24; 32)	28 (25; 33)
Gestationsalter [Tage] Median (IQR)	48 (42; 55)	47 (42; 54)
Vorheriger Abbruch (%)	18	18
Telemedizinische Nachbetreuung vs. andere telemedizinische Nachbetreuung		
Fok 2021		
N	43	45
Alter [Jahre] Median (Min; Max)	27 (18; 44)	29 (18; 44)
Gestationsalter [Tage] Median (Min; Max)	44 (34; 60)	45 (37; 59)
Vorheriger Abbruch (%)	44	44
IQR: Interquartilsabstand; Min: Minimum; Max: Maximum; N: Anzahl randomisierter Personen (Werte, die auf anderen Personenzahlen basieren, werden in der entsprechenden Zeile gekennzeichnet, wenn Abweichungen relevant)		

5.3 Übersicht der Endpunkte

Aus 2 Studien konnten Daten zu kritischen und wichtigen Endpunkten extrahiert werden. Die Tabelle 8 zeigt die Übersicht der verfügbaren Daten zu Endpunkten aus den eingeschlossenen Studien. Zu dem Endpunkt nicht diagnostizierte fortbestehende Schwangerschaft wurden in keiner Studie Daten berichtet.

Tabelle 8: Matrix der Endpunkte

Studie	Kritische Endpunkte		Wichtige Endpunkte		
	nicht diagnostizierte fortbestehende Schwangerschaft	Zufriedenheit der Frauen mit der Behandlung	Befolgung der Nachsorgestrategie durch die Frauen	ungeplante Besuche oder Telefonanrufe in einer medizinischen Einrichtung	unvollständiger Schwangerschaftsabbruch mit Notwendigkeit eines gynäkologisch-chirurgischen Eingriffs
Endler 2022	–	●	●	●	–
Fok 2021	–	●	–	●	●
●: Daten wurden berichtet und waren verwertbar. –: Es wurden keine bzw. keine für den Evidenzbericht verwertbare Daten berichtet.					

5.4 Ergebnisse zu Endpunkten

5.4.1 Ergebnisse zum kritischen Endpunkt Zufriedenheit der Frauen mit der Behandlung

Tabelle 9: Evidenzprofil für den kritischen Endpunkt Zufriedenheit der Frauen mit der Behandlung

Faktoren der Qualität der Evidenz							Anzahl der Personen mit Ereignis/Anzahl der Personen		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikationsbias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD [95 %-KI]	
Telemedizinische Nachbetreuung versus persönliche ärztliche Nachbetreuung												
Zufriedenheit – Nachbeobachtungszeitpunkt 6 Wochen												
Interpretation der Effektschätzung: Ein OR größer als 1 bzw. eine positive RD bedeuten einen Vorteil für die Interventionsgruppe.												
RCT, 1 [24]	schwerwiegend ^d	nicht zutreffend	nicht schwerwiegend	unentdeckt	schwerwiegend ^e	keine	369/370	342/347	99	5,39 [0,63; 46,41]	1 [0; 3]	niedrig
Telemedizinische Nachbetreuung versus andere telemedizinische Nachbetreuung												
Präferenz – Nachbeobachtungszeitpunkt 14 Tage												
Interpretation der Effektschätzung: Ein OR größer als 1 bzw. eine positive RD bedeuten einen Vorteil für die Interventionsgruppe.												
RCT, 1 [26]	sehr schwerwiegend ^f	nicht zutreffend	schwerwiegend ^g	unentdeckt	sehr schwerwiegend ^h	keine	32/36	42/44	95	0,38 [0,07; 2,21]	-7 [-21; 6]	sehr niedrig

Tabelle 9: Evidenzprofil für den kritischen Endpunkt Zufriedenheit der Frauen mit der Behandlung

Faktoren der Qualität der Evidenz							Anzahl der Personen mit Ereignis/Anzahl der Personen		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD [95 %-KI]	
a. Risiko der Vergleichsgruppe der eingeschlossenen Studie b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden. c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang D zu entnehmen. d. Die Verblindung und die Umsetzung des ITT-Prinzips waren nicht adäquat. e. Das 95 %-KI des OR überdeckt 1 und 2. Somit können weder Effekte zuungunsten noch mittelgroße Effekte zugunsten der Prüfintervention ausgeschlossen werden. f. Die Randomisierung und die Verdeckung der Gruppenzuteilung waren unklar. Die Verblindung, die Umsetzung des ITT Prinzips und die ergebnisunabhängige Berichterstattung waren nicht adäquat. g. Die Studie berichtet den Endpunkt Zufriedenheit nicht direkt. Die indirekte Operationalisierung Präferenz wurde aufgrund der Vergleichbarkeit zur Operationalisierung in der NICE 2019 ausgewählt. h. Das 95 %-KI des OR überdeckt 0,5 und 2. Somit können weder mittelgroße Effekte zugunsten noch mittelgroße Effekte zuungunsten der Prüfintervention ausgeschlossen werden. C: Vergleichsgruppe; I: Prüfinterventionsgruppe; ITT: Intention to treat; KI: Konfidenzintervall; N: Anzahl der Studien; OR: Odds Ratio, RCT: randomisierte kontrollierte Studie, RD: absolute Risikodifferenz												

5.4.2 Ergebnisse zum kritischen Endpunkt Befolgung der Nachsorgestrategie

Tabelle 10: Evidenzprofil für den kritischen Endpunkt Befolgung der Nachsorgestrategie

Faktoren der Qualität der Evidenz							Anzahl der Personen mit Ereignis/Anzahl der Personen		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikationsbias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD [95 %-KI]	
Telemedizinische Nachbetreuung versus persönliche ärztliche Nachbetreuung												
Adhärenz – Nachbeobachtungszeitpunkt 5 Tage												
Interpretation der Effektschätzung: Ein OR größer als 1 bzw. eine positive RD bedeuten einen Vorteil für die Interventionsgruppe.												
RCT, 1 [24]	schwerwiegend	nicht zutreffend	nicht schwerwiegend	unentdeckt	schwerwiegend	keine	358/378	340/354	96	0,74 [0,37; 1,48]	-1 [-5; 2]	niedrig
a. Risiko der Vergleichsgruppe der eingeschlossenen Studie b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden. c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang D zu entnehmen. d. Die Verblindung und die Umsetzung des ITT-Prinzips waren nicht adäquat. e. Das 95 %-KI des OR überdeckt 0,5 und 1. Somit können weder mittelgroße Effekte zuungunsten noch Effekte zugunsten der Prüfintervention ausgeschlossen werden.												
C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; OR: Odds Ratio, RCT: randomisierte kontrollierte Studie, RD: absolute Risikodifferenz												

5.4.3 Evidenzprofil für den wichtigen Endpunkt ungeplante Besuche oder Telefonanrufe in einer medizinischen Einrichtung

Tabelle 11: Evidenzprofil für den wichtigen Endpunkt ungeplante Besuche oder Telefonanrufe in einer medizinischen Einrichtung (mehrsseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Personen mit Ereignis/Anzahl der Personen		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD [95 %-KI]	
Telemedizinische Nachbetreuung versus persönliche ärztliche Nachbetreuung												
Hospitalisierung – Nachbeobachtungszeitpunkt 5 Tage												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative RD bedeuten einen Vorteil für die Interventionsgruppe.												
RCT, 1 [24]	schwer-wiegend ^d	nicht zutreffend	nicht schwer-wiegend	unentdeckt	sehr schwer-wiegend ^e	keine	3/377	2/359	1	1,43 [0,36; 8,62]	0 [-1; 2]	sehr niedrig
Aufsuchen einer Notaufnahme – Nachbeobachtungszeitpunkt 5 Tage												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative RD bedeuten einen Vorteil für die Interventionsgruppe.												
RCT, 1 [24]	schwer-wiegend ^d	nicht zutreffend	nicht schwer-wiegend	unentdeckt	sehr schwer-wiegend ^e	keine	6/373	5/350	1	1,13 [0,34; 3,74]	0 [-2; 2]	sehr niedrig
Aufsuchen einer Klinik – Nachbeobachtungszeitpunkt 6 Wochen												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative RD bedeuten einen Vorteil für die Interventionsgruppe.												
RCT, 1 [24]	schwer-wiegend ^f	nicht zutreffend	nicht schwer-wiegend	unentdeckt	nicht schwer-wiegend	keine	42/369	34/350	10	1,19 [0,74; 1,93]	2 [-3; 6]	moderat

Tabelle 11: Evidenzprofil für den wichtigen Endpunkt ungeplante Besuche oder Telefonanrufe in einer medizinischen Einrichtung (mehrsseitige Tabelle)

Faktoren der Qualität der Evidenz							Anzahl der Personen mit Ereignis/Anzahl der Personen		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD [95 %-KI]	
Telemedizinische Nachbetreuung versus andere telemedizinische Nachbetreuung												
Kontaktaufnahme mit einem Arzt												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative RD bedeuten einen Vorteil für die Interventionsgruppe.												
RCT, 1 [26]	sehr schwerwiegend ^g	nicht zutreffend	nicht schwerwiegend	unentdeckt	schwerwiegend ^h	keine	11/39	18/45	40	0,59 [0,24; 1,48]	-12 [-30; 8]	sehr niedrig
Aufsuchen einer Klinik oder Notfallambulanz												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative RD bedeuten einen Vorteil für die Interventionsgruppe.												
RCT, 1 [26]	sehr schwerwiegend ^h	nicht zutreffend	nicht schwerwiegend	unentdeckt	sehr schwerwiegend ^e	keine	3/39	3/45	7	1,17 [0,22; 6,14]	1 [-11; 14]	sehr niedrig
a. Risiko der Vergleichsgruppe der eingeschlossenen Studie												
b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.												
c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang D zu entnehmen.												
d. Die Verblindung und die Umsetzung des ITT-Prinzips waren nicht adäquat.												
e. Das 95 %-KI des OR überdeckt 0,5 und 2. Somit können weder mittelgroße Effekte zugunsten noch mittelgroße Effekte zuungunsten der Prüfintervention ausgeschlossen werden.												
f. Die Verblindung, die Umsetzung des ITT-Prinzips und die ergebnisunabhängige Berichterstattung waren nicht adäquat.												
g. Die Randomisierung und die Verdeckung der Gruppenzuteilung waren unklar. Die Verblindung, die Umsetzung des ITT-Prinzips und die ergebnisunabhängige Berichterstattung waren nicht adäquat.												
h. Das 95 %-KI des OR überdeckt 0,5 und 1. Somit können weder mittelgroße Effekte zugunsten noch Effekte zuungunsten der Prüfintervention ausgeschlossen werden.												
C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; OR: Odds Ratio, RCT: randomisierte kontrollierte Studie, RD: absolute Risikodifferenz												

5.4.4 Evidenzprofil für den wichtigen Endpunkt unvollständiger Schwangerschaftsabbruch mit der Notwendigkeit eines gynäkologisch-chirurgischen Eingriffs

Tabelle 12: Evidenzprofil für den wichtigen Endpunkt unvollständiger Schwangerschaftsabbruch mit der Notwendigkeit eines gynäkologisch-chirurgischen Eingriffs

Faktoren der Qualität der Evidenz							Anzahl der Personen mit Ereignis/Anzahl der Personen		Basisrisiko in % ^a	Effekt		Qualität der Evidenz ^b
Studien-design; N	Studien-limitationen ^c	Inkonsistenz	Indirektheit	Publikations-bias	Fehlende Genauigkeit	Andere Faktoren	[I]	[C]		OR [95 %-KI]	RD [95 %-KI]	
Telemedizinische Nachbetreuung versus andere telemedizinische Nachbetreuung												
Durchführung einer Vakuumaspiration												
Interpretation der Effektschätzung: Ein OR kleiner als 1 bzw. eine negative RD bedeuten einen Vorteil für die Interventionsgruppe.												
RCT, 1 [26]	sehr schwerwiegend ^d	nicht zutreffend	schwerwiegend ^e	unentdeckt	sehr schwerwiegend ^f	keine	2/39	3/45	7	0,76 [0,12; 4,78]	−2 [−13; 11]	sehr niedrig
a. Risiko der Vergleichsgruppe der eingeschlossenen Studie												
b. Unter Qualität der Evidenz wird die studienübergreifende endpunktbezogene Qualität der Evidenz verstanden.												
c. Die Bewertung der Studienlimitationen pro Studie pro Endpunkt ist Anhang D zu entnehmen.												
d. Die Randomisierung und die Verdeckung der Gruppenzuteilung waren unklar. Die Umsetzung des ITT-Prinzips und die ergebnisunabhängige Berichterstattung waren nicht adäquat.												
e. Die Gründe für die Durchführung der Vakuumaspiration werden nicht genannt.												
f. Das 95 %-KI des OR überdeckt 0,5 und 2. Somit können werden mittelgroße Effekte zugunsten noch mittelgroße Effekte zuungunsten der Prüfintervention ausgeschlossen werden.												
C: Vergleichsgruppe; I: Prüfinterventionsgruppe; KI: Konfidenzintervall; N: Anzahl der Studien; OR: Odds Ratio, RCT: randomisierte kontrollierte Studie, RD: absolute Risikodifferenz												

6 Anmerkungen zu den Ergebnissen

Für die untersuchte Fragestellung zum Vergleich einer telemedizinischen Nachbetreuung mit der üblichen persönlichen Nachbehandlung oder einer anderen telemedizinischen Nachbetreuung bei Frauen, bei denen ein medikamentöser Schwangerschaftsabbruch bis zur Schwangerschaftswoche 10 + 0 durchgeführt wurde, wurden in einer Update-Recherche der NICE 2019 2 RCTs identifiziert.

Die Basis-SÜ NICE 2019 hat insgesamt 6 RCTs eingeschlossen, wovon 4 RCTs dem Vergleich telemedizinische Nachbetreuung versus persönliche ärztliche Nachbehandlung und 2 RCTs dem Vergleich telemedizinische Nachbetreuung versus andere telemedizinische Nachbetreuung zugeordnet wurden. Die Qualität der Evidenz in der SÜ war für alle untersuchten Endpunkte niedrig und sehr niedrig. Eine Ausnahme stellen die Endpunkte Zufriedenheit der Frauen mit der Behandlung und ungeplante Besuche oder Telefonanrufe in der medizinischen Einrichtung dar, die mit moderat bewertet wurden.

Vergleich: Telemedizinische Nachbetreuung versus persönliche ärztliche Nachbehandlung

Für den Vergleich einer telemedizinischen mit einer persönlichen ärztlichen Nachbehandlung wurde in der Update-Recherche 1 neue RCT (Endler 2022) mit 900 Teilnehmerinnen identifiziert. Aus dieser RCT konnten lediglich die Ergebnisse zu den Endpunkten Befolgung der Nachsorgestrategie und ungeplante Besuche oder Telefonanrufe in der medizinischen Einrichtung (Operationalisierung: Aufsuchen einer Klinik) mit den Ergebnissen aus dem relevanten Vergleich der NICE 2019 verglichen werden. Es zeigen sich für beide Endpunkte sowohl in der neu identifizierten Studie Endler 2022 als auch in der NICE 2019 keine signifikanten Gruppenunterschiede (siehe Anhang C) und die Punktschätzungen sowie die zugehörigen 95 %-KI lagen in einem ähnlichen Bereich. Die Ergebnisse widersprechen sich somit nicht. Die Qualität der Evidenz der neu identifizierten Studie Endler 2022 war für die beiden Endpunkte niedrig.

Für den Endpunkt Zufriedenheit der Frauen konnte kein Vergleich mit Endler 2022 erfolgen, da in der NICE 2019 mit der Präferenz bei einem zukünftigen Schwangerschaftsabbruch eine indirekte Operationalisierung ausgewählt wurde und keine der eingeschlossenen Studien in der SÜ Daten zur Zufriedenheit mit der Behandlung berichteten. In Endler 2022 werden diese Daten ebenfalls berichtet. Sie können allerdings nicht mit den Daten aus der NICE 2019 zusammen betrachtet werden, da die Frauen zusätzlich als Präferenz „beide Nachbetreuungen gemischt“ angeben konnten. In den Studien der NICE 2019 gab es nur die Auswahl zwischen einer der beiden Nachbetreuungen oder keiner Präferenz als Antwortmöglichkeit. Die Studie Endler 2022 berichtete jedoch Daten zu der direkten Operationalisierung Zufriedenheit, die im Bericht dargestellt wurden und mit einer niedrigen Qualität der Evidenz bewertet wurden.

Da die Studie Endler 2022 in Südafrika durchgeführt wurde, bestehen mögliche Einschränkungen bei der Übertragung der Ergebnisse auf den deutschen Versorgungskontext. Dies betrifft zum Beispiel die Lebenssituation der untersuchten Population: Ein substanzieller Anteil der Frauen lebte unter sehr einfachen sozioökonomischen Bedingungen, ersichtlich dadurch, dass die Hälfte der Frauen gelegentlich, aber regelmäßig Lebensmittelknappheit erfuhr. Zusätzlich wohnte ein Viertel der Frauen in sehr einfachen Gegebenheiten (in der Studie als Hütten bezeichnet). Da der Zugang zu einer persönlichen Nachbetreuung nach einer Abtreibung für Frauen aus solchen ressourcenarmen Gebieten mit höheren Hürden z. B. aufgrund einer längeren Anreisedauer verbunden ist, wird ein Effekt der alternativen telemedizinischen Nachsorge möglicherweise überschätzt.

Vergleich: Telemedizinische Nachbetreuung versus andere telemedizinische Nachbetreuung

Die neue RCT Fok 2021 mit 88 Teilnehmerinnen wurde für den Vergleich der anderen telemedizinischen Nachbetreuung identifiziert. Ziel des Vergleichs ist die Gegenüberstellung der Sicherheit und Effektivität verschiedener telemedizinischer Protokolle. In der Studie Fok 2021 unterschieden sich jedoch lediglich die durchgeführten Schwangerschaftstests in den beiden Gruppen, nicht jedoch die telemedizinischen Anteile. Formal erfüllt die Studie somit nicht die Einschlusskriterien. Sie wurde lediglich aufgrund der Vergleichbarkeit zu einer eingeschlossenen Primärstudie der NICE 2019 (Blum 2016 [28]), für die im vorliegenden Evidenzbericht ein Update vorgenommen wurde, berücksichtigt.

Die Studie Fok 2021 kann mit keiner anderen eingeschlossenen Primärstudie aus dem relevanten Vergleichsarm der NICE 2019 ausgewertet werden. Die Primärstudie Blum 2016 [28] untersucht zwar ebenfalls den Vergleich von 2 verschiedenen Schwangerschaftstests. Jedoch wurde in der Studie der Multi-Level Test mit einem hochsensitiven Test verglichen, während in Fok 2021 der Vergleich mit einem niedrig sensitiven Test vorgenommen wurde. Daher können die Ergebnisse der beiden Studien nicht miteinander verglichen werden.

7 Literatur

1. Bundestag. Gesetz für eine bessere Versorgung durch Digitalisierung und Innovation (Digitale-Versorgung-Gesetz – DVG). Bundesgesetzblatt Teil 1 2019; (49): 2562-2584.
2. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. Sicherer Schwangerschaftsabbruch [online]. 2021 [Zugriff: 17.05.2022]. URL: <https://www.awmf.org/leitlinien/detail/anmeldung/1/II/015-094.html>.
3. National Institute for Health and Care Excellence. Abortion care; Follow-up after medical abortion up to 10+0 weeks; NG140 [online]. 2019 [Zugriff: 17.08.2022]. URL: <https://www.nice.org.uk/guidance/ng140>.
4. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. J Clin Epidemiol 1991; 44(11): 1271-1278. [https://dx.doi.org/10.1016/0895-4356\(91\)90160-b](https://dx.doi.org/10.1016/0895-4356(91)90160-b)
5. Schünemann H, Brożek J, Guyatt G et al. GRADE Handbook [online]. 2013 [Zugriff: 19.02.2021]. URL: <https://gdt.gradeapro.org/app/handbook/handbook.html>.
6. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden; Version 6.1 [online]. 2022 [Zugriff: 27.01.2022]. URL: <https://www.iqwig.de/methoden/allgemeine-methoden-v6-1.pdf>.
7. ICH Expert Working Group. ICH harmonised tripartite guideline: structure and content of clinical study reports; E3 [online]. 1995 [Zugriff: 19.02.2021]. URL: https://database.ich.org/sites/default/files/E3_Guideline.pdf.
8. Moher D, Hopewell S, Schulz KF et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c869. <https://dx.doi.org/10.1136/bmj.c869>.
9. Waffenschmidt S, Navarro-Ruan T, Hobson N et al. Development and validation of study filters for identifying controlled non-randomized studies in PubMed and Ovid MEDLINE. Res Synth Methods 2020; 11(5): 617-626. <https://dx.doi.org/10.1002/jrsm.1425>.
10. Guyatt GH, Oxman AD, Santesso N et al. GRADE guidelines: 12. Preparing summary of findings tables—binary outcomes. J Clin Epidemiol 2013; 66(2): 158-172. <https://dx.doi.org/10.1016/j.jclinepi.2012.01.012>.
11. Newcombe RG. Interval estimation for the difference between independent proportions: comparison of eleven methods. Stat Med 1998; 17(8): 873-890. [https://dx.doi.org/10.1002/\(sici\)1097-0258\(19980430\)17:8<873::aid-sim779>3.0.co;2-i](https://dx.doi.org/10.1002/(sici)1097-0258(19980430)17:8<873::aid-sim779>3.0.co;2-i).
12. Martín Andrés A, Silva Mato A. Choosing the optimal unconditioned test for comparing two independent proportions. Comput Stat Data Anal 1994; 17(5): 555-574. [https://dx.doi.org/10.1016/0167-9473\(94\)90148-1](https://dx.doi.org/10.1016/0167-9473(94)90148-1).

13. Sutton AJ, Abrams KR, Jones DR et al. Methods for meta-analysis in medical research. Chichester: Wiley; 2000.
14. Veroniki AA, Jackson D, Viechtbauer W et al. Recommendations for quantifying the uncertainty in the summary intervention effect and estimating the between-study heterogeneity variance in random-effects meta-analysis. *Cochrane Database Syst Rev* 2015; (10 Suppl 1): 25-27.
15. Guyatt GH, Oxman AD, Kunz R et al. GRADE guidelines: 7. Rating the quality of evidence—inconsistency. *J Clin Epidemiol* 2011; 64(12): 1294-1302.
<https://dx.doi.org/10.1016/j.jclinepi.2011.03.017>.
16. Balshem H, Helfand M, Schunemann HJ et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011; 64(4): 401-406.
<https://dx.doi.org/10.1016/j.jclinepi.2010.07.015>.
17. Meerpohl JJ, Langer G, Perleth M et al. GRADE-Leitlinien: 3. Bewertung der Qualität der Evidenz (Vertrauen in die Effektschätzer). *Z Evid Fortbild Qual Gesundhwes* 2012; 106(6): 449-456. <https://dx.doi.org/10.1016/j.zefq.2012.06.013>.
18. Guyatt GH, Oxman AD, Vist G et al. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). *J Clin Epidemiol* 2011; 64(4): 407-415.
<https://dx.doi.org/10.1016/j.jclinepi.2010.07.017>.
19. Guyatt GH, Oxman AD, Kunz R et al. GRADE guidelines: 8. Rating the quality of evidence—indirectness. *J Clin Epidemiol* 2011; 64(12): 1303-1310.
<https://dx.doi.org/10.1016/j.jclinepi.2011.04.014>.
20. Guyatt GH, Oxman AD, Montori V et al. GRADE guidelines: 5. Rating the quality of evidence—publication bias. *J Clin Epidemiol* 2011; 64(12): 1277-1282.
<https://dx.doi.org/10.1016/j.jclinepi.2011.01.011>.
21. Guyatt GH, Oxman AD, Kunz R et al. GRADE guidelines 6. Rating the quality of evidenced—imprecision. *J Clin Epidemiol* 2011; 64(12): 1283-1293.
<https://dx.doi.org/10.1016/j.jclinepi.2011.01.012>.
22. Guyatt GH, Oxman AD, Sultan S et al. GRADE guidelines: 9. Rating up the quality of evidence. *J Clin Epidemiol* 2011; 64(12): 1311-1316.
<https://dx.doi.org/10.1016/j.jclinepi.2011.06.004>.
23. Schmidt-Hansen M, Cameron S, Lohr PA et al. Follow-up strategies to confirm the success of medical abortion of pregnancies up to 10 weeks' gestation: a systematic review with meta-analyses. *Am J Obstet Gynecol* 2020; 222(6): 551-563.e13.
<https://dx.doi.org/10.1016/j.ajog.2019.11.1244>.

24. Endler M, Petro G, Gemzell Danielsson K et al. A telemedicine model for abortion in South Africa: a randomised, controlled, non-inferiority trial. Lancet 2022; 400(10353): 670-679. [https://dx.doi.org/10.1016/s0140-6736\(22\)01474-x](https://dx.doi.org/10.1016/s0140-6736(22)01474-x).
25. University of Cape Town. Telemedicine for First Trimester Abortion in South Africa [online]. 2021 [Zugriff: 17.10.2022]. URL: <https://ClinicalTrials.gov/show/NCT04336358>.
26. Fok WK, Lerma K, Shaw KA et al. Comparison of two home pregnancy tests for self-confirmation of medication abortion status: A randomized trial. Contraception 2021; 104(3): 296-300. <https://dx.doi.org/10.1016/j.contraception.2021.05.003>.
27. Stanford University. Medical Abortion Self-Confirmation (MASC) [online]. 2019 [Zugriff: 17.10.2022]. URL: <https://ClinicalTrials.gov/show/NCT03148587>.
28. Blum J, Sheldon WR, Ngoc NT et al. Randomized trial assessing home use of two pregnancy tests for determining early medical abortion outcomes at 3, 7 and 14days after mifepristone. Contraception 2016; 94(2): 115-121. <https://dx.doi.org/10.1016/j.contraception.2016.04.001>.
29. Shea BJ, Grimshaw JM, Wells GA et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007; 7: 10. <https://dx.doi.org/10.1186/1471-2288-7-10>.
30. Wong SS, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455.
31. Lefebvre C, Glanville J, Briscoe S et al. Cochrane Handbook for Systematic Reviews of Interventions; Version 6; Technical Supplement to Chapter 4: Searching for and selecting studies [online]. 2019 [Zugriff: 19.02.2021]. URL: <https://training.cochrane.org/handbook/version-6/chapter-4-tech-suppl>.

8 Studienlisten

8.1 Liste der identifizierten systematischen Übersichten

1. Baiju N, Acharya G, D'Antonio F et al. Effectiveness, safety and acceptability of self-assessment of the outcome of first-trimester medical abortion: a systematic review and meta-analysis. BJOG 2019; 126(13): 1536-1544. <https://dx.doi.org/10.1111/1471-0528.15922>.
2. DeNicola N, Grossman D, Marko K et al. Telehealth Interventions to Improve Obstetric and Gynecologic Health Outcomes: A Systematic Review. Obstet Gynecol 2020; 135(2): 371-382. <https://dx.doi.org/10.1097/aog.0000000000003646>.
3. National Institute for Health and Care Excellence. Abortion care; Follow-up after medical abortion up to 10+0 weeks; NG140 [online]. 2019 [Zugriff: 17.08.2022]. URL: <https://www.nice.org.uk/guidance/ng140>.
4. Schmidt-Hansen M, Cameron S, Lohr PA et al. Follow-up strategies to confirm the success of medical abortion of pregnancies up to 10 weeks' gestation: a systematic review with meta-analyses. Am J Obstet Gynecol 2020; 222(6): 551-563.e13. <https://dx.doi.org/10.1016/j.ajog.2019.11.1244>.

8.2 Liste der ausgeschlossenen Publikationen mit Ausschlussgründen

Nicht E7

1. Reynolds-Wright JJ, Norrie J, Cameron ST. UTAH: Using Telemedicine to improve early medical Abortion at Home: a protocol for a randomised controlled trial comparing face-to-face with telephone consultations for women seeking early medical abortion. BMJ Open 2021; 11(6): e046628. <https://dx.doi.org/10.1136/bmjopen-2020-046628>.

Anhang A Bewertung der Qualität der Informationsbeschaffung der Basis-SÜ aus der fokussierten Recherche

Die Bewertung der Qualität der Informationsbeschaffung erfolgte in Anlehnung an AMSTAR, Item 3 [29].

Tabelle 13: Bewertung der Informationsbeschaffung zur systematischen Übersicht NICE 2019

NICE 2019 [3,23]	Bewertung	Erläuterung
1. Wurden mindestens 2 verschiedene Informationsquellen durchsucht (z. B. bibliografische Datenbanken und Studienregister)?	ja	1. Bibliografische Datenbanken 2. Übermittlungen von Stakeholdern (siehe Methods Supplement) 3. Sichtung von Referenzlisten (siehe Methods Supplement)
2. Wurden mindestens 2 verschiedene bibliografische Datenbanken durchsucht?	ja	Medline, Medline In-Process, CCTR, CDSR, DARE, HTA, Embase
3. Werden Suchzeitraum oder Suchdatum genannt?	ja	Appendix A: zeitliche Beschränkung ab 2000 Appendix B: 19th November 2018
4. Werden zumindest die wichtigsten Freitextbegriffe oder Schlagwörter der Suchstrategie genannt?	ja	Darstellung der Suchstrategien im Appendix E
Frage 1.–4. mit ja: umfassend; mindestens 1 Frage mit nein: fraglich	ja	Umfassend, mit folgenden Einschränkungen: ▪ fehlende Suche in Studienregistern ▪ Apply standard animal/non-English language exclusion zeitliche Einschränkung ab 2000

Anhang B Bewertung der methodischen Qualität der berücksichtigten systematischen Übersicht aus der fokussierten Recherche

Die Bewertung der methodischen Qualität erfolgte mit dem Oxman-Guyatt-Index [4]

Tabelle 14: Bewertung der methodischen Qualität der Basis-SÜ NICE 2019

Item	Bewertung
Were the search methods used to find evidence (original research) on the primary question(s) stated?	ja
Was the search for evidence reasonably comprehensive?	ja
Were the criteria used for deciding which studies to include in the overview reported?	ja
Was bias in the selection of studies avoided?	unklar
Were the criteria used for assessing the validity of the included studies reported?	ja
Was the validity of all studies referred to in the text assessed using appropriate criteria (either in selecting studies for inclusion or in analysing the studies that are cited)?	unklar
Were the methods used to combine the findings of the relevant studies (to reach a conclusion) reported?	ja
Were the findings of the relevant studies combined appropriately relative to the primary question the overview addresses?	ja
Were the conclusions made by the author(s) supported by the data and/or analysis reported in the overview?	ja
Gesamtbewertung	6 Punkte

Anhang C Gegenüberstellung der Ergebnisse

In der Tabelle 15 werden lediglich die Effektschätzungen der Operationalisierungen der NICE 2019 und der neu identifizierten Studie Endler 2022 dargestellt, die inhaltlich miteinander verglichen werden konnten.

Tabelle 15: Gegenüberstellung der Effektschätzungen von der Basis-SÜ mit Endler 2022

Endpunkt	NICE-2019 <i>Operationalisierung</i> RR [95 %-KI]	Endler 2022 <i>Operationalisierung</i> RR [95 %-KI]
Befolgung der Nachsorgestrategie durch die Frauen	<i>Adhärenz^a</i> ▪ Studie 1 RR: 0,95 [0,87, 1,03] ▪ Studie 2 RR 1,07 [1,05; 1,1] ▪ Studie 3 RR: 0,99 [0,98; 1,01]	<i>Adhärenz</i> ▪ RR: 0,99 [0,96; 1,02]
Ungeplante Besuche oder Telefonanrufe in der medizinischen Einrichtung	<i>ungeplante Besuche in der medizinischen Einrichtung</i> ▪ RR: 1,20 [0,91; 1,59]	<i>Aufsuchen einer Klinik</i> ▪ RR: 1,17 [0,76; 1,80]
a. Aufgrund bedeutsamer Heterogenität ($I^2 = 93\%$) wurden in der SÜ nur die einzelnen Effektschätzungen der Studien dargestellt. KI: Konfidenzintervall; RR: Relatives Risiko		

Anhang D Bewertung der Studienlimitationen

Tabelle 16: Studienlimitationen zum Endpunkt Zufriedenheit der Frauen mit der Behandlung

Studie	Randomisierung	Verdeckung der Gruppenzuteilung	Verblindung	Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Sonstiges
Telemedizinische Nachbetreuung vs. persönliche ärztliche Nachuntersuchung						
Endler 2022	adäquat	adäquat	nicht adäquat	nicht adäquat	adäquat	adäquat
Telemedizinische Nachbetreuung vs. andere telemedizinische Nachbetreuung						
Fok 2021	unklar	unklar	nicht adäquat	nicht adäquat	nicht adäquat	adäquat
ITT: Intention to treat						

Tabelle 17: Studienlimitationen zum Endpunkt Befolgung der Nachsorgestrategie durch die Frauen

Studie	Randomisierung	Verdeckung der Gruppenzuteilung	Verblindung	Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Sonstiges
Telemedizinische Nachbetreuung vs. persönliche ärztliche Nachuntersuchung						
Endler 2022	adäquat	adäquat	nicht adäquat	nicht adäquat	adäquat	adäquat
ITT: Intention to treat						

Tabelle 18: Studienlimitationen zum Endpunkt ungeplante Besuche oder Telefonanrufe in einer medizinischen Einrichtung

Studie	Randomisierung	Verdeckung der Gruppenzuteilung	Verblindung	Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Sonstiges
Telemedizinische Nachbetreuung vs. persönliche ärztliche Nachuntersuchung						
Endler 2022	adäquat	adäquat	nicht adäquat	nicht adäquat	adäquat / nicht adäquat ^a	adäquat
Telemedizinische Nachbetreuung vs. andere telemedizinische Nachbetreuung						
Fok 2021	unklar	unklar	nicht adäquat	nicht adäquat	nicht adäquat	adäquat
a. abweichende Bewertung für die Operationalisierung Aufsuchen einer Klinik ITT: Intention to treat						

Tabelle 19: Studienlimitationen zum Endpunkt unvollständiger Schwangerschaftsabbruch mit der Notwendigkeit eines gynäkologisch-chirurgischen Eingriffs

Studie	Randomisierung	Verdeckung der Gruppenzuteilung	Verblindung	Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Sonstiges
Telemedizinische Nachbetreuung vs. andere telemedizinische Nachbetreuung						
Fok 2021	unklar	unklar	unklar	nicht adäquat	nicht adäquat	adäquat
ITT: Intention to treat						

Anhang E Suchstrategien

E.1 Bibliografische Datenbanken

Suche nach systematischen Übersichten

1. MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) ALL 1946 to July 19, 2022

Es wurde folgender Filter übernommen:

- Systematische Übersicht: Wong [30] – High specificity strategy

#	Searches
1	exp Self Concept/
2	exp Telemedicine/
3	Postoperative Care/
4	Aftercare/
5	(follow-up or postoperative*).ti,ab.
6	(telecommunicat* or telemed* or telehealth* or teleconsult*).ti,ab.
7	or/1-6
8	exp Abortion Induced/
9	Abortion, Spontaneous/
10	Abortion, Missed/
11	abortion*.ti,ab.
12	(terminat* adj3 pregnanc*).ti,ab.
13	miscarriage*.ti,ab.
14	(early adj pregnancy adj failure*).ti,ab.
15	or/8-14
16	and/7,15
17	Cochrane database of systematic reviews.jn.
18	(search or MEDLINE or systematic review).tw.
19	meta analysis.pt.
20	or/17-19
21	20 not (exp animals/ not humans.sh.)
22	and/16,21
23	22 and (english or german or multilingual or undetermined).lg.

2. International HTA Database

Suchoberfläche: INAHTA

#	Searches
1	"Self Concept" [mhe]
2	"Telemedicine" [mhe]
3	"Postoperative Care" [mh]
4	"Aftercare" [mhe]
5	(follow-up OR postoperative*)[title] OR (follow-up OR postoperative*)[abs]
6	(telecommunicat* OR telemed* OR telehealth* OR teleconsult*)[title] OR (telecommunicat* OR telemed* OR telehealth* OR teleconsult*)[abs]
7	#6 OR #5 OR #4 OR #3 OR #2 OR #1
8	"Abortion Induced"[mhe]
9	"Abortion Spontaneous"[mh]
10	"Abortion Missed"[mh]
11	(abortion*) [title] OR (abortion*) [abs]
12	(terminat* AND pregnanc*) [title] OR (terminat* AND pregnanc*) [abs]
13	(miscarriage*) [title] OR (miscarriage*) [abs]
14	(early AND pregnancy AND failure*) [title] OR (early AND pregnancy AND failure*) [abs]
15	#14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #8
16	#15 AND #7

Suche nach Primärstudien

1. MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) ALL 1946 to September 02, 2022

Es wurden folgende Filter übernommen:

- RCT: Lefebvre [31] – Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision)
- Suchblock zur Intervention (Suchzeilen 5-19): in Anlehnung an NICE 2019 [3]

#	Searches
1	exp Abortion, Induced/
2	abortion*.ti,ab.
3	(terminat* adj3 pregnanc*).ti,ab.
4	1 or 2 or 3
5	exp Pregnancy Tests/

#	Searches
6	((pregnan\$ or LSPT or MSPT or HSPT or MLPT or HLPT or LSUP or MSUP or HSUP) adj test\$).tw.
7	Self-Assessment/
8	*Self Report/
9	Checklist/
10	((self-assess\$ or selfassess\$ or self assess\$ or self-evaluat\$ or selfevaluat\$ or self evaluat\$) adj3 (success or outcome\$ or complet\$ or home or remote)).tw.
11	checklist\$.tw.
12	exp Chorionic Gonadotropin/
13	((beta-hcg\$ or hcg\$) adj (test\$ or level\$ or measurement\$)).tw.
14	exp Telemedicine/
15	(telemed\$ or teleconsult\$).tw.
16	Self Administration/
17	((follow-up or followup or follow up) adj (care or model\$ or procedure\$)).tw.
18	((simple\$ or standard\$ or traditional\$ or mToP) adj (follow-up or followup or follow up)).tw.
19	((in-person\$ or in-clinic\$ or in-office\$ or remote\$ or telephone\$ or ultrasound\$ or ultrasonograph\$ or sonogra\$ or endosonogra\$) adj3 (follow-up or followup or follow up or assess\$)).tw.
20	or/5-19
21	randomized controlled trial.pt.
22	controlled clinical trial.pt.
23	(randomized or placebo or randomly).ab.
24	clinical trials as topic.sh.
25	trial.ti.
26	or/21-25
27	and/4,20,26
28	(animals/ not humans/) or comment/ or editorial/ or exp review/ or meta analysis/ or consensus/ or exp guideline/
29	hi.fs. or case report.mp.
30	or/28-29
31	27 not 30
32	31 and (english or german or multilingual or undetermined).lg.
33	32 and 20181101:3000.(dt).

2. The Cochrane Library

Suchoberfläche: Wiley

- Cochrane Central Register of Controlled Trials: Issue 8 of 12, August 2022

#	Searches
#1	[mh "Abortion, Induced"]
#2	abortion*:ti,ab

#	Searches
#3	(terminat* NEAR/3 pregnanc*):ti,ab
#4	#1 or #2 or #3
#5	[mh "Pregnancy Tests"]
#6	((pregnan* or LSPT or MSPT or HSPT or MLPT or HLPT or LSUP or MSUP or HSUP) NEXT test*):ti,ab
#7	[mh ^"Self-Assessment"]
#8	[mh ^"Self Report"[mj]]
#9	[mh ^"Checklist"]
#10	((self-assess* or selfassess* or self assess* or self-evaluat* or selfevaluat* or self evaluat*) NEAR/3 (success or outcome* or complet* or home or remote)):ti,ab
#11	checklist*:ti,ab
#12	[mh "Chorionic Gonadotropin"]
#13	((beta-hcg* or hcg*) NEXT (test* or level* or measurement*)):ti,ab
#14	[mh "Telemedicine"]
#15	(telemed* or teleconsult*):ti,ab
#16	[mh ^"Self Administration"]
#17	((follow-up or followup or follow up) NEXT (care or model* or procedure*)):ti,ab
#18	((simple* or standard* or traditional* or mToP) NEXT (follow-up or followup or follow up)):ti,ab
#19	((in-person* or in-clinic* or in-office* or remote* or telephone* or ultrasound* or ultrasonograph* or sonogra* or endosonogra*) NEAR/3 (follow-up or followup or follow up or assess*)):ti,ab
#20	#5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19
#21	#4 and #20
#22	(*clinicaltrial*gov* or *trialsearch*who* or *clinicaltrialsregister*eu* or *anzctr*org*au* or *trialregister*nl* or *irct*ir* or *isrctn* or *controlled*trials*com* or *drks*de*):so
#23	#21 not #22
#24	#23 not ((language next (afr or ara or aze or bos or bul or car or cat or chi or cze or dan or dut or es or est or fin or fre or gre or heb or hrv or hun or ice or ira or ita or jpn or ko or kor or lit or nor or peo or per or pol or por or pt or rom or rum or rus or slo or slv or spa or srp or swe or tha or tur or ukr or urd or uzb)) not (language near/2 (en or eng or english or ger or german or mul or unknown)))
#25	#24 with Publication Year from 2018 to 2022, in Trials

E.2 Studienregister

1. ClinicalTrials.gov

Anbieter: U.S. National Institutes of Health

- URL: <http://www.clinicaltrials.gov>
- Eingabeoberfläche: Expert Search

Suchstrategie
AREA[ResultsFirstSubmitDate] NOT MISSING AND AREA[ConditionSearch] (abortion OR miscarriage)

2. International Clinical Trials Registry Platform Search Portal

Anbieter: World Health Organization

- URL: <https://trialsearch.who.int>
- Eingabeoberfläche: Standard Search

Suchstrategie
abortion OR pregnancy termination OR miscarriage OR early pregnancy failure [with Results only]