



## review

# Diagnosi domiciliare dell'apnea del sonno. Revisione sistematica della letteratura\*

Una revisione basata sull'evidenza sponsorizzata  
dall'American Academy of Sleep Medicine, American College  
of Chest Physician e American Thoracic Society

W. Ward Flemons, MD; Michael R. Littner, MD, FCCP;  
James A. Rowley, MD, FCCP; Peter Gay, MD, FCCP;  
W. McDowell Anderson, MD, FCCP; David W. Hudgel, MD, FCCP;  
R. Douglas McEvoy, MBBS, MD; Daniel I. Loube, MD, FCCP

(CHEST Edizione Italiana 2004; 3:42-78)

**Parole chiave:** diagnosi; rapporto di probabilità; polisonnografia/metodi; protocollo di ricerca; revisione della letteratura; sensibilità e specificità; sindrome dell'apnea del sonno

**Abbreviazioni:** AASM = American Academy of Sleep Medicine, ACCP = American College of Chest Physicians; AHRQ = Agency for Healthcare Research and Quality; ATS = American Thoracic Society; CPAP = ventilazione a pressione positiva continua; EMG = elettromiogramma; ERC = commissione per la revisione delle evidenze; LR = rapporto di probabilità; RDI = indice di disturbo respiratorio; REM = movimenti oculari rapidi; ROC = receiver operating characteristics; RTI-UNC = Research Triangle Institute-University of North Carolina

### 1.0 INTRODUZIONE/BACKGROUND

L'apnea del sonno è una patologia comune che colpisce bambini e adulti. La patologia è caratterizzata da periodi ricorrenti di cessazione del flusso aereo (apnea) e da periodi di riduzione del flusso aereo (ipopnea). I meccanismi fisiologici rela-

tivi ai due tipi di eventi sono simili così come simile è la loro rilevanza clinica.<sup>1</sup> La forma più comune di apnea del sonno, la cosiddetta apnea ostruttiva del sonno, è causata dal collassamento delle vie aeree superiori. Esistono diversi metodi per valutare la severità di questa condizione tra i quali la misurazione del numero di apnee e ipopnee per ora di sonno (indice di apnea-ipoapnea [AHI]), la severità delle desaturazioni durante il sonno o la severità della sonnolenza diurna, il sintomo più comunemente associato a questa condizione. Nel Winsconsin Sleep Cohort Study<sup>2</sup> la prevalenza di AHI  $\geq 5$  era del 24% negli uomini e nel 9% delle donne di età compresa tra il 30 ed i 60 anni. La prevalenza di apnea del sonno sintomatica (AHI  $\geq 5$  associato ad eccessiva sonnolenza diurna) negli uomini e nelle donne era pari al 4% e 2%, rispettivamente.<sup>2</sup> L'esame standard per diagnosticare la malattia è la polisonnografia, eseguita in un laboratorio del sonno in presenza di un tecnico, che permette di monitorare i parametri respiratori ed il tempo di sonno. La polisonnografia è un esame impegnativo che richiede tempo e la presenza di un tecnico specialista. Il lungo tempo d'attesa per accedere ai laboratori del sonno rappresenta un problema per molti pazienti, nella maggior parte dei quali l'apnea del sonno continua a rimanere non diagnosticata. Il Winsconsin

\*Dalla Faculty of Medicine (Dott. Flemons), University of Calgary, Calgary, AB, Canada; David Geffen School of Medicine (Dott. Littner), University of California Los Angeles, Los Angeles, CA, USA; Division of Pulmonary, Critical Care, and Sleep Medicine (Dott. Rowley), Department of Medicine, Wayne State University School of Medicine, Detroit, MI, USA; Pulmonary, Critical Care, and Sleep Medicine (Dott. Gay), Mayo Clinic, Rochester, MN, USA; the University of South Florida College of Medicine (Dott. Anderson), Tampa, FL, USA; Case Western Reserve University (Dott. Hudgel), Cleveland, OH, USA; Adelaide Institute for Sleep Health (Dott. Mc-Evoy), Repatriation General Hospital, Daw Park, Australia; the Swedish Medical Center (Dott. Loube), Seattle, WA, USA. Tavole di evidenza supplementari sono disponibili in linea: <http://www.chestjournal.org/cgi/content/full/124/4/1543/DC1>.

Manoscritto ricevuto il 1 maggio 2003; revisione accettata il 2 maggio 2003.

La riproduzione di questo articolo è vietata in assenza di autorizzazione scritta dell'American College of Chest Physicians (e-mail: [permissions@chestnet.org](mailto:permissions@chestnet.org)).

Corrispondenza: W. Ward Flemons, MD, Faculty of Medicine, University of Calgary, 1403 Twenty-Ninth St NW, Calgary, AB, Canada T2N 2T9; e-mail: [flemons@ucalgary.ca](mailto:flemons@ucalgary.ca)

(CHEST 2003; 124:1543-1579)

Sleep Cohort Study<sup>3</sup> dimostrava che nel 93% delle donne e nell'82% degli uomini con apnea del sonno di grado moderato-severo la diagnosi non era mai stata posta. Pertanto, si sta osservando un interesse crescente per gli approcci diagnostici alternativi, come per esempio i sistemi di monitoraggio portatili, che sono stati proposti come una valida alternativa per diagnosticare l'apnea del sonno. Il termine "sistema di monitoraggio portatile" include un'ampia serie di apparecchi che possono registrare molteplici segnali, così come avviene per la polisomnografia convenzionale, o un unico segnale come per esempio l'ossimetria (vedi paragrafo 1.1). Quando un sistema permette di registrare i segnali di EEG e l'elettromiogramma (EMG), può essere effettuata la stadiazione del sonno che consente di calcolare l'AHI. Solitamente, tuttavia, i segnali di EEG ed EMG non vengono registrati dai comuni sistemi di monitoraggio portatili ed in questo caso il numero di eventi respiratori viene calcolato per ora di monitoraggio e chiamato "Indice di disturbo respiratorio" (RDI). L'utilizzo dei sistemi di monitoraggio portatili per la diagnosi di apnea del sonno è un argomento controverso che è stato oggetto, in passato, di alcune review.<sup>4-8</sup> Da quando è stata prodotta l'ultima review sull'argomento sono stati pubblicati ulteriori studi scientifici e sono stati messi a punto metodi standardizzati per valutare le evidenze prodotte dagli studi sui test diagnostici.

L'American Thoracic Society (ATS), l'American College of Chest Physician (ACCP) e l'American Academy of Sleep Medicine (AASM) hanno progettato individualmente di revisionare e aggiornare le evidenze disponibili sulla validità dei sistemi portatili per la diagnosi dell'apnea del sonno negli adulti. Durante un congresso dell'ACCP, nel settembre del 2000, è stata discussa la possibilità che le tre organizzazioni potessero collaborare alla realizzazione di questo progetto e alla fine è stato raggiunto e formalizzato un accordo che prevedeva la creazione di un gruppo di studio, e il coinvolgimento di un centro specializzato nella revisione delle evidenze, per produrre una ricerca bibliografica dettagliata ed una revisione basata sulla evidenza sull'uso dei sistemi di monitoraggio portatili per la diagnosi dell'apnea del sonno. Altre due organizzazioni, la National Association for the Medical Direction of Respiratory Care e l'Australasian Sleep Association, hanno aderito a partecipare al progetto e hanno designato i loro membri nella commissione. Sono state preparate linee guida dettagliate sul conflitto di interessi per impedire la partecipazione a questo progetto di membri legati alle aziende produttrici di sistemi di monitoraggio portatili per la diagnosi dell'apnea del sonno. L'ACCP si è incaricato della parte amministrativa per il gruppo di studio. Sono state create tre commissioni che includevano, in ciascuna di esse, almeno un rappresentante per ciascuna organizzazione: (1) Steering Committee, Nancy Collop (Chair), Patrick Strollo, and John Shepard; (2) Evidence

Review Committee (ERC). Ward Flemons (Chair), James Rowley, Michael Littner, William Anderson, David Hudgel, Dan Loube, Peter Gay, and Doning McEvoy; and (3) Guideline Committee, Andrew Chesson (Chair), Alan Pack, and Richard Berry.

I fondi per il finanziamento di questo progetto, inclusa una revisione basata sull'evidenza commissionata ad una squadra di ricercatori specializzati presso la "RTI International and the University of North Carolina at Chappel Hill (RTI-UNC)" sono stati stanziati dall'ATS, ACCP, AASM.

In passato sono state pubblicate tre review sui sistemi di monitoraggio portatili per la diagnosi dell'apnea del sonno. Nel 1994 l'AASM (precedentemente nota come American Sleep Disorders Association) ha pubblicato la descrizione di 23 studi<sup>4</sup> che riportavano alcuni aspetti del monitoraggio con i sistemi portatili. In questa review gli studi sul sonno erano suddivisi nelle seguenti 4 categorie: tipo 1, polisomnografia standard; tipo 2, polisomnografia completa con sistema portatile; tipo 3, monitoraggio con sistema portatile modificato; tipo 4, registrazione in continuo di uno o due parametri (vedi sessione 1.1.1). Nel 1997 l'AASM ha pubblicato delle linee guida<sup>5</sup> ed una review<sup>6</sup> contenenti le indicazioni per la polisomnografia e procedure simili, nelle quali era inclusa una sezione dedicata agli studi di tipo 3 e 4. Basandosi sulla review, le linee guida affermavano che i monitoraggi di tipo 3, sorvegliati da un tecnico, erano potenzialmente appropriati in quei pazienti per i quali la diagnosi di apnea del sonno era altamente probabile (> 70%). In queste linee guida si raccomandava di eseguire una polisomnografia standard nei pazienti sintomatici per quali il monitoraggio di tipo 3 aveva dato esito negativo. Gli studi di tipo 4 non venivano invece presi in considerazione per la diagnosi di apnea del sonno. Nel 1997 l'Agency For Health Care and Quality (AHRQ) (precedentemente nota come Agency For Care Policy And Research), negli Stati Uniti, aveva commissionato una revisione sistematica della letteratura sulla diagnosi dell'apnea del sonno<sup>7,8</sup>. Una parte di questa revisione, che includeva gli articoli pubblicati dal 1980 al 1° Novembre 1987, prendeva in considerazione gli studi con i sistemi di monitoraggio portatili (25 studi), inclusa l'ossimetria (12 studi). In questa revisione sistematica la qualità di ciascuno studio veniva valutata attraverso una scala sviluppata dagli autori. Ciò era di potenziale aiuto alle revisioni dell'AASM perché si tentava di identificare quei problemi che potevano inficiare la validità dei dati e le conclusioni di uno studio.

Negli ultimi 10 anni si è osservato un interesse crescente per i metodi di valutazione della qualità degli studi scientifici, specialmente al momento di redigere una revisione sistematica. In tal senso è stato pubblicato un maggior numero di lavori sia sui metodi di valutazione degli studi sulla terapia che sui metodi di valutazione degli studi sui sistemi diagnostici. Il gruppo di studio dell'ACCP/ATS/AASM

ha considerato importante aggiornare la revisione della letteratura dal 1997 così come il sistema di valutazione della qualità degli studi sui sistemi di monitoraggio portatili. Il metodo pubblicato da Sacket et al<sup>9</sup> nel 2000, per valutare il grado di evidenza degli studi sui test diagnostici, è stato usato, in questo caso, perché si avvicina molto ai metodi comunemente usati per valutare la qualità degli studi sulla terapia e la prognosi. Inoltre, questo metodo si concentra sui seguenti aspetti che riguardano la progettazione degli studi sui test diagnostici: evitare errori di inclusione (usando un campione di pazienti consecutivi); praticare un disegno in cieco; evitare errori di verifica (usando riferimenti standard per tutti i pazienti).

Viene pubblicato un numero sempre crescente di lavori che paragonano i sistemi di monitoraggio portatili per la diagnosi dell'apnea del sonno con la polissonnografia convenzionale. Dal 1990 al 2001 sono stati pubblicati, nella letteratura in lingua inglese, 51 articoli che soddisfacevano i criteri per essere inclusi in questa revisione sistematica. Questi articoli venivano classificati in base al livello di evidenza (I, II, III o IV) basandosi sul metodo di Sacket et al. (Vedi paragrafo 1.5). La maggior parte di questi studi (30 su 51) era di qualità molto alta (livello I e II), senza che ci fosse un trend verso il miglioramento. In Fig. 1 sono mostrati gli studi di livello I (migliore qualità) e di livello II, così come il numero totale di studi pubblicati sui sistemi portatili nel corso degli anni.

L'obiettivo di una revisione sistematica è quello di riassumere l'insieme dei dati presenti in letteratura in modo da poter giungere più facilmente a delle conclusioni su un aspetto particolare della medicina. Un approccio comunemente usato per sintetizzare dati di studi diversi è la meta-analisi. Questo approccio è stato adoperato per la stesura della review commissionata dall'AHRQ nella quali i risultati venivano riportati come curve "receiver operating characteristic" (ROC).<sup>8</sup> Questo gruppo di studio ha deciso di non effettuare una meta-analisi dei risultati perché gli studi analizzati erano troppo diversi tra loro per ciò che riguarda alcuni aspetti quali i segnali misurati (Tabella 1<sup>10-60</sup> e paragrafo 1.1.1), i criteri usati per definire un evento respira-

torio (paragrafo 1.1.2), l'analisi dei segnali (paragrafo 1.1.3) e per la diversa qualità degli studi stessi (paragrafo 1.5). Pertanto il gruppo di studio ha deciso di riassumere e riportare singolarmente i dati di ciascuno studio, senza combinarli in forma di meta-analisi, in modo da permettere al lettore di trarre le conclusioni sulle evidenze. I dati di ciascun studio sono stati sintetizzati in tabelle e classificati come segue: (1) tipo di monitoraggio (paragrafo 1.1); (2) setting dello studio (domicilio vs laboratorio) (paragrafo 1.4.1); (3) livello di evidenza e livello di qualità (paragrafo 1.5).

Il report è stato suddiviso in 3 sezioni primarie e 4 secondarie. Le sezioni primarie sono le seguenti:

1. Utilità dei sistemi di monitoraggio portatili nel ridurre la probabilità che il paziente abbia un AHI anormale (escludere la patologia) (paragrafo 4.1.1);
2. Utilità dei sistemi di monitoraggio portatili nell'aumentare la probabilità che il paziente abbia un AHI anormale (individuare la patologia) (paragrafo 4.1.2);
3. Utilità dei sistemi di monitoraggio portatili nel ridurre ed aumentare la probabilità che un paziente abbia un AHI anormale (escludere ed individuare la patologia) (paragrafo 4.1.3).

Le sezioni secondarie sono le seguenti:

1. Riproducibilità dei risultati ottenuti con i sistemi di monitoraggio portatili (paragrafo 4.2.1);
2. Rapporto costo-beneficio dei sistemi di monitoraggio portatili (paragrafo 4.2.2);
3. Percentuale di fallimenti dei sistemi di monitoraggio portatili (paragrafo 4.2.3);
4. Popolazioni di pazienti studiati e possibilità di generalizzare i risultati (paragrafo 4.2.4).

Infine l'obiettivo di questo gruppo di studio era quello di evidenziare i problemi che emergono dalla letteratura corrente sui sistemi di monitoraggio portatili per l'apnea del sonno, di descrivere le possibilità per una futura ricerca e di sottolineare alcuni degli aspetti metodologici fondamentali che dovrebbero essere presi in considerazione dai ricercatori, dagli editori delle riviste scientifiche, dai revisori e dai lettori.

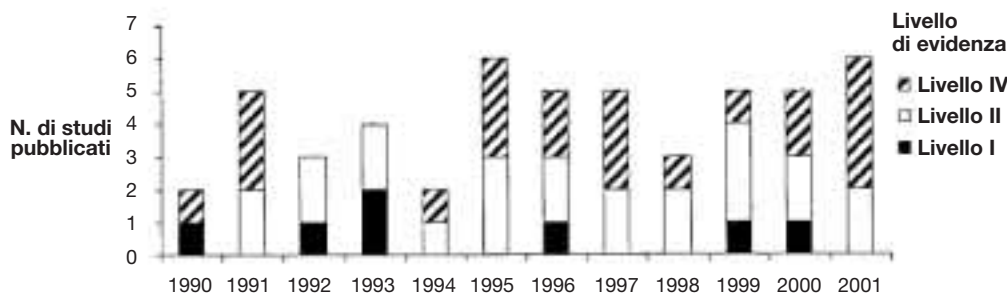


FIGURA 1. Qualità degli studi pubblicati sui sistemi di monitoraggio portatili per la diagnosi di apnea del sonno dal 1990 al 2001.

**Tabella 1—Definizione degli eventi respiratori e informazioni sugli studi\***

Scoring del canale e criteri		Criteri secondari	Oxim Samp Freq Hz	Definizione eventi PSG	Livello di evidenza	Livello di qualità	Studio/anno	Maschi %	AHI medio	BMI medio	PM tipo	PM nome†	Metodo di analisi	Ref. n.
Criteri primari	Scoring del canale e criteri													
Flusso (NP)	↓ > 50%			Completa ↓ flusso (term) o ↓ tor/add RIP > 50% ↓ Flusso (term) > 50%	II	a	Rees et al/1998 (n = 20)	100	39	31	4	AutoSet	Auto	10
				↓ Flusso (PT) + ↓ O <sub>2</sub> > 3% o risveglio ↓ Flusso (term) > 50%	II	a	Gugger et al/1997 (n = 67)	87	26	31	4	AutoSet	Auto	11
				NR	IV	b	Mayer et al/1998 (n = 95)	83	43	31	4	AutoSet	Auto	12
				↓ Flusso (NS) > 80% o ↓ flusso > 50% + ↓ O <sub>2</sub> > 3% o risveglio	IV	b	Kiely et al/1996 (n = 36)	75	19,4	28	4	AutoSet	Auto	13
				NR	IV	b	Bradley et al/1995 (n = 31)	84	25	30	4	AutoSet	Auto	14
				NR	IV	a	Bagnato et al/2000 (n = 56)	80	38	31,3	4	AutoSet	Auto	15
No Flusso				NR	IV	c	Gugger et al/1995 (n = 27)	NR	NR	29	4	AutoSet	Auto	16
				Completa ↓ flusso (PT) o ↓ flusso— non specificato Valutabile ↓ flusso (term) + ↓ O <sub>2</sub> > 3%	IV	d	Fleury et al/1996 (n = 44)	77	NR	29	4	AutoSet	Auto	17
Flusso (term)	Valutabile ↓	e ↓ O <sub>2</sub> > 1%	NR	Completa ↓ flusso (term) o valutabile ↓ flusso + ↓ O <sub>2</sub> > 3% o risveglio	II	a	Redline et al/1991 (n = 25)	83	NR	31	3	Edentec	Mannale	18
				↓ Flusso (term) > 25%	IV	a	Fry et al/1998 (n = 93)	64	29	NR	2	DigitTrace	Mannale	19
				↓ Flusso (term) > 40%	IV	b	Portier et al/2000 (n = 78)	82	26	31	2	MiniSomno	Mannale	20
				↓ Flusso (term) > 50%	I	a	Zucconi et al/1996 (n = 29)	69	32	31	3	Microdigitrap	Comb	21
				↓ Flusso (term) > 50%	I	a	Ensellem et al/1990 (n = 63)	NR	NR	NR	3	EdenTrace	Mannale	22
				↓ Flusso (term) > 50%	II	a	Man and Kang/1995 (n = 104)	80	NR	NR	3	Poly G	Mannale	23
				↓ Flusso (term) > 50%	IV	a	Ancoli-Israel et al/ 1997 (n = 34)	89	42	NR	3	Night Watch	Auto + manuale	24
				Completa ↓ flusso (term) o ↓ flusso > 50% or ↓ tor/ add RIP > 50%	II	a	Myklytyn et al/1999 (n = 20)	100	NR	32	2	Compumedics	Mannale	25
				Completa ↓ flusso (term) o valutabile ↓ flusso o paradosso tor/add + ↓ O <sub>2</sub> (NS)	II	a	Parra et al/1997 (n = 89)	82	34	29	3	EdenTrace	Mannale	26
				o risveglio Completa ↓ flusso (term) o valutabile ↓ flusso + ↓ O <sub>2</sub> > 2% o risveglio	II	b	Ballester et al/2000 (n = 116)	56	10	26	3	Sibel-home monitor	Mannale	27
				Completa ↓ flusso (term) o ↓ flusso > 50% + ↓ O <sub>2</sub> > 3%/1% o risveglio	II	a	Ficker et al/2001 (n = 51)	86	24	29	3	Somnocheck	Mannale	28
Flusso (term)	No Flusso	o Flusso + ↓ > 3% e O <sub>2</sub> > 1%	4	Completa ↓ flusso (term) o ↓ flow > 50% + ↓ O <sub>2</sub> > 3%/1% o risveglio	II	a	White et al/1995 (n = 100)	28/31	31/33	31/33	3	Night Watch	Auto + manuale	29
				Completa ↓ flusso (term) o ↓ tor/add RIP > 50%	IV	b	Whittle et al/1997 (n = 78)	93	NR	31	3	EdenTrace	Mannale	30
				↓ Flusso (NS) > 50%	I	a	Verse et al/1992 (n = 53)	92	18	27	3	Poly-MESAM	Auto + manuale	31
				NR	IV	d	Orr et al/1994 (n = 40)	NR	26	NR	2	Sleep I/T	Auto +	32

Tabella 1—Continuazione \*

Scoring del canale e criteri		Oxim Samp Freq Hz	Definizione eventi PSG	Livello di evidenza	Livello di qualità	Studio/anno	Maschi % AHI medio	BMI medio	PM tipo	PM nome†	Metodo di analisi	Ref. n.
Criteri primari	Criteri secondari											
Flusso (term) + ↓ O <sub>2</sub> Saturazione di O <sub>2</sub>		10 NR	↓ Flusso (term) > 90% o ↓ flusso > 50% + ↓ O <sub>2</sub> > 3% NR	II II	a a	Baltzan et al/2000 (n = 97) Stoohs e Guilleminault/ 1992 (n = 56) Curubhagavatula et al/2001 (n = 359) Vazquez et al/2000 (n = 241)	75 82	28,4 NR	4 4	Oxiflow Mesam IV	Comb Auto	33 34
		NR	Completa ↓ flusso (NS) or ↓ flusso > 50% + ↓ O <sub>2</sub> > 3% o risveglio	II	d		68	32	4	NR	Manuale	35
e	Risaturazione > 2%	1	Completa ↓ flusso (term) o ↓ flusso > 50% + ↓ O <sub>2</sub> > 3% o risveglio	I	a		78	32	4	SnoreSat	Auto	36
		2	NR	II	b	Yamashiro et al/ 1995 (n = 269)	70	NR	4	Biox 3740	Manuale	37
	Saturazione > 1% e HR di O <sub>2</sub> Δ	0,5 e 0,08	↓ O <sub>2</sub> > 3%	IV	a	Wiltshire et al/2001 (n = 100)	36	NR	4	Biox 3740	Auto	38
> 3% ↓		0,5	Completa ↓ flusso (NS) o ↓ tor/add > 30% + ↓ O <sub>2</sub> > 3% o risveglio	IV	b	Schafer et al/1997 (n = 114)	88	29	31	Mesam IV	Manuale	39
		0,08	↓ Flusso (NS) > 50%	I	d	Gyulay et al/1993 (n = 98)	79	NR	4	Biox 3700	Auto o manuale	40
o	CT < 90% > 1% CT < 90% > 0,8%	0,2	Completa ↓ flusso (term) o valutabile ↓ flusso + ↓ O <sub>2</sub> > 3% o risveglio	II	c	Golpe et al/1999 (n = 116)	90	24	4	Minolta Pulseox 7	Auto + manuale	41
	Saturazione < 90% di O <sub>2</sub>	0,16	Completa ↓ flusso (NS)	IV	c	Williams et al/1991 (n = 36)	NR	NR	4	Biox 3700 IV/ Nellcor N100	Manuale	42
	Pausa di russamento o variabilità dell'HR	NR	Completa ↓ flusso (term) o valutabile ↓ flusso + ↓ O <sub>2</sub> > 3% o risveglio	II	a	Esmola et al/1996 (n = 150)	89	30	4	Mesam IV	Auto o manuale	43
	> 10% HR Δ	NR	Completa ↓ flusso (NS) o ↓ tor/add > 50%	II	a	Koziej et al/1994 (n = 56)	91	NR	4	Mesam IV	Auto o manuale	44
> 4% ↓		0,16	Completa ↓ flusso (NS) o ↓ flusso o tor/add > 50% + ↓ O <sub>2</sub> > 3% o risveglio	I	b	Chiner et al/1999 (n = 275)	90	42	4	Nellcor N200	Comb	45
		NR	NR	IV	d	Bonsignore et al/ 1990 (n = 73)	74	NR	4	NR	NR	46
	Risaturazione > 2%	1	Completa ↓ flusso (term) o ↓ tor/add > 50% + ↓ O <sub>2</sub> > 2%	IV	c	Rauscher et al/1991 (n = 53)	NR	NR	4	Mesam with Minolta Pulseox 7	Auto	47
> 5% ↓		NR	NR	II	c	Cooper et al/1991 (n = 41)	61	NR	4	Biox Ila	Manuale	48
Indice Delta > 0,6		0,08	Completa ↓ flusso (NS) o ↓ flusso > 50% + ↓ O <sub>2</sub> (NS) o risveglio	IV	c	Levy et al/1996 (n = 301)	NR	NR	4	Biox 3700	Auto	49
> 0,8		0,08	NR	IV	d	Pepin et al/1991 (n = 26)	NR	51	4	Biox 3700	Auto	50
> 0,4 o	CT < 90% > 1%	0,08	NR	IV	d	Olson et al/1999 (n = 906)	NR	40	4	Biox 4700	Auto	51

**Tabella 1—Continuazione\***

Scoring del canale e criteri		Oxim Samp Freq Hz	Definizione eventi PSG	Livello di evidenza	Livello di qualità	Studio/anno	Maschi %	AHI medio	BMI medio	PM tipo	PM nome†	Metodo di analisi	Ref. n.
Criteri primari	Criteri secondari												
Saturazione di O <sub>2</sub>	PA soglia 0,7% o Nessun criterio	0,2 2 NR NR 1 0,2	Completa ↓ flusso (NS) o ↓ flusso > 50% + ↓ O <sub>2</sub> > 3% Completa ↓ flusso (NS) o ↓ tor/add > 50% + ↓ O <sub>2</sub> > 4% NR Completa ↓ flusso (NS) o ↓ tor/add > 50% + ↓ O <sub>2</sub> > 2% Completa ↓ flusso (term) + ↓ O <sub>2</sub> > 4% ↓ flusso (term) > 25% + ↓ tor > 25% + ↓ add > 15% + tor/add ↓ tor/add > 50% + ↓ O <sub>2</sub> > 3% Completa ↓ flusso (NS) o ↓ flusso (NS) + ↓ O <sub>2</sub> > 3% NR Completa ↓ flusso (term) o ↓ flusso > 50% + ↓ O <sub>2</sub> > 3%	II I II II IV	a b a b c	Zamarron et al/1999 (n = 253) Series et al/1993 (n = 240) Douglas et al/1992 (n = 200) Ranschler et al/1993 (n = 116) Ryan et al/1995 (n = 69)	80 90 80 82 83	40 38 NR NR NR	30 32 NR NR 30	4 4 4 4 4	Criticare 504 BioX IVA Ohmeda 3700 Minolta Pulsox Auto 7 Minolta Pulsox 7	Auto Manuale Auto Auto Manuale	52 53 54 55 56
Russamento	Presente	1	↓ tor/add > 50% + ↓ O <sub>2</sub> > 3%	II	a	Issa et al/1993 (n = 29)	78	NR	31	4	SnoreSat	Auto	57
Analisi spettrale della frequenza cardiaca	Picco di potenza in FFT in finestra di 30-70 sec.	0,2	Completa ↓ flusso (NS) o ↓ flusso (NS) + ↓ O <sub>2</sub> > 3%	IV	b	Zamarron et al/2001 (n = 197)	82	40	31	4	Criticare 504	Auto	58
Pressione faringoesofagea	NR	NR	NR	IV	d	Reda et al/2001 (n = 59)	81	40	32	4	NR	NR	59
NR	NR	NR	Completa ↓ flusso (term) o ↓ flusso > 50% + ↓ O <sub>2</sub> > 3%	IV	d	Claman et al/2001 (n = 42)	74	26	31	3	Bedbugg	Auto	60

\*Oxim Samp Freq = frequenza di campionamento dell'ossimetria; PSG = polisomnografia; Pub date = data di pubblicazione; BMI = indice di massa corporea; PM = sistema di monitoraggio portatile; NR = non riportato; NS = non specificato; Flusso (NP) = flusso aereo misurato con pressione nasale; Flusso (term) = flusso aereo misurato con pneumotacografo; FFT = trasformata rapida di Fourier; tor = torace; add = addome; paradosso = movimento paradosso del torace e dell'addome; CT = tempo cumulativo di saturazione di O<sub>2</sub>; Auto = gli eventi erano identificati usando un sistema di analisi automatico basato su algoritmi; Manuale = gli eventi erano identificati manualmente dall'operatore; Auto+manuale = gli eventi erano identificati attraverso un'analisi automatica con algoritmi e manuale dell'operatore; Comb = gli eventi erano identificati attraverso un'analisi automatica con algoritmi con o senza analisi manuale dell'operatore, con entrambi i risultati riportati; PA = arteria polmonare; ↓ = riduzione; ETCO<sub>2</sub> = CO<sub>2</sub> di fine espirazione; RIP = pletismografia respiratoria induttiva.  
†Informazioni sulle ditte produttrici: AutoSet (Resmed; Poway, CA); Edentec (Nellcor; Pleasanton, CA); DigiTrace (SleepMed; Peabody, MA); MimiSomno (Nellcor); Microdigitrap; EdenTrace (Nellcor Puritan Bennett [Melville] Ltd; Kanata, ON, Canada); Poly G (CNS Inc; Whippany, NJ); Night Watch (Respironics; Murrayville, PA); Compumedics (Abbotsford, Australia); Sibel (Gemblox, Belgium); Somnocheck (Weinmann; Hamburg, Germany); Poly-MESAM (MAP GmbH; Martinsreid, Germany); Sleep I/T (CNS Inc); Oxiflow (Nellcor); Mesam IV (MAP GmbH); SnoreSat (Saga Tech; Calgary, AB, Canada); Biox monitors (Ohmeda; Madison, WI); Pulsox 7 (Minolta; Diessehofen, Switzerland); Criticare 504 (Criticare Systems, Inc; Waukesha, WI); Ohmeda 3700 (Ohmeda); Bedbugg (Sleep Solutions Inc; Palo Alto, CA).

## 1.1. Sistemi di monitoraggio portatili

### 1.1.1. Tipo di monitoraggio

I sistemi di monitoraggio portatili sono stati classificati secondo il metodo usato dall'America Sleep Disorder Association nella review pubblicata nel 1994.<sup>4</sup> Il monitoraggio di tipo 1 (polisonnografia standard) era considerato il sistema standard con il quale tutti gli altri sistemi venivano paragonati. I segnali fisiologici che venivano registrati dal sistema portatile ed usati per definire un evento respiratorio variavano tra i diversi studi e con il variare del sistema di monitoraggio (Tabella 1<sup>10-60</sup>). Come spiegato sotto in dettaglio i monitoraggi di tipo 2 includono la stadiazione del sonno e la misurazione dei parametri respiratori, i monitoraggi di tipo 3 utilizzano almeno 3 canali respiratori, mentre i monitoraggi di tipo 4 utilizzano almeno un canale respiratorio, generalmente o la saturazione di ossigeno o il flusso aereo.

#### 1.1.1.1. Tipo 2: polisonnografia con sistema portatile completo

Questi sistemi includono un minimo di sette canali quali EEG, elettrooculogramma, EMG del mento, ECG o frequenza cardiaca, flusso aereo, sforzo respiratorio e saturazione di ossigeno. Questo tipo di monitoraggio permette di stadiare il sonno e pertanto di calcolare un AHI.

#### 1.1.1.2. Tipo 3: sistema portatile modificato

Questo sistema include un minimo di 4 canali quali ventilazione o flusso aereo (almeno due canali per i movimenti respiratori o movimenti respiratori e flusso aereo) frequenza cardiaca o ECG, e saturazione di ossigeno.

#### 1.1.1.3. Tipo 4: sistemi ad uno o due canali

La maggior parte di questi sistemi misura uno o due parametri per esempio la saturazione di ossigeno o il flusso aereo. I sistemi che non avevano i requisiti per essere inclusi nel tipo 3 (es. un sistema che misurava da 1 a 3 canali o che, pur avendo 4 canali, non includeva il flusso aereo) venivano classificati di tipo 4.

### 1.1.2. Segnali per identificare gli eventi

Un importante compito del gruppo di studio era quello di osservare le similitudini e le differenze nel modo in cui i diversi sistemi registravano i segnali e nel modo in cui questi segnali venivano utilizzati per definire un evento respiratorio. Come per la polisonnografia, anche per il monitoraggio con il sistema portatile il concetto di "evento respiratorio anormale" era eterogeneo (Tabella 1). I metodi più comuni per identificare gli eventi respiratori erano la riduzione del flusso aereo misurato attraverso un termistore, o con un segnale di pressione nasale, e la desaturazione di ossigeno (molteplici approcci diversi). In alcune circostanze queste metodiche venivano usate in combinazione.

### 1.1.2.1. Flusso

Il metodo standard per definire una apnea o ipoapnea è la riduzione del flusso aereo o del volume corrente. Per definire una ipoapnea la riduzione del flusso deve essere  $> 50\%$  rispetto al basale.<sup>1</sup> Il metodo migliore per quantizzare il flusso è il pneumotacografo, tuttavia nessun sistema portatile si avvale di questa tecnica.

#### 1.1.2.1.1. Termistore

I termistori misurano la differenza di temperatura e non esiste una relazione lineare con il flusso aereo. Pertanto possono non essere abbastanza sensibili da individuare le ipoapnee. Per questo motivo è stato raccomandato di non usare i termistori quando si effettua una polisonnografia a scopo di ricerca. Uno dei vantaggi dei termistori è che essi possono rilevare il flusso aereo attraverso il naso e la bocca, e pertanto questo rimane il metodo più comunemente usato per definire gli eventi respiratori basati sulla misurazione del flusso (Tabella 1).

#### 1.1.2.1.2. Pressione nasale

La misurazione della pressione nasale fornisce una approssimazione lineare del flusso aereo nel suo range completo, tranne alle estremità. La relazione lineare può essere migliorata con un'estrazione della radice quadrata. Tuttavia ciò può non essere necessario se l'obiettivo primario della misurazione è l'individuazione dell'evento. Questo metodo può non essere accurato come un termistore nel distinguere un'apnea da un'ipoapnea. Tuttavia, nella pratica clinica, questa distinzione non è considerata importante.<sup>1</sup> Il segnale può produrre falsi positivi se il paziente respira ad intermittenza dalla bocca e la qualità del segnale potrebbe essere scarsa se il paziente respira dalla bocca per lunghi periodi di tempo. Pertanto, in questo caso può essere necessario avere una conferma visiva delle apnee e delle ipoapnee e ciò rende potenzialmente difficile usare questo sistema per gli studi domiciliari.<sup>11-14,16,17</sup>

#### 1.1.2.2. Pletismografia respiratoria induttiva

La pletismografia respiratoria induttiva, quando opportunamente calibrata, può fornire una misura del volume corrente. Quando il segnale non è calibrato, può solo essere utile ad identificare gli eventi respiratori. Questo metodo viene usato principalmente durante la polisonnografia. Solo in uno studio la pletismografia respiratoria induttiva è stata usata in un sistema portatile come segnale secondario (Tabella 1).

#### 1.1.2.3. Saturazione di ossigeno

Gli ossimetri sono molti diversi dagli altri sistemi di misurazione, in particolare per ciò che riguarda la frequenza di campionamento e gli algoritmi usati per misurare la saturazione di ossigeno. Alcuni ossimetri effettuano letture multiple, le conservano in memoria, calcolano la media e riportano un valore

ogni 3-12 secondi. Altri ossimetri campionano e riportano ciascun valore ad una frequenza fino a 10 Hz.<sup>33</sup> In uno studio è stato dimostrato che una frequenza di campionamento di 1/12 Hz riduceva il numero degli artefatti.<sup>38</sup> I metodi per l'analisi automatica della saturazione di ossigeno sono anch'essi variabili. La maggior parte di questi metodi si basa sull'identificazione delle cadute della saturazione di ossigeno, alcuni identificano la risaturazione, mentre altri utilizzano entrambi i criteri (Tabella 1). Alcuni sistemi di analisi automatica definiscono il livello basale di saturazione di ossigeno, ma la maggior parte dei sistemi non ha questa funzione. In alcuni studi è stata misurata la percentuale di tempo cumulativo che il paziente passa con una saturazione di ossigeno < 90% per vedere se ciò poteva essere utile ad identificare i pazienti con apnea del sonno. In altri studi è stato calcolato "l'indice delta" che permette di valutare la variabilità della saturazione di ossigeno durante un intero studio. Questi due ultimi metodi non permettono di identificare gli eventi specifici ma piuttosto di individuare quei pazienti nei quali c'è la probabilità di trovare apnee o ipoapnee.

L'analisi dell'ossimetria, diretta a identificare le riduzioni transitorie della saturazione di ossigeno, dovrebbe essere più sensibile nelle situazioni in cui il livello di saturazione basale di ossigeno è più basso, cioè in relazione alla curva di dissociazione ossiemoglobinica. Pertanto, nei pazienti studiati in altitudine o nei pazienti con patologie polmonari (es. BPCO), può essere evidenziato un numero più elevato di desaturazioni e ciò potrebbe migliorare la sensibilità del monitoraggio e ridurne la specificità. A questo proposito sono stati pubblicati due studi che valutavano i pazienti con BPCO.<sup>45,50</sup> Tuttavia questi due studi non chiarivano come la presenza di BPCO potesse influenzare la sensibilità e la specificità del sistema di monitoraggio portatile.

#### 1.1.2.4. Altri

In uno studio<sup>57</sup> veniva usato il russamento come metodo primario per identificare gli eventi e combinato con la desaturazione di ossigeno come evento secondario. Altri studi<sup>43-44</sup> hanno usato il russamento in associazione alla variabilità della frequenza cardiaca come criterio per l'identificazione degli eventi. In uno studio<sup>58</sup> è stata usata l'analisi spettrale della frequenza cardiaca, mentre un altro studio prendeva in considerazione la misurazione della pressione faringo-esofagea<sup>59</sup> come parametro per identificare gli eventi respiratori (Tabella 1).

#### 1.1.3. Metodi di analisi

Gli studi differivano tra di loro per i canali fisiologici monitorati, i criteri usati per definire gli eventi e i metodi usati per lo scoring degli eventi (Tabella 1). Nella maggior parte degli studi in cui veniva utilizzato il termistore per misurare il flusso l'analisi veniva eseguita manualmente, mentre nella maggior

parte degli studi in cui si utilizzava la pressione nasale l'analisi era automatica. Alcuni sistemi effettuano l'analisi automatica attraverso il computer o attraverso una stampa che permette anche il controllo manuale da parte dell'operatore, per eventuali modifiche. Negli studi alcuni autori spiegavano in che modo era stata eseguita l'analisi (automatica, automatica e manuale, solo manuale) e, nel caso in cui l'analisi era solo manuale, da chi veniva effettuata. Tuttavia la maggior parte degli studi non dava informazioni chiare su questi aspetti. L'analisi automatica ha il vantaggio di eliminare una fonte di variabilità nei risultati, cioè il riconoscimento degli eventi da parte di un operatore. Tuttavia, i polisinnogrammi usati come riferimento standard per individuare i pazienti con o senza apnea del sonno vengono analizzati manualmente. Pertanto si potrebbe obiettare che l'analisi automatica dei sistemi di monitoraggio portatili non è confrontabile. Inoltre, lo scoring polisinnografico può includere il risveglio come criterio secondario (Tabella 1). L'analisi con i sistemi di monitoraggio portatili, quando eseguita manualmente, richiede un periodo di tempo compreso tra 2 e  $\geq$  di 10 minuti. Alcuni studi hanno preso in considerazione l'analisi automatica con la correzione manuale e hanno riportato risultati diversi (Tabella 1). La variabilità dei risultati ottenuti con analisi manuale non è stata riportata (paragrafo 4.2.1). Alcuni operatori considerano i sistemi di analisi automatici "scatole nere", perché non identificano su una traccia di appropriata risoluzione gli eventi classificati, cosicché un tecnico o un clinico li possa rivedere e modificare, individuare gli artefatti e valutare la qualità dello studio. I dettagli sulla possibilità, con ciascun sistema di monitoraggio, di vedere gli eventi su un monitor, in modo da permettere al tecnico o al clinico di revisionarli, non erano sempre riportati negli studi che utilizzavano l'analisi automatica.

#### 1.2. Setting dello studio e sorveglianza

I sistemi di monitoraggio portatili possono essere usati in molti setting come l'ospedale, il laboratorio del sonno o il domicilio del paziente. Inoltre il monitoraggio può essere effettuato in presenza del tecnico o senza sorveglianza. Il ruolo del tecnico è quello di stabilire se il sistema di monitoraggio portatile sta funzionando bene, di aiutare il paziente, per esempio incoraggiandolo a dormire supino, e di intervenire in caso di problemi. Con poche eccezioni, negli studi effettuati presso i laboratori del sonno, il monitoraggio veniva effettuato in presenza di un tecnico e simultaneamente con il polisinnogramma, mentre il monitoraggio domiciliare veniva effettuato senza sorveglianza. In un piccolo studio<sup>25</sup> effettuato in un laboratorio del sonno, veniva valutata la performance di un sistema portatile di tipo 2 in presenza e in assenza del tecnico. In un secondo studio,<sup>48</sup> effettuato presso un ospedale, la sorveglianza era parziale e non si aveva la registrazione simultanea del polisinnogramma.

nogramma. In uno studio<sup>29</sup> eseguito con un sistema di monitoraggio portatile a casa del paziente il tecnico aveva la possibilità di osservare lo studio a distanza e di intervenire chiamando il paziente qualora insorgessero problemi tecnici.

### 1.3. *Concordanza tra le misurazioni*

Esistono diversi metodi per valutare il grado di concordanza tra due metodi designati per la misurazione di uno stesso fenomeno. Tra questi ricordiamo il coefficiente di correlazione di Pearson, il coefficiente di correlazione intraclasse, l'approccio di Bland e Altman delle differenze medie e dei limiti di concordanza, ed infine i rapporti sensibilità/specificità/probabilità (LRs). Sebbene il coefficiente di correlazione di Pearson sia ampiamente utilizzato, in questo caso questo approccio non è raccomandato perché misura l'associazione, e non la concordanza.<sup>61</sup> I coefficienti di correlazione intraclasse potrebbero essere usati per valutare la concordanza,<sup>62</sup> tuttavia la maggior parte dei clinici non ha familiarità con questo approccio e pertanto esso non viene usato comunemente. L'approccio di Bland e Altman, che consiste nel calcolare la media delle differenze tra due misurazioni, è preferibile al coefficiente di correlazione di Pearson, tuttavia, i limiti di concordanza, che rappresentano il fattore chiave che descrive come due misure concordano tra loro, possono essere ingannevoli, se non ben calcolati.

Sensibilità, specificità ed LRs hanno il vantaggio di essere usate comunemente e sono facilmente comprensibili. Questi parametri permettono di risolvere il problema di una specifica classificazione dei pazienti. Utilizzare il rapporto sensibilità/specificità/LRs richiede che il paziente venga classificato affetto o non affetto dalla patologia basandosi su un valore cut-off di AHI, che varia nei diversi studi. La severità degli eventi respiratori durante la notte è estremamente variabile e l'AHI ne descrive solo un singolo aspetto. Poiché molti pazienti hanno indici che oscillano attorno al valore di cut-off abituale è possibile che la classificazione del paziente possa variare in seguito alla variabilità nelle misurazioni (paragrafo 4.2.1). Inoltre sono stati sollevati dubbi legittimi sul fatto che l'AHI, che è una misura derivata dalla polisonnografia eseguita nei laboratori del sonno, possa essere considerato un riferimento standard corretto. L'AHI rappresenta il riferimento standard più comunemente usato ed il sistema per valutare la severità dell'apnea del sonno per il quale esistono più dati pubblicati in relazione alla morbilità (disfunzione neurocognitiva, ipertensione e qualità di vita). Per questo motivo questo parametro costituisce la base per la revisione sistematica condotta dal gruppo di studio dell'ATS/ASM/ACCP.

L'analisi dei risultati, effettuata utilizzando la sensibilità, la specificità e l'LRs, dovrebbe tener conto della precisione delle stime (calcolo degli intervalli di confidenza) che riflettono direttamente la misura

del campione ed il protocollo di studio. Gli studi con livello IV di evidenza e quelli che includono un piccolo numero di pazienti (e ampi intervalli di confidenza) dovrebbero essere interpretati con cautela (vedi Tabelle 3 e 4). Sensibilità, specificità ed LRs, descrivono le caratteristiche operative di un test (misura in cui la probabilità di essere in presenza della malattia è modificata da un risultato positivo o negativo). Tuttavia poiché un clinico ha bisogno di conoscere la reale probabilità che il paziente sia affetto o meno da una determinata patologia (probabilità post-test o valore predittivo) le caratteristiche operative di un test devono essere interpretate alla luce della conoscenza della probabilità pre-test (prevalenza) della patologia (Tabella 2). L'utilità dei test diagnostici alternativi alla polisonnografia, per i pazienti con sospetta apnea del sonno, può essere vista come la percentuale di pazienti che hanno un test positivo o negativo e la percentuale di pazienti per i quali si è ottenuto un risultato falso positivo o falso negativo, rispettivamente (vedi Tabelle da 3 a 5). Poiché il numero di risultati veramente positivi dipende dalla sensibilità e il numero di falsi positivi dipende dalla specificità, entrambi questi parametri sono importanti per valutare l'utilità di un test, oltre alla probabilità pre-test.

### 1.4. *Validazione dei sistemi di monitoraggio portatili*

Per validare i sistemi di monitoraggio portatili sono stati usati diversi approcci. L'approccio standard, effettivamente usato fino ad ora, è stato quello di paragonare il sistema portatile con un sistema standard, come descritto in precedenza. Il limite di questo approccio sta nel fatto che esso considera la polisonnografia standard, eseguita in un laboratorio del sonno, come l'esame ottimale per la diagnosi dell'apnea del sonno. Tuttavia ciò non è completamente vero, per diverse ragioni. Da un punto di vista tecnico il paziente non dorme bene in un laboratorio come a casa sua e trascorre in media più tempo dormendo in posizione supina. Inoltre, l'AHI non correla bene con outcome quali la qualità di vita e la sonnolenza diurna e non è predittivo di quali pazienti alla fine utilizzeranno e trarranno beneficio dalla terapia. Pertanto uno studio di validazione più appropriato dovrebbe paragonare la rilevanza del monitoraggio con sistemi portatili e della polisonnografia standard per le decisioni cliniche del medico e per gli outcome importanti per i pazienti. Fino ad oggi non sono stati pubblicati studi che si sono avvalsi di questo approccio per validare l'uso dei sistemi di monitoraggio portatili.

Ci sono diversi aspetti nel protocollo e nei metodi di uno studio che, se non attentamente controllati, possono inficiarne i risultati e le conclusioni. In questa revisione, per ciascuno studio, abbiamo valutato un livello di evidenza e assegnato un livello di qualità basandoci su come venivano controllati alcuni aspetti, potenzialmente problematici, del protocollo

di studio (paragrafo 2.3). Altri aspetti del protocollo di studio che possono influenzare l'interpretazione dei risultati sono presi in considerazione qui di seguito.

#### 1.4.1. Monitoraggio sorvegliato/non sorvegliato

Quando il monitoraggio con un sistema portatile viene effettuato in un setting con sorveglianza (più spesso un laboratorio del sonno) si può valutare la sua performance in circostanze ideali, eliminando importanti fonti di variabilità che non hanno nulla a che fare con il sistema di monitoraggio, quali ad esempio la variabilità da notte a notte. La possibilità di eseguire in simultanea una polisonnografia permette di valutare un aspetto importante, cioè l'efficienza del sistema portatile. Se però il sistema non viene poi testato in un setting non sorvegliato, possibilmente la casa del paziente, rimarrà il problema di stabilire se il sistema veramente funziona nel setting per il quale esso è stato ideato. Inoltre, quando i dati di uno studio domiciliare vengono paragonati a quelli ottenuti con uno studio polisunnografico condotto in laboratorio, le limitazioni della polisunnografia, come esame standard, devono essere tenute in considerazione.

#### 1.4.2. Metodologia dello studio

##### 1.4.2.1. Descrizione della popolazione dello studio

Un'adeguata descrizione della popolazione dei pazienti studiati è fondamentale per aiutare il lettore a capire se i risultati dello studio sono applicabili anche alla popolazione dei loro pazienti. Idealmente, un'ampia casistica di pazienti (severità della malattia, età, razza, sesso) deve essere ottenuta senza che l'investigatore partecipi alla selezione dei pazienti. Quest'ultimo punto serve ad evitare difetti di selezione. Se gli investigatori studiano un gruppo di pazienti selezionati da loro stessi (es. pazienti inviati ad un laboratorio al quale gli investigatori solitamente inviano i loro pazienti) i risultati sulla prevalenza, sensibilità specificità, LR e limiti di concordanza potrebbero essere inficiati e potrebbero non essere applicabili ad altre popolazioni di pazienti. Negli studi ci dovrebbe essere una chiara descrizione su chi invia i pazienti al centro del sonno, sulla quantità di pazienti inviati e sul tipo di centro del sonno (comunità, università, o Veterans Affairs Hospital). Se alcuni pazienti non vengono inclusi nello studio, dovrebbero essere usati criteri di esclusione e giustificata la scelta di questi criteri. Bisogna inoltre descrivere il numero di pazienti inviati nel periodo di reclutamento, il numero di pazienti che sono eleggibili per partecipare allo studio, il numero di pazienti realmente inclusi nello studio (elencando i pazienti esclusi e le ragioni dell'esclusione), quanti pazienti hanno completato lo studio (con il numero ed i motivi del drop-out) ed infine la percentuale di pazienti per i quali i dati non sono interpretabili.

##### 1.4.2.2. Descrizione dei metodi di monitoraggio con sistemi portatili e con la polisunnografia

Dovrebbe essere riportata una descrizione chiara degli apparecchi usati per registrare i segnali con i sistemi portatili e con il polisunnografo. La definizione degli eventi respiratori dovrebbe essere fornita nei dettagli per permettere ad altri di ripetere i metodi. Se vengono adoperati indici quali l'RDI, la definizione dovrebbe essere data chiaramente. Dovrebbe essere dichiarato inoltre il metodo con cui è stata effettuata l'analisi (automatica, manuale, durata delle epoche, tipo di monitoraggio o analisi automatica con revisione manuale che permette di modificare i risultati). Quando i metodi rappresentano una parte importante nella valutazione del sistema portatile non è accettabile usare uno studio precedente come riferimento.

##### 1.4.2.3. Riproducibilità

Esistono molte fonti di variabilità che fanno sì che i risultati di uno studio non siano generalizzabili. Un aspetto molto importante è la variabilità nel riconoscimento degli eventi da parte dell'operatore. Se il sistema di monitoraggio prevede un'analisi automatica (basata su algoritmi) che non permette la revisione manuale, allora questo aspetto non rappresenta un problema. Tuttavia se c'è la possibilità di intervenire manualmente nell'analisi allora dovrebbe essere valutata l'abilità di due operatori, e quella di un singolo operatore, a ripetere la revisione di una precedente analisi e ciò dovrebbe essere riportato nell'articolo. La riproducibilità all'interno di una analisi e tra le analisi viene riportata come coefficiente  $\kappa$ . Un'altra fonte di variabilità che bisogna prendere in considerazione è la differenza da notte a notte. Questa forma di variabilità può essere analizzata usando l'approccio di Bland ed Altman<sup>61</sup>, o il coefficiente di correlazione intraclasse.<sup>62</sup> I coefficienti di correlazione di Pearson vengono utilizzati spesso per valutare la riproducibilità, ma, come abbiamo detto in precedenza per la concordanza tra due metodi diversi, questo approccio non è raccomandato.

##### 1.4.2.4. Evitare errori

Ci sono diversi aspetti in uno studio che possono indicare se i risultati dello stesso sono attendibili. Esiste evidenza che gli studi sui test diagnostici, che presentano difetti nel protocollo, tendono a sovrastimare l'accuratezza del test.<sup>63</sup> Per evitare errori di selezione bisognerebbe usare una serie appropriata di pazienti consecutivi. Errori di verifica si hanno quando i risultati di un monitoraggio determinano se effettuare il monitoraggio successivo. Questi errori si possono evitare assicurandosi che il test standard e il test diagnostico studiato vengano effettuati in tutti i pazienti eleggibili. È anche importante che il test diagnostico e il test standard di riferimento vengano analizzati separatamente, senza che l'operatore conosca i risultati dell'altro test.

L'analisi *post hoc* dei risultati permette all'investi-

gatore di ottimizzare l'utilità del test. Effettuare analisi multiple riduce l'affidabilità dei test statistici perché ogni volta che viene effettuata un'analisi c'è una probabilità casuale che il risultato sia positivo (l'uso di analisi multiple aumenta la probabilità di risultati spuri). Anche se esistono metodi statistici per correggere questi fattori, viene comunque raccomandato di selezionare prima dello studio i valori soglia per stabilire la positività o la negatività del test. Idealmente questi valori soglia vengono definiti in uno studio iniziale e poi confermati in uno studio prospettico indipendente. Un altro approccio consiste nell'individuare i valori soglia usando una parte di una popolazione di pazienti e di validarli nella rimanente parte di pazienti. Pochi studi hanno adottato questi approcci e cercato di correggere per le analisi multiple. Solo due studi<sup>30,35</sup> hanno validato i valori soglia usati per stabilire la probabilità di apnea del sonno.

#### 1.4.2.5. Presentazione dei risultati

I risultati degli studi sull'accuratezza di un test diagnostico dovrebbero indicare il numero di pazienti nei quali sono stati effettuati entrambi i test, stabilire chiaramente la prevalenza della patologia in questione, il numero di veri positivi e falsi positivi, veri negativi e falsi negativi, che permetteranno di calcolare sensibilità, specificità ed LRs, così come i valori predittivi positivi e negativi. Se per definire la presenza o l'assenza della malattia vengono utilizzati valori soglia diversi, gli autori dovrebbero dichiarare chiaramente l'effetto sulla prevalenza e le caratteristiche operative del test diagnostico. Se per un test diagnostico positivo/negativo vengono riportati diversi valori soglia gli effetti della variazione del valore soglia dovrebbero essere riportati con una curva ROC e/o il calcolo di un LR per ciascun valore soglia. Infine, come avviene con tutte le analisi statistiche, dovrebbero essere riportati gli intervalli di confidenza 95% per ciascuna stima (sensibilità, specificità, LRs). Le misure effettuate su numeri ridotti di pazienti sono imprecise come si può vedere dagli ampi intervalli di confidenza.

#### 1.5. Valutazione del livello di evidenza

La revisione dell'AHRQ<sup>7,8</sup> sulla diagnosi di apnea del sonno pubblicata nel 1999 è stata la prima revisione sistematica che valutava la qualità degli studi pubblicati. I metodi per valutare il livello di evidenza degli studi sui test diagnostici non sono molto utilizzati e gli autori del report dell'AHRQ hanno utilizzato un loro personale approccio.

Questi autori assegnavano dei punti se lo studio rispondeva a determinati criteri di qualità (un totale di 44 punti per 18 criteri). Con questo metodo decidevano di escludere il 20% degli articoli con i punteggi più bassi. La qualità degli studi sui test diagnostici è stata presa in considerazione anche da altre pubblicazioni. Il *Journal of the American Medical Association* ha pubblicato numerose guide per aiu-

tare il clinico a riconoscere gli studi clinici di elevata qualità. Le guide che riguardano i test diagnostici elencano alcuni importanti criteri per stabilire la validità di uno studio.<sup>64,65</sup> I criteri principali sono i seguenti: (1) il paragone con il test di riferimento standard è indipendente ed in cieco; (2) il campione di pazienti ai quali viene applicato il test include una casistica ampia. I criteri secondari sono i seguenti: (1) i risultati del test che viene valutato non influenzano la decisione di effettuare il test di riferimento standard (difetto di verifica); (2) i metodi di esecuzione del test vengono descritti in dettaglio in modo da poter essere ripetuti. Anche se gli articoli pubblicati dal *Journal of the American Medical Association* rappresentano una guida utile per i clinici, non forniscono una metodologia o un sistema per il rating degli studi. L'approccio proposto da Sackett et al<sup>9</sup> usava un numero piccolo ma essenziale di caratteristiche per il rating degli studi sui test diagnostici: (1) il paragone indipendente in cieco tra il test studiato ed il test di riferimento standard; (2) una casistica appropriata di pazienti consecutivi (evitare errori di selezione); (3) l'uso di un test di riferimento standard per tutti i pazienti dello studio. Questi criteri sono stati organizzati in livelli di evidenza (da I a V), dei quali il livello I è considerato il migliore.<sup>9</sup>

Il gruppo di studio dell'ATS/ASM/ACCP ha deciso di non applicare il metodo di valutazione della qualità degli studi pubblicati usato dall'AHRQ, per una serie di ragioni. In primo luogo non è chiaro il razionale per cui ad alcuni criteri è stato assegnato un punteggio elevato (es. studio randomizzato controllato, 10 punti; studio in cieco sullo stato clinico, 5 punti) e ad altri un punteggio basso (es. difetti di verifica [i risultati dello studio non determinano chi deve essere sottoposto a polisinnogramma], 1 punto; pazienti inclusi con un ampio spettro di severità dell'apnea del sonno, 1 punto). Secondo, per stabilire la qualità dello studio, è stato usato un criterio (disegno randomizzato controllato) che non è riportato né tra i criteri della serie del *Journal of the American Medical Association*, né tra i criteri dell'articolo di Sackett et al,<sup>9</sup> e a questo criterio è stato attribuito un punteggio 10 volte più grande di quello attribuito ad altri importanti criteri di qualità, come per esempio l'elusione dell'errore di verifica.<sup>7</sup> Esiste una importante differenza tra uno studio controllato randomizzato e l'esecuzione in ordine random del test con il sistema portatile e della polisinnografia standard. Quest'ultimo aspetto può essere importante, perché l'ordine di esecuzione dei test può influenzare i risultati. Per esempio, se durante la prima notte si effettua la polisinnografia standard, che dà risultati diversi dal monitoraggio con il sistema portatile, ciò potrebbe influenzare il confronto tra i risultati. Il gruppo di studio dell'ATS/ASM/ACCP, per valutare il livello di evidenza degli articoli inclusi in questa revisione sistematica, ha adattato il metodo proposto da Sackett et al,<sup>9</sup> (vedi paragrafo 2.3).

Dopo che il gruppo di studio ATS/ASM/ACCP

aveva completato la valutazione della letteratura sui sistemi di monitoraggio portatili, l'AHRQ ha prodotto un report intitolato "Sistemi per valutare il livello di evidenza scientifica".<sup>66</sup> Questo report sottolinea i seguenti cinque aspetti per valutare la qualità di articoli sui test diagnostici: (1) popolazione dello studio; (2) adeguata descrizione del test; (3) test standard di riferimento appropriato; (4) confronto tra i due test in cieco; (5) elusione degli errori di verifica. Noi assumiamo che il primo punto (popolazione dello studio) si riferisce ad una casistica ampia di pazienti consecutivi (elusione degli errori di selezione). Pertanto i criteri pubblicati da Sackett et al<sup>9</sup>, includono quattro di questi 5 aspetti (1, 3, 4, 5). Perché gli studi potessero essere inclusi in questa revisione i risultati ottenuti con il sistema di monitoraggio portatile dovevano essere stati paragonati a quelli ottenuti con una polisinnografia standard. Poiché non è stato possibile valutare i diversi metodi/definizioni relativi alla polisinnografia, tutti gli articoli erano considerati uguali per ciò che

riguarda questo aspetto. Il gruppo di studio ha valutato se negli articoli c'era una descrizione adeguata dei metodi usati per registrare i segnali con i sistemi portatili, dei metodi per definire un evento (inclusa la definizione degli eventi), e se i metodi usati per lo scoring degli eventi erano descritti adeguatamente (2.3.2). Pertanto questi due criteri di qualità si riflettono nel punteggio di qualità. Dei trenta studi con livello di evidenza I o II, solo uno studio non rispondeva ad entrambi i criteri ed aveva un livello di qualità pari a "d" (Tabella 2). Altri sette studi non rispondevano ad uno dei due criteri; uno di questi aveva un livello di qualità pari a "d", mentre gli altri tre erano di livello "a" o "b". La nostra opinione è che se avessimo incorporato l'elemento "adeguata descrizione del test" (punto 2 del report dell'AHRQ), per calcolare il livello di evidenza, i risultati e le conclusioni di questo report non sarebbero cambiati.

Oltre che a definire gli elementi per valutare la qualità dell'evidenza degli studi clinici randomizzati, degli studi osservazionali e degli studi sui test dia-

**Tabella 2—Numero di studi pubblicati sui sistemi di monitoraggio portatili per la diagnosi di apnea del sonno**

Tipo di monitoraggio	Setting dello studio	Livello di evidenza	Livello di qualità				Referenze	N. di studi totali
			a	b	c	d		
2	Laboratorio del sonno	I						
		II	1				25	
		IV	1			1	19* / 32†	
	Subtotale						3	
	Domicilio	I						
		II						
		IV	1	1			19* / 20	
	Subtotale						2	
	Totale						5	
3	Laboratorio del sonno	I	3				21, 22, 31	
		II	4	1			18‡, 23, 28, 29* / 27	
		IV				1	60	
	Subtotale						9	
	Domicilio	I						
		II	2				26, 29*	
		IV	1	1			24 / 30	
	Subtotale						4	
	Totale						13	
4	Laboratorio del sonno	I	1	1			36 / 45	
		II	10	3	1	1	10, 11, 12, 33,* 34, 43, 44, 52, 54, 57 / 13, 37, 55 / 48 / 35	
		IV	2	2	3	5	15, 38* / 14, 58 / 16, 47, 49 / 46, † 17, 51,* 50, 59	
	Subtotale						29	
	Domicilio	I		1			53 / 40	
		II	1		1		33* / 41	
		IV	1	1	2	1	38* / 39 / 42, 56, 46 / 51*	
	Subtotale						9	
	Totale						38	

\*Lo studio paragonava la polisinnografia effettuata in laboratorio e il monitoraggio con sistema portatile effettuato sia in laboratorio che a domicilio.  
 †Setting dello studio non chiaro; classificato come studio in laboratorio.  
 ‡Lo studio paragonava la polisinnografia effettuata in laboratorio e il monitoraggio con sistema portatile effettuato sia in laboratorio che a domicilio ma combinati in un unico gruppo per l'analisi. Classificato come laboratorio del sonno, visto che la maggior parte dei monitoraggi con sistema portatile veniva effettuata in laboratorio.

gnostici, il report dell'AHRQ<sup>7,8</sup> elencava aspetti importanti per valutare il livello di evidenza da applicare alle revisioni sistematiche. Per le revisioni sistematiche il report dell'AHRQ raccomandava di considerare i seguenti aspetti:

1. Razionale dello studio;
2. Strategia di ricerca;
3. Criteri di inclusione ed esclusione;
4. Interventi;
5. Outcome;
6. Estrazione di dati;
7. Qualità e validità dello studio;
8. Sintesi ed analisi dei dati;
9. Risultati;
10. Discussione;
11. Finanziamento.

In questa revisione sistematica noi abbiamo sottolineato il razionale dello studio (paragrafi 1.0, 4.1, 4.2) e la strategia della ricerca ed abbiamo inserito dei criteri di inclusione ed esclusione (paragrafo 2.1). Abbiamo revisionato gli studi che valutavano i monitoraggi con sistemi portatili effettuati nei laboratori del sonno (setting sorvegliato, solitamente in simultanea con una polisonnografia) e a casa dei pazienti (setting non sorvegliato). Gli outcome revisionati includono l'accuratezza diagnostica (sensibilità, specificità, LR), i risultati dei test su una popolazione con sospetta apnea del sonno (percentuale di pazienti con risultato positivo o negativo e percentuale di pazienti mal classificati dal monitoraggio col sistema portatile), i costi, le percentuali di fallimento e la riproducibilità. Abbiamo incluso anche una descrizione sul sistema di estrazione dei dati (paragrafo 2.2) e il processo dettagliato per valutare la qualità e la validità degli studi (paragrafo 2.3). Sono stati così sintetizzati i dati di 51 studi ed i risultati sono stati presentati in una serie di tabelle. Anche se non è stata eseguita una meta-analisi, dai dati forniti nelle tabelle è possibile determinare il livello di evidenza disponibile su alcuni particolari problemi e la qualità di tale evidenza. Inoltre è possibile valutare se i risultati degli studi di qualità elevata sono consistenti tra loro. Nel paragrafo 5.0 sono fornite inoltre delle raccomandazioni per migliorare la qualità della ricerca sui metodi diagnostici per l'apnea del sonno.

## 2.0. MATERIALI E METODI

### 2.1. Revisione della letteratura

Il gruppo di studio dell'ATS/ASM/ACCP ha incaricato il centro dell'RTI-UNC di condurre una revisione sistematica della letteratura e di riportare i dati degli studi più rilevanti in forma di abstract. Per condurre la revisione sistematica il gruppo dell'RTI-UNC ha applicato i metodi raccomandati, che enfatizzano una ricerca estesa e l'utilizzo di una procedura standard per la revisione degli articoli selezionati.

Una revisione sistematica della letteratura sulla diagnosi per

l'apnea del sonno è stata effettuata nel 1997 dall'AHRQ.<sup>7</sup> La nostra ricerca comprendeva invece gli articoli pubblicati dopo il 1997. La ricerca iniziale è stata completata il 26 Giugno 2001. È stata anche cercata la bibliografia di due review<sup>4,6</sup> dell'American Sleep Disorder Association per potere identificare gli articoli più rilevanti. Sono state utilizzate diverse strategie di ricerca che si basavano sullo screening, sulla diagnosi e sui costi. La ricerca veniva condotta usando le parole "screening" (inclusi i termini "riproducibilità dei risultati", "valore predittivo dei test" o "sensibilità e specificità") e "diagnosi" per trovare le citazioni che includevano i termini "sindrome dell'apnea da sonno", "apnea ostruttiva del sonno", "ossimetria", "polisonnografia", "monitoraggio fisiologico", "resistenza delle vie aeree", "sindrome dell'aumentata resistenza delle alte vie aeree", "indice di disturbo respiratorio", "autoset", "russamento" o "risvegli correlati agli eventi respiratori". Il termine "home care services" veniva anche usato per identificare le citazioni.

Per la parola "screening" la ricerca MEDLINE ha identificato 157 citazioni mentre per la parola "diagnosi" ne sono state identificate 180. Con i termini "Home care services" e "polisonnografia" sono stati identificate altre 14 citazioni. Per le ricerche sui costi, le parole "costi" e "analisi dei costi" sono state associate ai termini "analisi costo/beneficio", "assegnazione dei costi", "controllo dei costi", "risparmio nei costi", "divisione dei costi", "costo della malattia", "costi sanitari" e "spese sanitarie". La ricerca MEDLINE dal 1997 ha prodotto 35 citazioni.

I criteri di inclusione erano i seguenti:

- Paziente uomo/donna, età  $\geq 18$  anni con diagnosi di apnea del sonno;
- Studio pubblicato in lingua inglese, nessuna restrizione di razza o sesso;
- Utilizzo di sistemi portatili per la diagnosi;
- Polisonnografia o altri test accettabili utilizzati per la diagnosi di apnea del sonno;
- Dopo il completamento dello studio ogni gruppo analizzato era costituito da  $\geq 10$  pazienti.

I criteri di esclusione erano i seguenti:

- Studi nei bambini;
- Studi in lingua diversa dall'inglese;
- Review, meta-analisi, casi clinici, abstract, lettere ed editoriali.

Sono stati revisionati i titoli degli articoli che contenevano l'espressione "diagnosi dell'apnea del sonno" e i relativi abstract sono stati valutati per stabilirne la rilevanza. Se c'erano dubbi sui criteri di inclusione/esclusione veniva revisionato l'intero articolo. Gli elenchi delle referenze contenute in questi articoli venivano visionati per identificare altri possibili articoli da includere nella revisione.

### 2.2. Tabelle di evidenza

L'RTI-UNC ha lavorato congiuntamente con l'ERC per identificare gli aspetti chiave, per produrre un modello di abstract e per sviluppare un modulo con gli elementi per l'estrazione dei dati (vedi "Appendice"; disponibile on line al sito <http://www.chestjournal.org/cgi/content/full/124/4/1543/DC1>). Due revisori del centro successivamente hanno estratto i dati da ciascun studio. Questi revisori hanno poi confrontato i loro risultati, per ciascuno elemento sul modulo dell'estrazione dei dati, e per ciascuno studio, e, nei casi in cui discordavano, è stato raggiunto un accordo tra i revisori. Successivamente sono stati compilati i moduli con i dati finali e sono stati inviati ai membri dell'ERC che hanno deciso come sintetizzare le evidenze disponibili. L'ERC ha stabilito che la ricerca doveva includere gli articoli fino al dicembre 2001; ciò ha fatto sì che fossero identificati due ulteriori articoli che i membri dell'ERC hanno estratto.

### 2.3. Valutazione della qualità degli articoli

La valutazione della qualità degli studi è stata effettuata attraverso il rating di dieci caratteristiche diverse di ciascun articolo. Ciò ha permesso la classificazione del livello di evidenza dell'articolo come I, II, III, IV (basandosi su tre di queste caratteristiche). Un'ulteriore classificazione della qualità dello studio (a, b, c, d) si basava sulle caratteristiche rimanenti. Ciascuna delle dieci caratteristiche era valutata indipendentemente da due revisori dell'RTI-UNC e da due membri dell'ERC. Il livello di evidenza finale e il livello di qualità erano determinati dall'ERC, che si basava su definizioni standard preselezionate.

#### 2.3.1. Livello di evidenza (I, II, III e IV)

L'ERC, adottando il metodo di Sackett et al, si è basata sulla presenza o assenza di tre indicatori di qualità per l'assegnazione del livello di evidenza.<sup>9</sup> Le definizioni di questi livelli di evidenza sono elencate di seguito:

- I**, confronto in cieco, pazienti consecutivi, test di riferimento standard eseguito su tutti i pazienti;
- II**, confronto in cieco, pazienti non consecutivi, test di riferimento standard eseguito su tutti i pazienti;
- III**, confronto in cieco, pazienti consecutivi, test di riferimento standard non eseguito su tutti i pazienti;
- IV**, il test di riferimento standard non era eseguito in cieco o indipendentemente dal test studiato.

Le definizioni dei tre indicatori usati per assegnare il livello di evidenza erano le seguenti:

**Confronto in cieco:** l'analisi degli studi eseguiti con il sistema portatile e con la polisonnografia veniva effettuata separatamente senza che l'operatore, che analizzava un test, conoscesse i risultati dell'altro test. Se nell'articolo l'autore non specificava che l'analisi era stata effettuata in cieco lo studio non soddisfaceva il criterio.

**Pazienti consecutivi:** gli investigatori non partecipavano alla decisione di quali pazienti includere nello studio. Il criterio era soddisfatto se i pazienti venivano inviati a una clinica del sonno piuttosto che a un laboratorio del sonno (a meno che gli investigatori non dichiaravano chiaramente di non aver partecipato a selezionare i pazienti da inviare al laboratorio).

**Test di riferimento standard eseguito su tutti i pazienti:** tutti i pazienti inclusi nello studio dovevano essere stati sottoposti al monitoraggio con il sistema portatile e alla polisonnografia. Se i risultati di un test influenzavano la decisione di effettuare l'altro test il criterio non veniva soddisfatto.

#### 2.3.2. Valutazione della qualità (a, b, c, d)

Altri sette aspetti della metodologia dello studio venivano sottoposti a rating, e il livello di qualità era assegnato basandosi sul numero di aspetti (indicatori) per i quali i criteri erano soddisfatti. Sebbene l'assegnazione random del test fosse un importante indicatore, questo non era un criterio applicabile agli studi per i quali il monitoraggio con il sistema portatile veniva eseguito simultaneamente alla polisonnografia. Pertanto in alcuni casi per il rating dello studio venivano usati sei indicatori anziché sette. Il livello di qualità (da a a b) veniva basato sul numero di indicatori per i quali i criteri non erano soddisfatti, così come segue:

- a**, criteri non soddisfatti per uno o per nessun indicatore;
- b**, criteri non soddisfatti per due indicatori;
- c**, criteri non soddisfatti per tre indicatori;
- d**, criteri non soddisfatti per quattro indicatori.

I sette indicatori, e le rispettive definizioni, sono riportati qui di seguito:

1. **Reclutamento prospettico dei pazienti:** il monitoraggio con il sistema portatile e la polisonnografia erano eseguiti appena il paziente veniva arruolato nello studio e non veniva eseguita la revisione di una serie di pazienti studiati precedentemente.

2. **Ordine random dei test:** l'ordine di esecuzione del monitoraggio con il sistema portatile e della polisonnografia era random e non a discrezione dell'investigatore. Se il monitoraggio con il sistema portatile veniva eseguito in contemporanea con il polisonnogramma questo indicatore non veniva valutato.

3. **Bassa percentuale di dati persi (< 10%):** la percentuale di pazienti per i quali i risultati non potevano essere confrontati per la perdita dei dati della polisonnografia o del monitoraggio con sistema portatile era < 10%.

4. **Alta percentuale di pazienti che completavano lo studio (> 90%):** dei pazienti inizialmente arruolati nello studio (senza contare le esclusioni a priori) > 90% completava il protocollo di studio.

5. **Metodi/definizioni della polisonnografia descritti in dettaglio:** i metodi della polisonnografia devono includere:

- a. *descrizione della strumentazione usata;*
- b. *definizione e criteri relativi a tutti i tipi di eventi classificati e usati per il confronto.*

6. **Metodi/definizioni del monitoraggio con sistema portatile descritti in dettaglio:**

i metodi della polisonnografia devono includere:

- a. *descrizione della strumentazione usata;*
- b. *definizione e criteri relativi a tutti i tipi di eventi classificati e usati per il confronto.*

7. **Sistema di analisi del test eseguito con sistema portatile descritto:** include una chiara dichiarazione sul sistema di analisi usato, manuale o automatico, e, in caso di analisi automatica, se c'è stata la revisione manuale.

### 2.4. Sintesi delle evidenze sul monitoraggio con sistema portatile

Per valutare l'accuratezza diagnostica dei sistemi portatili in quasi tutti gli studi si sceglieva di riportare la sensibilità e la specificità. Molti studi prendevano in considerazione valori soglia multipli per definire un risultato positivo ed ottenere la combinazione della specificità e della sensibilità. Nel tentativo di stabilire se i monitoraggi con sistema portatili potevano ridurre la probabilità che un paziente avesse l'apnea del sonno (paragrafo 4.1.1), negli articoli veniva ricercata la migliore sensibilità riportata, visto che ciò forniva il più basso numero di falsi negativi ed il più basso LR. Nel caso in cui venivano riportate varie combinazioni di sensibilità e specificità, e i valori di due sensibilità erano simili, la "migliore sensibilità riportata" era quella associata al più alto valore di specificità corrispondente. Se gli autori riportavano una definizione diversa di apnea del sonno, il gruppo di studio sceglieva la definizione di  $AHI \leq 15$  (quando veniva riportato) perché si pensava che la maggior parte dei clinici fosse interessata a sapere che probabilità aveva il paziente di avere un AHI al di sotto di questo valore, proprio perché questo è il valore che influenza la scelta terapeutica. Al contrario, quando il gruppo di studio doveva stabilire se il monitoraggio con il sistema portatile aumentava la probabilità che il paziente avesse l'apnea del sonno, veniva allora ricercata la migliore specificità riportata, perché ciò forniva il più alto valore di LR ed il più basso numero di falsi positivi. Nel caso in cui varie combinazioni di specificità e sensibilità venivano riportate, ed i valori di due specificità erano simili, la "migliore specificità riportata" era considerata quella associata alla più alta sensibilità.

Quando si usavano diversi valori soglia per la migliore sensibilità riportata e la migliore specificità riportata (diversi punti sulla curva ROC), allora alcuni pazienti soddisfacevano l'uno o l'altro criterio, ma alcuni non soddisfacevano nessun criterio e quindi i risultati erano indeterminati. È importante esaminare la percentuale di pazienti che soddisfa i criteri per un risultato negativo (migliore sensibilità riportata) e la percentuale di pazienti che soddisfa il criterio per un risultato positivo (migliore specificità

**Tabella 3—Sensibilità più alta riportata ed LR negativo\***

Tipo di monitoraggio	Setting dello studio	Livello di evidenza	Livello di qualità	Primo autore	Refere- nza	Paz. n.	Prev	LR neg. e CI			Migliore sens. e CI/Corr. Spec.				Paz. con PM risult. neg. %	Falsi neg. %	
								Basso		Alto	Basso		Alto				AHI Dfn
								LR	CI	CI	Sens.	CI	CI	Spec.			
2	Domicilio Laboratorio del sonno	IV	b	Portier F	20	78	47,0	0,19	0,10	0,38	81	68,3	93,7	98	15	61	15
		II	a	Mykytyn I	25	20	50,0	0,22	0,06	0,78	80	55,2	100	90	10	55	18
3	Domicilio Laboratorio del sonno	IV	d	Orr WC	32	40	62,5	†	‡	‡	100	‡	‡	93	15	35	0
		II	a	White D	29	70	61,4	0,13	0,05	0,34	91	82,4	99,6	70	10	32	17
				Parra O	26	89	84,3	0,15	0,04	0,52	95	90,1	99,9	33	10	9	45
		IV	a	Ancoli-Israel S	24	34	73,5	†	‡	‡	100	‡	‡	66	10	18	0
			b	Whittle A	30	58	55,2	0,43	0,22	0,85	75	60,0	90,0	58	15	40	35
		I	a	Zucconi M	21	29	65,5	†	‡	‡	100	‡	‡	100	10	34	0
				Emsellem H	22	63	61,9	0,05	0,01	0,21	95	88,2	100	96	5	40	8
				Verse T	31	53	47,2	0,08	0,02	0,3	92	81,8	100	96	10	54	7
		II	a	White D	29	30	63,3	†	‡	‡	100	‡	‡	64	10	24	0
				Ficker J	28	51	56,9	0,03	0,00	0,24	97	90,8	100	100	10	45	4
4	Domicilio Laboratorio del sonno	IV	d	Redline S	18	25	84,0	0,05	0,01	0,32	95	85,7	100	100	10	20	21
				Man G	23	104	26,9	0,15	0,06	0,37	86	73,1	98,9	95	15	73	5
		IV	b	Ballester E	27	116	24,1	0,12	0,04	0,34	89	77,4	100	92	10	72	4
		I	d	Claman D	60	42	50,0	0,15	0,05	0,43	86	71,2	100	95	15	54	13
			b	Series F	53	240	45,8	0,04	0,01	0,16	98	95,4	100	48	10	27	3
		II	d	Gyulay S	40	98	43,9	0,14	0,04	0,42	93	85,4	100	51	15	32	10
			a	Baltzan M	33	66	41,2	0,31	0,09	1,06	90	78,7	100	32	15	23	18
		IV	c	Golpe R	41	116	61,2	0,33	0,18	0,62	84	75,5	92,5	48	10	28	34
			b	Schafer H	39	114	70,2	0,12	0,04	0,34	95	90,2	99,8	41	10	16	22
			c	Williams A	42	36	55,5	0,35	0,19	0,64	65	44,1	85,9	100	10	64	30
			d	Ryan P	56	69	46,3	0,69	0,55	0,87	31	15,0	47,0	100	15	86	37
		I	d	Olson L	51	793	44,0	0,17	0,09	0,31	97	95,2	98,8	18	15	11	12
			a	Vazquez J	36	241	59,0	0,04	0,01	0,10	97	94,2	99,8	80	10	35	5
		II	b	Chiner E	45	275	77,7	0,24	0,17	0,33	82	76,8	87,2	76	15	31	45
			a	Koziej M	44	56	66,1	†	‡	‡	100	‡	‡	63	10	21	0
				Esnaola S	43	150	60,0	0,03	0,01	0,11	98	95,1	100	78	10	32	4
				Gugger M	11	67	46,3	0,04	0,01	0,29	97	91,0	100	77	20	43	3
				Mayer P	12	95	93,7	0,06	0,01	0,25	97	93,5	100	50	5	6	47
				Stoohs R	34	56	46,4	0,08	0,02	0,30	92	81,6	100	97	10	56	7
				Baltzan M	33	86	41,2	0,09	0,01	0,64	97	91,4	100	32	15	20	6
		Zamarron C	52	233	53,2	0,09	0,05	0,19	94	89,8	98,2	65	10	34	10		
		Issa F	57	129	54,3	0,12	0,06	0,23	89	81,7	96,3	95	7	49	12		
		Douglas N	54	200	45,5	0,36	0,27	0,48	67	57,3	76,7	92	15	65	23		
	b	Kiely J	13	36	36,1	†	‡	‡	100	‡	‡	92	15	59	0		
		Yamashiro Y	37	269	50,9	0,08	0,04	0,16	94	90,0	98,0	74	5	39	8		
		Rauscher H	55	116	40,5	0,10	0,03	0,30	94	87,2	100	62	10	39	6		
	c	Cooper B§	48	41	29,2	0,29	0,11	0,78	75	50,5	99,5	86	15	68	11		
	d	Gurubhagavatula I	35	359	69,4	0,15	0,11	0,21	86	81,7	90,3	91	5	38	26		
IV	a	Bagnato M	15	56	78,6	0,05	0,01	0,44	98	93,9	100	40	10	10	16		
	b	Bradley P	14	31	48,4	†	‡	‡	100	‡	‡	92	15	48	0		
		Zamarron C	58	197	46,2	0,21	0,14	0,32	81	72,9	89,1	91	10	58	15		
	c	Levy P	49	301	64,1	0,04	0,02	0,12	98	96,0	100	46	15	18	7		
		Gugger M	16	27	70,3	0,20	0,07	0,54	82	64,7	99,3	90	20	39	32		
	d	Fleury B	17	38	46,9	†	‡	‡	100	‡	‡	87	10	46	0		
		Pepin J	50	26	57,7	0,11	0,02	1,11	95	84,0	100	45	5	22	13		
		Olson L	51	113	22	0,11	0,03	0,42	92	81,3	100	73	15	59	3		
		Bonsignore G	46	83	56,6	0,26	0,16	0,42	75	62,0	87,0	100	10	58	25		

\*Totale di 49 studi su 46 studi di ricerca. Cinque studi non sono stati inclusi (15, 33, 35, 37, 47) perché i dati di sensibilità/specificità non erano riportati. Pts = pazienti; Prev = prevalenza; Neg LR = LR negativo; CI = intervallo di confidenza; Basso CI = intervallo di confidenza più basso; Alto CI = intervallo di confidenza più alto; Migliore sens. = migliore sensibilità riportata nell'articolo (per un cutoff di AHI ≤ 15, tranne se non riportato) con il più basso LR calcolato; Corr. Spec. = specificità del test con la migliore sensibilità riportata; Sens. = sensibilità; Spec. = specificità; AHI Dfn = valore cutoff di AHI (dalla polisomnografia) usato per definire i pazienti con apnea del sonno alla migliore sensibilità riportata. Vedi Tabella 1 per le altre abbreviazioni non presenti nel testo.

†Il totale è 0 perché la sensibilità è uguale a 100%.

‡Non può essere calcolato.

§Non effettuato simultaneamente alla PSG.

||Indice di apnea, non AHI.

**Tabella 4—Sensibilità più alta riportata ed LR positivo\***

Tipo di monitoraggio	Setting dello studio	Livello di evidenza	Livello di qualità	Primo autore	Referenza	Paz. n.	Prev	LR pos. e CI				Migliore Spec. e CI/Corr. Sens.				Paz. con PM risult. pos. %	Falsi pos. %	
								Basso		Alto		Basso		Alto				
								LR	CI	CI	Spec.	CI	CI	Sens.	AHI Dfn			
2	Domicilio	IV	b	Portier F	20	78	47	40,5	4,8	343	98	93,7	100	81	15	39	3	
	Laboratorio del sonno	II	a	Mykytyn I	25	20	50	8	1,2	52,4	90	71,4	100	80	10	45	11	
3	Domicilio	IV	d	Orr WC	32	40	62,5	14,3	2,3	90,4	93	80,1	100	100	15	65	4	
		II	a	Parra O	26	89	84,3	9	1,3	61,0	93	79,6	100	63	10	54	2	
	Laboratorio del sonno	IV	a	Ancoli-Israel S	24	34	73,5	2,9	1,2	7,3	66	35,1	96,9	100	10	82	11	
			b	Whittle A	30	58	55,2	1,8	1,1	2,9	58	39,0	77,0	75	15	60	31	
		I	a	Zucconi M	21	29	65,5	†	‡	‡	100	‡	‡	100	10	66	0	
				Verse T	31	53	47,2	23,3	3,7	141	96	88,7	100	92	10	46	5	
		II	a	Emsellem H	22	63	61,9	23,8	3,3	168	96	88,2	100	95	5	60	2	
				Ficker J	28	51	56,9	†	‡	‡	100	‡	‡	97	10	55	0	
				Redline S	18	25	84	†	‡	‡	100	‡	‡	95	10	80	0	
				Man G	23	104	26,9	17,2	6,4	46,3	95	90,1	99,9	86	15	27	14	
	White D		29	30	43,3	6,4	1,3	32,3	88	68,8	100	77	20	40	17			
	Ballester E		27	116	24,1	11,1	5,4	22,8	92	86,3	97,7	89	10	28	22			
4	Domicilio	IV	d	Claman D	60	42	50	17,2	2,6	112	95	85,7	100	86	15	46	6	
			b	Series F	53	240	45,8	1,9	1,6	2,2	48	39,4	56,6	98	10	73	39	
		II	d	Gulay S	40	98	43,9	20	3,1	129	98	94,3	100	40	15	18	6	
			a	Baltzan M	33	66	41,2	10,3	1,7	63,5	97	91,6	100	31	15	14	12	
		IV	c	Golpe R	41	116	61,2	10,7	2,0	56,8	97	92,0	100	32	10	21	6	
			b	Schafer H	39	114	70,2	4,0	1,8	9,1	85	73,0	97,0	60	10	47	10	
	Laboratorio del sonno	c		Williams A	42	36	55,5	†	‡	‡	100	‡	‡	65	10§	36	0	
				Ryan P	56	69	46,4	†	‡	‡	100	‡	‡	31	15	14	0	
		d		Olson L	51	793	44	1,8	1,6	2,0	54	49,4	58,6	83	5	62	41	
			a	Vazquez J	36	241	49	8,2	5,1	13,2	88	82,3	93,7	98	15	54	11	
		II	b	Chiner E	45	275	77,7	8,9	3,5	22,1	93	86,6	99,4	62	15	50	3	
			a	Issa F	44	129	31	45	10,5	193	98	95,1	100	90	20	29	5	
				Esnaola S	43	150	60	34,5	5,8	203	98	94,5	100	69	10	42	2	
				Stoohs R	34	56	46,4	30,7	4,0	235	97	90,9	100	92	10	44	4	
				Douglas N	54	200	45,5	15,8	6,0	51,6	97	93,8	100	52,7	10	26	6	
				Baltzan M	33	86	41,2	14	2,0	98,4	98	94,1	100	28	15	13	9	
		IV	a		Mayer P	12	95	54,7	11,2	3,8	33,8	93	85,4	100	79	30	46	7
					Gugger M	11	67	46,3	4,2	2,3	7,7	77	63,3	90,7	97	20	57	22
c			Zamarron C	52	233	53,2	3,6	2,3	5,6	84	77,1	90,9	57	10	38	20		
			Koziej M	44	56	66,1	2,7	1,5	4,9	63	41,3	84,7	100	10	79	16		
b			Kiely J	13	36	36,1	12,5	3,1	50,0	92	80,9	100	100	15	41	12		
			Yamashiro Y	37	269	50,9	3,6	2,7	4,8	74	66,5	81,5	94	5	61	21		
d			Rauscher H	55	116	40,5	2,5	1,8	3,4	62	50,5	73,5	94	10	61	37		
			Cooper B	48	41	48,8	12	1,8	79,0	95	85,7	100	60	5	32	8		
a			Gurubhagavatula I	35	359	33	18	10,3	31,3	95	92,2	97,8	90	30	33	10		
			Bagnato M	15	56	50,7	2,4	1,5	3,8	60	41,7	78,3	97	20	69	29		
		Bradley P	14	31	48,4	12,5	2,4	65,9	92	78,7	100	100	15	52	8			
		Zamarron C	58	197	46,2	9,0	4,9	16,6	91	85,6	96,4	81	10	42	12			
		Levy P	49	301	64,1	†	‡	‡	100	‡	‡	46	15	30	0			
		Gugger M	16	27	70,3	8,2	1,0	66,0	90	69,2	100	82	20	61	5			
b		Pepin J	50	26	57,7	†	‡	‡	100	‡	‡	50	5	29	0			
		Bonsignore G	46	83	56,6	†	‡	‡	100	‡	‡	75	10	42	0			
		Fleury B	17	38	46,9	7,7	2,2	26,3	87	71,0	100	100	10§	54	13			
		Olson L	51	113	22	6	1,3	27,1	98	95,1	100	12	15	4	37			

\*Totale di 49 studi su 46 studi di ricerca. Cinque studi non sono stati inclusi (15, 33, 35, 37, 47) perché i dati di sensibilità/specificità non erano riportati. Pos LR = LR se il test con sistema portatile dava risultati positivi; Migliore spec. = migliore specificità riportata nell'articolo con il più alto LR calcolato; Corr. Sens. = sensibilità del test con la migliore specificità riportata. Vedi Tabella 1 e 3 per le altre abbreviazioni.

†Il totale è infinito perché la specificità è uguale a 100%.

‡Non può essere calcolato.

§Indice di apnea, non AHI.

||Non effettuato simultaneamente alla PSG.

riportata); queste percentuali sono riportate per ciascun studio nelle Tabelle 3 e 4. È anche importante determinare la percentuale di pazienti che soddisfano i criteri per i quali si sono avuti risultati falsi (classificati male dal test diagnostico). Questo dato, riportato in Tabella 3 e 4, è influenzato dalla prevalenza così come dalle caratteristiche operative del test (sensibilità, specificità ed LR). Gli studi, sintetizzati per il valore di LR, il tipo di monitoraggio, il setting, il livello di evidenza, per la migliore sensibilità riportata e la migliore specificità riportata, sono mostrati nelle Tabelle 6 e 7, rispettivamente (disponibili on line, <http://www.chestjournal.org/cgi/content/full/124/4/1543/DC1>). Un'altra sintesi simile degli articoli, con entrambi i valori alti e bassi di LR, è presentata in Tabella 8 (disponibile on line, <http://www.chestjournal.org/cgi/content/full/124/4/1543/DC1>).

Idealmente, un test eseguito con un sistema portatile dovrebbe avere un singolo valore cut-off che abbia una elevata sensibilità e specificità, cosicché i pazienti possano essere classificati come positivi o negativi, senza "zone grigie". Per stabilire se alcuni sistemi di monitoraggio si avvicinano al "test ideale", con valori cut-off che minimizzano il numero di pazienti nella zona grigia e la percentuale di errore, sono stati esaminati gli studi con entrambi un alto e un basso LR calcolato (i risultati sono

riportati in Tabella 5 e sintetizzati nel paragrafo 4.1.3). Come le Tabelle 3 e 4, la Tabella 5 evidenzia la percentuale di pazienti con un risultato falso e le percentuali di pazienti che non soddisfacciano i criteri per un risultato positivo/negativo (quelli nella zona grigia). Se lo studio aveva un singolo valore soglia per la migliore sensibilità e specificità, allora la percentuale di pazienti senza un risultato negativo o positivo era pari a zero.

In Tabella 3 e 4 sono riportati gli intervalli di confidenza 95% per la migliore sensibilità riportata e la migliore specificità riportata, rispettivamente, e per i corrispondenti valori di LR. Gli intervalli di confidenza 95% erano calcolati dalla sensibilità, specificità, prevalenza e numero di pazienti riportati, secondo il metodo di Simel et al.<sup>67</sup> In alcuni studi la prevalenza non era riportata ma stimata dalle figure.

Nel valutare gli studi sui sistemi portatili, per evitare errori nell'interpretazione dei dati è importante considerare i seguenti punti:

- Di quanto la migliore sensibilità riportata (basso LR) riduce la probabilità che il paziente sia affetto da apnea del sonno?
- Di quanto la migliore specificità riportata (alto LR) aumenta la probabilità che il paziente sia affetto da apnea del sonno?
- Il monitoraggio con il sistema portatile può sostanzialmente

**Tabella 5—Studi con alto (> 0,5) e basso (< 0,2) LR\***

Tipo di monitoraggio	Setting dello studio	Livello di evidenza	Livello di qualità	Primo autore	Risult. PM negativo				Risult. PM positivo				Paz. con falso pos. o neg. %	Paz. senza risult. neg. o pos. %				
					Refe- renza	Paz. n.	LR neg.	Migl. Sens.	Spec. Corr.	Paz. con neg. %	LR pos.	Migl. Sens.			Spec. Corr.	Paz. con pos. %		
2	Domicilio Laboratorio del sonno	IV	b	Portier F	20	78	0,19	81	98	61	40,5	98	81	39	10	0		
				Orr WC	32	40	†	100	93	35	14,3	93	100	65	3	0		
3	Domicilio Laboratorio del sonno	II	a	White D	29	70	0,13	91	70	32	5,1	83	86	46	16	22		
				Parra O	26	89	0,15	95	33	9	9	93	63	54	5	37		
				Zucconi M	21	29	†	100	100	34	‡	100	100	66	0	0		
				Emsellem H	22	63	0,05	95	96	40	23,8	96	95	60	5	0		
				Verse T	31	53	0,08	92	96	54	23,3	96	92	46	6	0		
	Laboratorio del sonno	II	a	White D	29	30	†	100	64	24	6,4	88	77	40	7	36		
				Ficker J	28	51	0,03	97	100	45	‡	100	97	55	2	0		
				Redline S	18	25	0,05	95	100	20	‡	100	95	80	4	0		
				Man G	23	104	0,15	86	95	73	17,2	95	86	27	7	0		
				Ballester E	27	116	0,12	89	92	72	11,1	92	89	28	9	0		
4	Domicilio Laboratorio del sonno	IV	d	Claman D	60	42	0,15	86	95	54	17,2	95	86	46	10	0		
				Gyulay S	40	98	0,14	93	51	32	20	98	40	19	4	49		
				Vazquez J	36	241	0,04	97	80	35	8,2	88	98	54	8	11		
				II	Esnaola S	43	150	0,03	98	78	32	34,5	98	69	42	2	26	
					Mayer P	12	95	0,06	97	50	6	11,2	93	79	46	6	48	
					Stoohs R	34	56	0,08	92	97	56	30,3	97	92	44	5	0	
					Baltzan M	33	86	0,09	97	32	20	14	98	28	13	2	67	
					Issa F	57	129	0,12	89	95	49	45	98	90	29	7	22	
				IV	b	Kiely J	13	36	†	100	92	59	12,5	92	100	41	5	0
						Gurubhagavatula I	35	359	0,15	86	91	38	18	95	90	33	13	29
						Bradley P	14	31	†	100	92	48	12,5	92	100	52	4	0
				IV	c	Levy P	49	301	0,04	98	46	18	‡	100	46	30	1	52
						Gugger M	16	27	0,20	82	90	39	8,2	90	82	61	16	0
Fleury B	17	38	†			100	87	46	7,7	87	100	54	7	0				
Pepin J	50	26	0,11			95	45	22	‡	100	50	29	3	49				
IV	d	Olson L	51	113	0,11	92	73	59	6	98	12	4	3	37				

\*Totale di 28 studi su 27 studi di ricerca. Vedi Tabelle 1 e 3 per le abbreviazioni non usate nel testo.

†Il totale è 0 perché la sensibilità è uguale a 100%.

‡Il total è infinito perché la specificità è uguale a 100%.

ridurre (se il risultato del test è negativo) e aumentare (se il risultato del test è positivo) la probabilità che un paziente sia affetto da apnea del sonno?

- Quale percentuale di pazienti, nello studio, supera i valori soglia per stabilire che il risultato di un test è positivo o negativo?
- Qual è la percentuale di pazienti per i quali, anche superando i valori soglia, la diagnosi è sbagliata (risultati falsi)?
- Che probabilità ci sono che i risultati dello studio siano validi (livello di evidenza e qualità)?
- Quanto sono precise le stime (ampiezza degli intervalli di confidenza) della sensibilità, specificità, e del LRs ?
- Il valore soglia di RDI, usato per definire un risultato positivo e negativo, era selezionato prima di iniziare lo studio o dopo l'analisi (retrospettivamente)?

È difficile da comprendere il perché dell'utilizzo di diversi valori soglia di RDI per valutare la positività/negatività di un test e per definire la migliore sensibilità riportata e specificità.

### 3.0. RISULTATI DELLA RICERCA BIBLIOGRAFICA

#### 3.1. Numero degli articoli revisionati/scartati

La ricerca bibliografica iniziale ha permesso di identificare 59 articoli originali che soddisfacevano i criteri di inclusione. Di questi, 46 articoli sono stati selezionati per la revisione.<sup>10-27,29-34,36-57</sup> Tredici articoli sono stati scartati<sup>68-80</sup> per le seguenti ragioni: erano studi su sistemi di monitoraggio più vecchi per i quali erano disponibili pubblicazioni più recenti (un articolo) o studi su apparecchi non più in commercio; erano studi che valutavano sistemi non portatili (cinque articoli); erano studi che valutavano sistemi non comunemente usati o non disponibili (quattro articoli); non erano studi su sistemi di monitoraggio di segnali fisiologici (un articolo); erano studi che esaminavano pazienti che erano stati sottoposti a intervento chirurgico (un articolo).

La ricerca in un primo momento è stata estesa fino al 31.10.2001, permettendo di identificare altri tre articoli,<sup>28,58,60</sup> e successivamente fino al 31.12.2001, permettendo di identificare altri due articoli,<sup>35,59</sup> per un totale di 51 articoli inclusi nella review.

#### 3.2. Tipo di monitoraggio (domicilio vs laboratorio del sonno) e livello di evidenza

Quattro studi descrivevano 5 gruppi di pazienti sottoposti a monitoraggio di tipo 2 (polisonnografia completa) (Tabella 2). Tre di questi quattro studi erano stati classificati di livello di evidenza IV e uno di livello II.

Dodici studi descrivevano 14 gruppi di pazienti sottoposti a monitoraggio di tipo 3. In cinque di questi il monitoraggio era effettuato a domicilio e in cinque in laboratorio. In totale, questi studi erano di qualità più alta (domiciliari: 3 studi di livello II; in laboratorio: 3 studi di livello I e 5 studi di livelli II) rispetto agli studi con monitoraggio di tipo 2 (Tabella 2).

La maggior parte degli studi pubblicati erano

basati sul monitoraggio di tipo 4 (35 su 51). Di questi, 29 includevano pazienti studiati in laboratorio e 9 pazienti studiati a domicilio. Per circa il 50% degli studi sul monitoraggio di tipo 4 il livello di evidenza era I o II (Tabella 2).

#### 3.3. Sensori per gli eventi respiratori

La classificazione dei sistemi portatili, in base al tipo di monitoraggio, potrebbe non essere molto rilevante. È chiaro che i monitoraggi di tipo 2 sono diversi dagli altri monitoraggi, perché danno la possibilità di registrare i segnali di EEG ed EMG. La distinzione tra il monitoraggio di tipo 3 e quello di tipo 4 è meno chiara. Senza tener conto del tipo di monitoraggio, generalmente, per identificare un evento respiratorio, ci si avvale di un solo segnale; occasionalmente viene usato un secondo segnale e raramente un terzo. Tuttavia, per tutti e tre i tipi di monitoraggio esiste la possibilità di avvalersi di più canali ed in molti sistemi è disponibile anche un sensore di posizione. Piuttosto che classificare i tipi di monitoraggio potrebbe essere più importante distinguere i metodi di rilevazione di ciascun segnale monitorato (es. il flusso aereo misurato con termistore o con cannula nasale e i vari metodi per misurare la saturazione di ossigeno) e il modo in cui i segnali vengono adoperati per identificare gli eventi respiratori. In Tabella 1 vengono raggruppati gli studi in base al segnale primario utilizzato per identificare gli eventi respiratori durante il sonno. Viene indicato inoltre il metodo per definire gli eventi nel polisonnogramma, le differenze nella frequenza di campionamento della saturazione di ossigeno (quando la saturazione di ossigeno veniva utilizzata per definire un evento) e i vari sistemi di analisi effettuati con gli apparecchi portatili (automatica, manuale o combinata). Il segnale più comunemente usato dai sistemi portatili era il flusso aereo misurato attraverso un termistore; 10 dei 15 studi nei quali si utilizzava questo segnale come segnale primario per lo scoring degli eventi erano classificati di livello I o II.

Solo due di questi studi si avvalevano degli stessi criteri per definire un evento respiratorio (Tabella 1). Il flusso misurato attraverso la cannula nasale veniva utilizzato in 8 studi (4 di livello II di evidenza). Tutti e quattro gli studi di livello I-II si avvalevano dello stesso sistema portatile e degli stessi criteri per definire un evento respiratorio. La saturazione di ossigeno veniva usata come segnale primario per definire l'evento respiratorio in 22 studi, 13 dei quali erano di livello I e II. Questi 13 studi si avvalevano di 11 criteri diversi per classificare gli eventi. Negli studi che si avvalevano del segnale di flusso misurato con la cannula nasale, o della saturazione di ossigeno, l'analisi era più comunemente eseguita in modo automatico, mentre negli studi nei quali il flusso veniva misurato con un termistore veniva eseguita l'analisi manuale o combinata.

### 3.4. Analisi di Bland ed Altman

Alcuni studi confrontano i dati ottenuti col monitoraggio portatile e quelli con la polisonnografia usando la media delle differenze e i limiti di concordanza (Tabella 9; dati disponibili on line, <http://www.chestjournal.org/cgi/content/full/124/4/1543/DCI>). In totale 24 studi utilizzavano questo tipo di analisi (monitoraggio di tipo 2, 2 studi; monitoraggio di tipo 3, 6 studi; monitoraggio di tipo 4, 16 studi). Alcuni autori riportavano invece gli intervalli di confidenza piuttosto che i limiti di concordanza (Tabella 9). I limiti di concordanza tendevano a essere piuttosto alti suggerendo che la concordanza tra i due metodi non era particolarmente buona, ma i limiti erano particolarmente ampi per valori più alti di RDI, per i quali era meno importante ottenere lo stesso numero, come lo era per valori di RDI più bassi. I limiti di concordanza possono essere aggiustati utilizzando una trasformazione logaritmica delle differenze, ma gli autori lo fanno raramente, pertanto è difficile interpretare questi dati e fare delle raccomandazioni sui sistemi di monitoraggio portatili basandosi su questo solo parametro.

## 4.0. DISCUSSIONE SULLE EVIDENZE DISPONIBILI SUI SISTEMI DI MONITORAGGIO PORTATILI

### 4.1. Argomenti primari

Molteplici sono le cause di variabilità nei risultati del confronto tra i sistemi portatili e la polisonnografia convenzionale. Il monitoraggio simultaneo con il sistema portatile e la polisonnografia (in laboratorio, sotto sorveglianza) permette di controllare una serie di condizioni che non sono controllabili negli studi non in simultanea, e pertanto permette di valutare che cosa ci si potrebbe aspettare durante il monitoraggio non sorvegliato al domicilio del paziente. Il confronto tra gli studi effettuati a domicilio senza sorveglianza e la polisonnografia effettuata in laboratorio potrebbe non evidenziare le differenze a favore di un setting rispetto ad un altro (casa vs laboratorio).

Le conclusioni riguardanti l'utilizzo dei sistemi di monitoraggio portatili sono per lo più applicabili alle popolazioni dei pazienti e i metodi per segnalare gli eventi rappresentano lo scopo di questo report. Come spiegato in dettaglio nei paragrafi successivi la prevalenza dell'apnea del sonno era elevata, pari a circa il 55%. I pazienti erano soprattutto uomini e, in generale, erano selezionati per gli studi da clinici esperti. I metodi per l'analisi degli eventi respiratori variavano da uno studio ad un altro.

Alcuni studi, per determinare le anomalie dell'AHI esaminavano i vari valori soglia di RDI. I valori soglia ottimali sono stati determinati usando un'analisi *post-hoc* e pertanto non sono necessariamente riproducibili. L'LR calcolato dai dati di sensi-

bilità e specificità in molti studi aveva degli intervalli di confidenza ampi, indicando una mancanza di precisione in questo tipo di calcolo. Una meta-analisi, usando le curve di ROC, avrebbe permesso di ottenere intervalli di confidenza più stretti. Tuttavia l'ERC ha tenuto in considerazione il fatto che le similitudini tra gli studi non erano abbastanza per questo tipo di approccio.

Il fatto di utilizzare il tempo trascorso a letto durante i monitoraggi con dispositivi portatili ha fatto sì che in generale i valori di AHI ottenuti con la polisonnografia fossero leggermente più alti rispetto all'RDI ottenuto con i monitoraggi di tipo 3, mentre i dati sul monitoraggio di tipo 4 erano inconsistenti (Tabella 10; dati disponibili on line, <http://www.chestjournal.org/cgi/content/full/124/4/1543/DCI>). L'effetto di un AHI più alto in generale è quello di ridurre la sensibilità ed aumentare la specificità del test.

#### 4.1.1. Evidenze che i sistemi portatili possono essere usati per ridurre la probabilità che un paziente abbia un AHI anormale

Se il monitoraggio con sistema portatile deve essere usato per escludere la diagnosi di apnea del sonno è necessario che il test abbia una alta sensibilità/basso LR per un risultato negativo. La percentuale di pazienti per i quali si avrà un risultato negativo è determinata da (1) la probabilità o la prevalenza della malattia (una caratteristica della popolazione studiata) e (2) dalla sensibilità/specificità o LR negativo (caratteristiche operative) del sistema che si sta usando. Gli studi pubblicati fino ad oggi hanno incluso popolazioni afferenti alle cliniche del sonno tra le quali c'è un'alta prevalenza della patologia. Il risultato è che, anche avvalendosi di sistemi portatili con buone caratteristiche operative, una piccola percentuale di pazienti sarà classificata come negativa con una più alta probabilità di diagnosi errata (falsi negativi). Dettagli sugli studi considerati in questi paragrafi possono essere trovati nelle Tabelle 3 e 6 (<http://www.chestjournal.org/cgi/content/full/124/4/1543/DCI>).

#### 4.1.1.1. Monitoraggio di tipo 2

Tre studi riportavano dati sulla sensibilità e specificità.<sup>20,25,32</sup>

##### 4.1.1.1.1. Laboratorio del sonno-sorvegliato

Tre studi venivano effettuati presso un laboratorio del sonno, uno dei quali con livello di evidenza II, ma tutti includevano un basso numero di pazienti.<sup>25</sup> L'LR calcolato per un risultato negativo non era molto basso (0,22) e gli intervalli di confidenza erano molto ampi.

##### 4.1.1.1.2. Domicilio-non sorvegliato

Uno studio è stato effettuato a domicilio<sup>20</sup>, ma il suo livello di evidenza era basso. Lo studio non riportava un LR molto basso per un risultato negativo (0,19), e

tra i pazienti con risultato negativo in una buona percentuale di casi (15%) si trattava di un falso negativo.

#### 4.1.1.2. Monitoraggio di tipo 3

##### 4.1.1.2.1. Laboratorio del sonno-sorvegliato

In 9 studi effettuati presso un laboratorio del sonno il monitoraggio con il sistema portatile veniva effettuato simultaneamente alla polisonnografia. Otto di questi studi<sup>18,21-23,27-29,31</sup> erano di livello I o II, ed in tutti gli studi veniva riportato un LR molto basso o ragionevolmente basso (da 0 a 0,15). Due studi<sup>21,29</sup> riportavano una sensibilità del 100%, con un LR pari a zero, ma entrambi gli studi includevano un piccolo numero di pazienti. Tutti e 8 gli studi riportavano risultati negativi, con i sistemi portatili, in buona parte dei pazienti (da 20 a 73%), ma solo una piccola percentuale di questi erano falsi negativi (da 4 a 8%). In 7 di questi studi i sistemi di monitoraggio portatili si avvalevano di un termistore per identificare gli eventi. Quattro degli studi consideravano la desaturazione di ossigeno un criterio necessario per la definizione delle ipoapnee.

##### 4.1.1.2.2. Domicilio-non sorvegliato

Quattro studi riportavano dati sul monitoraggio domiciliare e 2 di questi<sup>26,29</sup> erano di livello II. Entrambi questi studi, insieme ad un altro di livello IV<sup>24</sup>, riportavano un valore di LR basso in un piccolo numero di pazienti. In 2 di questi studi<sup>24,26</sup> la prevalenza di apnea del sonno era molto alta, cosicché il numero di pazienti con risultato negativo era basso (9% e 18% rispettivamente). Lo studio con la più alta percentuale di risultati negativi (32%) riportava anche una sostanziale percentuale di risultati falsi negativi (17% di quelli con risultato negativo).<sup>29</sup> Tutti e tre questi studi si avvalevano di un termistore per misurare il flusso, e due di questi studi consideravano la desaturazione di ossigeno come criterio necessario per la definizione delle ipoapnee.

La maggior parte delle evidenze sul monitoraggio non sorvegliato indica che con il monitoraggio di tipo 3, che si avvale di un termistore per misurare il flusso, si ha una ridotta probabilità di individuare la presenza di apnea del sonno in una importante percentuale di pazienti. Questo approccio non è stato ben validato nel setting domiciliare.

#### 4.1.1.3. Monitoraggio di tipo 4

I metodi usati per analizzare i monitoraggi di tipo 4 erano vari. I vari sistemi misuravano uno, due o tre variabili e le riportavano individualmente o in combinazione, in base al diverso tipo di sistema (Tabella 1). Le singole variabili misurate includevano l'ossimetria, la frequenza cardiaca, il flusso aereo, e in un caso, la pressione esofagea. I metodi di analisi si basavano sul criterio della desaturazione che oscillava dal 2 al 5%, sugli indici di desaturazione (indice delta), sul tempo passato al di sotto di un livello di saturazione (generalmente 90%) e sull'ispezione visiva della traccia ossimetrica.

Con i sistemi portatili l'analisi computerizzata,

senza revisione manuale, veniva usata in 16 studi, mentre l'analisi manuale in 7 studi (Tabella 1). In due studi l'approccio usato era la combinazione dell'analisi computerizzata, senza revisione, per una data variabile e l'analisi manuale per un'altra variabile. Uno studio ha dimostrato una marcata e significativa perdita di sensibilità di un ossimetro a una frequenza di campionamento di 12-s rispetto ad una frequenza di 2-s.<sup>38</sup> Tre studi riportavano una frequenza di campionamento pari a 12-s, quattro riportavano una frequenza di campionamento compresa tra 2 e 12 s, tre studi una frequenza  $\leq 2$  s mentre dieci studi non riportavano la frequenza di campionamento (Tabella 1). Le frequenze di campionamento nel resto degli studi non erano valutabili perché, per determinare l'ARDI o il suo equivalente per i sistemi portatili, venivano adoperati segnali quali il flusso aereo. La conseguenza di usare una bassa frequenza di campionamento è potenzialmente quella di ridurre la sensibilità ed aumentare la specificità del test. In tutti gli studi citati si utilizza il tempo passato a letto come riferimento per calcolare l'indice di apnea relativo ai monitoraggi con sistemi portatili rispetto al tempo di sonno, utilizzato per la polisonnografia standard. Ciò, come discusso prima, tende a ridurre i valori degli indici ottenuti con i sistemi portatili rispetto a quelli ottenuti con la polisonnografia. In due studi,<sup>33,44</sup> nei quali l'analisi computerizzata veniva paragonata all'analisi manuale, quest'ultima risultava superiore.

##### 4.1.1.3.1. Laboratorio del sonno-sorvegliato

Sedici studi sul monitoraggio di tipo 4 erano di livello I<sup>36,45</sup> o di livello II.<sup>11-13,33-35,37,43,44,48,52,54,55,57</sup> Altri nove studi revisionati erano di livello IV.<sup>14-17,46,49-51,58</sup> Nove dei sedici studi<sup>34-37,45,48,52,54,55</sup> si avvalevano della saturazione di ossigeno come metodo primario per identificare gli eventi respiratori (Tabella 1 e 3), tre studi utilizzavano il flusso misurato con la cannula nasale,<sup>11-13</sup> due studi associavano la saturazione di ossigeno al russamento o alla frequenza cardiaca,<sup>43,44</sup> uno studio utilizzava il flusso misurato con il termistore,<sup>33</sup> mentre uno studio usava il russamento in combinazione con la saturazione di ossigeno.<sup>57</sup> Tredici di 16 studi di livello I o II riportavano un LR  $\leq 0,15$  (Tabella 3). Uno studio di livello I riportava una sensibilità più bassa, pari all'82% (e un LR più alto per un risultato negativo), ma il criterio di desaturazione di ossigeno, per classificare un evento ( $> 4\%$ ), era più rigido rispetto ad altri studi.<sup>45</sup> La maggior parte degli studi riportava una percentuale elevata di pazienti con un risultato negativo ed una percentuale molto bassa di falsi negativi (Tabella 3). Due studi facevano eccezione: uno<sup>12</sup> perché riportava una prevalenza molto alta (93,7%) e l'altro<sup>45</sup> perché riportava una sensibilità più bassa. Gli studi non indicavano grosse differenze nella performance diagnostica tra i sistemi che usavano metodi diversi per identificare gli eventi. Gli studi classificati con livello di evidenza IV riportavano anche un LR abbastanza basso e, anche in que-

sto caso, non si osservavano grosse differenze tra i vari sistemi di analisi.

#### 4.1.1.3.2. Domicilio-non sorvegliato

Otto studi riportavano dati di specificità/sensibilità relativi ai monitoraggi di tipo 4 domiciliari. Quattro di questi studi<sup>33,40,41,53</sup> erano di livello I o II (Tabella 3). Tutti e otto gli studi si basavano sul monitoraggio della saturazione di ossigeno (7 studi usavano questo come criterio primario mentre uno studio usava questo come criterio secondario in associazione alla misurazione del flusso con termistore). Entrambi gli studi di livello I riportavano sensibilità più alte rispetto agli studi di livello II. Due degli studi di livello IV riportavano valori di LR piuttosto bassi. Nella maggior parte di questi studi la prevalenza era compresa tra il 40% ed il 70%. Tra i pazienti con test negativo (effettuato con sistema portatile) la percentuale di falsi negativi era più bassa nei due studi di livello I (3% e 10%) rispetto ai due studi di livello II (18% e 34%) o di livello IV (da 12% al 37%).

Come per il monitoraggio di tipo 3, la maggior parte delle evidenze sul monitoraggio di tipo 4 non sorvegliato suggerisce che, in una buona percentuale di casi, questi possono essere usati per ridurre la probabilità che un paziente abbia l'apnea del sonno. Alcuni studi effettuati a domicilio producevano risultati più variabili, tuttavia nei due studi con un livello di evidenza più elevato i risultati erano comparabili a quelli ottenuti nel setting non sorvegliato. Non sono state evidenziate differenze sostanziali nei risultati ottenuti con i sistemi di monitoraggio di tipo 4 che si avvalevano della saturazione di ossigeno e quelli che si avvalevano del flusso aereo misurato con la cannula per identificare gli eventi. Tuttavia, i sistemi con la cannula nasale non sono stati testati nel setting domiciliare non sorvegliato.

#### 4.1.2. Evidenze che i sistemi portatili possono essere usati per aumentare la probabilità che un paziente abbia un AHI anormale

Se i sistemi di monitoraggio portatili vengono usati per escludere la diagnosi di apnea del sonno, è necessario che il test abbia una elevata specificità/alto LR per un risultato positivo. Similmente a quanto avviene per ridurre la probabilità di apnea del sonno (paragrafo 4.1.1), la percentuale di pazienti con un risultato positivo è determinata da: (1) la probabilità o la prevalenza della malattia (caratteristica della popolazione studiata) e (2) la sensibilità/specificità o LR positivo (caratteristiche operative) del sistema che si sta usando. Visto che nelle cliniche del sonno esiste un'alta prevalenza della malattia, è probabile che una elevata percentuale di pazienti sarà classificata positiva a meno che la soglia per stabilire la positività non sia molto alta. Inoltre, a causa della elevata prevalenza della malattia, nei pazienti con risultati positivi c'è un rischio più basso di una falsa positività. I dettagli sugli studi considerati in questi paragrafi possono essere trovati in Tabella 4 e 7.

### 4.1.2.1. Monitoraggio di tipo 2

#### 4.1.2.1.1. Laboratorio del sonno-sorvegliato

Tre studi (livello II-a,<sup>25</sup> livello IV-b,<sup>20</sup> e livello IV-d<sup>32</sup>) paragonavano la registrazione effettuata con il sistema portatile con la polisonnografia effettuata simultaneamente. Nei due studi<sup>25,32</sup> che riportavano i dati relativi alla sensibilità e specificità (60 pazienti in totale), la specificità era elevata (90% e 93%, rispettivamente), così come lo era l'LR per un risultato positivo (8 e 14,3 rispettivamente). In uno di questi studi i pazienti venivano monitorati in un laboratorio del sonno, ma non erano sorvegliati da un tecnico e ciò aveva determinato, in alcuni di essi, la perdita dei segnali di flusso e di saturazione per un certo periodo di tempo.<sup>25</sup> In questo studio si otteneva un risultato positivo in circa il 50% dei pazienti, con un range di falsi positivi da 4% a 11%.

#### 4.1.2.1.2. Domicilio-non sorvegliato

Uno studio è stato effettuato a domicilio,<sup>20</sup> ma era di livello di evidenza basso. Lo studio riportava un LR molto alto per un risultato positivo (40,5) e, tra pazienti con risultato positivo (39%), i falsi positivi rappresentavano una piccolissima parte (3%).

### 4.1.2.2. Monitoraggio di tipo 3

#### 4.1.2.2.1. Laboratorio del sonno-sorvegliato

Nove studi, otto<sup>18,21-23,27-29,31</sup> di livello I e II e uno<sup>60</sup> di livello IV, paragonavano il monitoraggio con sistema portatile in un setting sorvegliato (laboratorio del sonno) con la polisonnografia registrata simultaneamente (Tabella 4). Questi studi riportavano valori di specificità elevati, > 90% nella maggior parte dei casi. In questi studi anche la sensibilità era elevata e di conseguenza il valore di LR per un risultato positivo era molto alto in tutti gli studi tranne in uno.<sup>29</sup> Il valore soglia prescelto per definire un AHI anormale, in generale, aveva poco effetto sulla specificità, con l'eccezione di uno studio nel quale la specificità aumentava sostanzialmente con l'aumentare della soglia di AHI.<sup>29</sup> I valori altamente positivi di LR indicano che uno studio effettuato con un sistema portatile con monitoraggio di tipo 3 sorvegliato può aumentare sostanzialmente la probabilità di un risultato positivo. Nella maggior parte dei casi questi studi utilizzavano definizioni simili per la classificazione degli eventi. L'eccezione consisteva nel fatto che in tre studi polisunnografici venivano usati gli arousal per definire gli eventi respiratori, in due studi, che si avvalevano di sistemi portatili, utilizzavano un periodo di tempo ridotto, in uno studio, che si avvaleva di un sistema portatile, veniva effettuata l'analisi computerizzata senza la revisione manuale, mentre il tempo speso a letto veniva utilizzato come riferimento per i sistemi portatili in tutti gli studi tranne in uno. In uno studio,<sup>21</sup> il tempo speso a letto veniva usato sia per la polisonnografia che per la registrazione effettuata con l'apparecchio portatile. In un altro studio<sup>29</sup> il sistema

portatile si avvaleva di una misura indiretta del sonno. In tre studi,<sup>21,28,31</sup> nei quali veniva confrontata l'analisi manuale con quella computerizzata, l'analisi manuale risultava superiore. La valutazione del tempo speso a letto, per calcolare l'indice di apnea con i sistemi portatili, può ridurre la sensibilità del test ed aumentarne la specificità in maniera artificiosa.

#### 4.1.2.2.2. Domicilio-non sorvegliato

Sono stati pubblicati quattro studi<sup>24,26,29,30</sup> (due studi di livello II e due studi di livello IV). Le specificità oscillavano tra il 58% e il 93% e, di conseguenza, i valori di LR per un risultato positivo erano modesti nei due studi di qualità più alta (5,1 e 9)<sup>26,29</sup> rispetto ai valori nei due studi di qualità più bassa (1,8 e 2,9).<sup>24,30</sup> La percentuale di pazienti con risultato positivo oscillava dal 46% all'82%, con un range corrispondente di falsi positivi che oscillava dal 2% al 31%.

Gli autori discutevano i possibili problemi dello studio. In due studi, nei quali la posizione corporea durante il sonno veniva misurata domiciliarmene, si menzionava il fatto che molti paziente con risultati falsi positivi passavano più tempo dormendo in posizione supina durante il monitoraggio domiciliare rispetto al monitoraggio polisonnografico in laboratorio. Se questi falsi positivi fossero stati dei positivi veri, dai dati può essere calcolato che la specificità sarebbe aumentata dal 66% al 100% in uno studio<sup>24</sup> e dal 70,4% al 79,2% (con un AHI < 10) e dall'82,9% all'87,2% (con un AHI < 20) in un altro studio.<sup>29</sup> I dati persi non erano molti, tranne per uno studio<sup>30</sup> nel quale dal 13% al 18% delle registrazioni non potevano essere interpretate.

Gli studi sul monitoraggio di tipo 3 effettuato in un setting sorvegliato riportavano alti valori di specificità e di LR, che potevano aumentare sostanzialmente la probabilità che un paziente avesse un AHI normale. Un numero limitato di studi sul monitoraggio domiciliare non sorvegliato ha riportato un range più ampio di specificità (e sensibilità) con valori di LR che, generalmente, non miglioravano la probabilità di apnea del sonno. Tuttavia è possibile che alcuni pazienti considerati falsi positivi siano stati classificati male a causa delle differenze, non controllabili, tra il setting domiciliare e il laboratorio, in relazione ad alcuni fattori come per esempio il tempo passato in posizione supina. La conseguenza di ciò è che i valori di specificità riportati dagli studi non sorvegliati potrebbero essere sottostimati e che l'utilità di questi monitoraggi nell'aumentare la probabilità di apnea del sonno potrebbe essere superiore a quella riportata. Tutti gli studi si avvalevano della misurazione del flusso aereo con un termistore per identificare gli eventi. Due dei quattro studi<sup>26,29</sup> includevano la desaturazione di ossigeno come criterio necessario per la definizione delle ipoapnee, mentre altri due studi non utilizzavano questo criterio<sup>24,30</sup> (Tabella 1).

### 4.1.2.3. Monitoraggio di tipo 4

#### 4.1.2.3.1. Laboratorio del sonno-sorvegliato

Venticinque studi si basavano sui monitoraggi di tipo 4 (2 studi livello I; 14 studi livello II; 9 studi livello IV) (Tabella 4 e 7). Tredici studi usavano la sola ossimetria,<sup>34-37,45,46,48-52,54,55</sup> 8 studi la misurazione del flusso (pressione nasale, 6 studi;<sup>11-14,16,17</sup> pressione nasale e desaturazione di ossigeno, 1 studio;<sup>15</sup> termistore con desaturazione di ossigeno, 1 studio<sup>33</sup>), 1 studio usava il russamento in associazione all'ossimetria,<sup>57</sup> 2 studi usavano l'ossimetria più la frequenza cardiaca ed il russamento,<sup>43,44</sup> mentre uno studio usava la frequenza cardiaca da sola.<sup>58</sup> In uno studio i valori di sensibilità non venivano riportati in modo interpretabile,<sup>59</sup> mentre altri 3 studi non riportavano affatto la sensibilità e la specificità.<sup>10,38,47</sup> Dei 16 studi di livello I o II, in 11 veniva riportato un valore di LR per un risultato positivo > 5, in 9 un LR > 10 e in 3 un LR > 20 (Tabella 4). Il tipo di segnale usato per identificare gli eventi non influenzava i risultati relativi alla specificità e sensibilità. In un certo numero di studi non veniva riportata la quantità di dati persi, mentre, in quelli in cui veniva riportata, la percentuale di dati persi oscillava tra lo 0% e l'11%. Nella maggior parte degli studi sui sistemi portatili la percentuale di pazienti con test positivo oscillava tra il 30% e 60%. Negli studi che riportavano valori di LR ≥ 10 la percentuale di risultati falsi positivi era bassa (range da 0% a 12%), mentre negli studi che riportavano valori di LR più bassi per un risultato positivo le percentuali di falsi positivi erano molte più alte (range da 3% a 37%).

#### 4.1.2.3.2. Domicilio-non sorvegliato

Otto studi riportavano i dati relativi alla sensibilità/specificità dei sistemi di monitoraggio portatili di tipo 4 effettuati nel setting domiciliare; quattro di questi erano di livello I o II,<sup>33,40,41,53</sup> mentre il resto era di livello IV. Tutti questi studi si basavano sul monitoraggio della saturazione di ossigeno (7 studi come criterio primario, mentre uno studio la utilizzava come criterio secondario in combinazione con la misurazione del flusso con il termistore). Tre dei 4 studi di livello di evidenza I/II<sup>33,39,41</sup> (saturazione di ossigeno) riportavano valori di LR > 10, mentre uno studio<sup>53</sup> riportava valori di LR molto bassi per un risultato positivo. Due degli studi di livello IV riportavano anche elevate specificità/LR. La prevalenza, nella maggior parte di questi studi, era compresa tra 40% e 70%. La percentuale di pazienti il cui test, effettuato col sistema portatile, era positivo, con un LR alto, oscillava tra il 14% ed il 42%, mentre il range di falsi positivi, negli studi con LR elevato, oscillava tra lo 0% ed il 12%. Il range di risultati positivi e quello di falsi positivi nei tre studi che riportavano un basso valore di LR andava da 47% a 73% e da 10% a 41%, rispettivamente.

Per determinare e classificare gli eventi venivano utilizzati approcci diversi, inclusi l'indice di desaturazione, le desaturazioni più i cambiamenti della

frequenza cardiaca, l'indice delta, il tempo cumulativo passato al di sotto di una determinata saturazione e la valutazione visiva del tracciato. L'analisi sui monitoraggi portatili veniva effettuata secondo approcci diversi. In due studi l'analisi era completamente computerizzata senza revisione manuale, in tre studi era manuale, in due studi era combinata e in uno studio il metodo d'analisi non era riportato (Tabella 1). In generale non c'erano dati che dimostravano che l'analisi computerizzata era superiore a quella manuale, anche se uno studio (usando il termistore più la desaturazione di ossigeno) dimostrava la superiorità dell'analisi manuale.<sup>33</sup>

#### *4.1.3. Evidenza che un singolo monitoraggio con sistema portatile può essere usato per aumentare e ridurre la probabilità che un paziente abbia un AHI anormale*

Se un sistema di monitoraggio portatile viene usato per ridurre (se il risultato è negativo) e aumentare (se il risultato è positivo) la probabilità che un paziente abbia l'apnea del sonno, il test deve avere alta sensibilità/basso LR per un risultato negativo e alta specificità/alto LR per un risultato positivo. Se il test permette di ottenere ciò con un singolo valore soglia, allora si otterranno sia risultati positivi che negativi, e se il test ha delle ottime caratteristiche operative ci saranno pochi falsi positivi o falsi negativi. Tuttavia, se si ottiene un basso LR ad un dato valore soglia ed un alto LR ad un altro valore soglia, allora per alcuni pazienti non sarà possibile ottenere una diagnosi con il sistema di monitoraggio portatile. I dati relativi ai valori bassi o alti di LR sono riassunti nelle Tabelle 5 e 8. In Tabella 5 è inclusa anche la percentuale di pazienti che non potevano essere classificati attraverso il monitoraggio portatile. Questo numero era pari a zero se veniva usato un singolo valore soglia per definire la positività e la negatività del test. La percentuale di falsi positivi e falsi negativi viene riportata anche in Tabella 5.

Ventotto gruppi di studi, estratti da un totale di 27 articoli, presentavano dati che indicavano che il monitoraggio con sistema portatile poteva ridurre (basso LR di un test negativo) e aumentare (alto LR di un test positivo) la probabilità di apnea del sonno nella popolazione studiata.

##### **4.1.3.1. Monitoraggio di Tipo 2**

Solo due studi sul monitoraggio di tipo 2 (entrambi di livello IV) riportavano valori di LR sia alti che bassi.

###### **4.1.3.1.1. Laboratorio del sonno-sorvegliato**

Un singolo studio<sup>32</sup> (livello di evidenza IV/livello di qualità d), che esaminava un piccolo numero di pazienti (40 pazienti), riportava la sensibilità ed la specificità allo stesso valore soglia, ciò produceva una bassa percentuale di falsi negativi o falsi positivi (mostrata in Tabella 5).

###### **4.1.3.1.2. Domicilio-non sorvegliato**

Uno studio<sup>20</sup> (livello di evidenza IV/livello di qualità b) ha evidenziato una sensibilità modesta ma una specificità molto alta, con un 10% di pazienti nei quali la diagnosi era errata. In questo studio la quantità di dati persi era elevata (20%)

##### **4.1.3.2. Monitoraggio di Tipo 3**

###### **4.1.3.2.1. Laboratorio del sonno-sorvegliato**

I nove studi che usavano il monitoraggio di tipo 3 nel laboratorio del sonno riportavano tutti dati che indicavano un alto positivo e basso negativo valore di LR. Come descritto in precedenza 8 di questi studi<sup>18,21-23,27-29,31</sup> erano di livello I o II. Solo in uno studio<sup>29</sup> venivano riportati pazienti che non venivano classificati come positivi o negativi, e nella maggior parte degli studi la percentuale di errore (risultato falso) era di circa il 5%.

###### **4.1.3.2.2. Domicilio-non sorvegliato**

Due di 4 studi domiciliari (entrambi di livello II), che usavano il monitoraggio di tipo 3,<sup>26,29</sup> riportavano alti valori di LR positivo e bassi valori di LR negativo. In entrambi gli studi per classificare gli eventi ci si avvaleva della misurazione del flusso con termistore e dei criteri di desaturazione di ossigeno. Entrambi gli studi adottavano valori soglia diversi per produrre un basso o alto LR, con il risultato che, nel 22% dei pazienti in uno studio<sup>29</sup> e nel 37% dei pazienti in un altro,<sup>26</sup> il risultato era intermedio e non poteva essere posta diagnosi. Le percentuali di errore erano pari a 16% e 5%, rispettivamente, per i 2 studi.

Come detto in precedenza, la maggior parte degli studi sui monitoraggi non sorvegliati indica che i monitoraggi di tipo 3 che si avvalgono della misurazione del flusso possono sostanzialmente ridurre ed aumentare la probabilità di apnea del sonno nella maggior parte dei pazienti. Il numero limitato di dati sul monitoraggio domiciliare indica che ulteriori studi saranno necessari per valutare questo tipo di approccio.

##### **4.1.3.3. Monitoraggio di Tipo 4**

###### **4.1.3.3.1. Laboratorio del sonno-sorvegliato**

Dei 25 studi pubblicati sul monitoraggio di tipo IV, presso il laboratorio del sonno, che riportavano i valori di sensibilità e specificità, 15 (livello I, 2 studi; livello II, 7 studi; livello IV, 6 studi) riportavano valori alti e bassi di LR. In dieci di questi studi, per ottenere livelli alti o bassi di LR, si adottavano valori soglia diversi (Tabella 5). Alcuni di questi<sup>13,34,36</sup> riportavano una percentuale bassa di pazienti non classificati, ma in molti di questi studi<sup>12,33,35,40,43,49,50,51,57</sup> questa percentuale era invece elevata. La percentuale di errore era bassa (< 7%). Dei 5 studi che adottavano un singolo valore soglia, quattro<sup>13,14,17,34</sup> riportavano una percentuale di risultati falsi < 7%.

Non c'erano dati chiari ad indicare che i diversi metodi di scoring degli eventi potessero migliorare il valore di LR e influenzare la percentuale di pazienti

non classificati o la percentuale di errori. Studi di livello I, II o IV riportavano percentuali variabili di pazienti non classificati o mal classificati, senza che ci fossero chiare differenze tra i diversi metodi utilizzati per l'identificazione degli eventi.

#### 4.1.3.3.2. Domicilio-non sorvegliato

Un singolo studio<sup>40</sup> (che usava la saturazione di ossigeno) sul monitoraggio domiciliare non sorvegliato riportava dati che permettevano di classificare i pazienti come positivi o negativi. Questo studio riportava una percentuale elevata di pazienti non classificati (50%), ma una bassa percentuale di errore (4%).

In generale, la maggior parte degli studi (19 su 27) che riportavano bassi ed alti valori di LR, erano considerati di buona qualità (livello di evidenza I/II, con la maggior parte dei criteri soddisfatti). Il tipo di monitoraggio o il canale primario registrato non influenzavano i risultati. Nella maggior parte degli studi, ma non in tutti, la prevalenza dell'apnea del sonno era elevata (> 40%). Il range di severità dell'apnea del sonno era ampio in tutti gli studi. In molti studi si eseguiva l'analisi manuale, ma in molti altri ciò non era possibile, perché il sistema portatile non lo permetteva (Tabella 1).

## 4.2. Aspetti secondari

### 4.2.1. Riproducibilità del monitoraggio con sistema portatile

La riproducibilità dei monitoraggi con sistema portatile può essere considerata da diversi punti di vista. In primo luogo bisogna considerare la riproducibilità dello scoring manuale degli eventi. Quando l'analisi viene effettuata manualmente, dovrebbero essere fornite delle indicazioni sulla probabilità che ha un operatore di ottenere risultati simili analizzando lo stesso test in occasioni diverse, così come sulla probabilità che due operatori possano ottenere risultati simili quando analizzano lo stesso test. Dei 51 studi revisionati, 19 riportavano un sistema di analisi esclusivamente manuale, mentre 10 riportavano un'analisi automatica revisionata manualmente o un'analisi combinata manuale ed automatica (Tabella 1). Nessuno di questi 29 studi riportava dati sulla riproducibilità dell'analisi.

Un'altra importante fonte di variabilità, nei risultati di un monitoraggio, è la variabilità che si osserva da notte a notte. Questa variabilità può essere random o dovuta all'effetto prima notte, come descritto precedentemente per la polisonnografia.<sup>81</sup> Solo 2 studi riportavano la variabilità da notte a notte. In entrambi questi studi il monitoraggio era domiciliare non sorvegliato, entrambi gli studi erano di livello di evidenza II e di livello di qualità "a" ed entrambi utilizzavano il flusso misurato con termistore e le desaturazioni di ossigeno come criteri per classificare gli eventi. Il primo studio<sup>18</sup> riportava dati ottenuti su 32 pazienti studiati domiciliariamente per 2 notti e dimostrava che non c'erano differenze nel

livello di desaturazione o nella durata degli eventi. Il coefficiente di correlazione tra il valore di RDI ottenuto durante la prima notte e quello ottenuto durante la seconda notte era pari a 0,94. L'analisi di Brand-Altmann non era riportata. I valori medi simili di RDI ottenuti nelle due notti, suggerivano la mancanza di un effetto prima-notte. Usando un valore soglia di RDI > 10 per definire un test anormale, gli autori riportavano che solo un paziente veniva classificato patologico nella seconda notte di studio, dopo che era stato classificato normale dopo la prima notte. Il secondo studio riportava invece delle differenze nei risultati del monitoraggio tra la prima notte, in laboratorio, e la seconda notte, a domicilio. Per 12 di 53 pazienti la differenza nei valori di RDI, tra le due notti, era  $\geq 10$ .

Un terzo criterio, per valutare la riproducibilità dei monitoraggi, consiste nel verificare i risultati di diversi gruppi di investigatori che utilizzano protocolli di studi simili. Anche se molti studi utilizzano sistemi di monitoraggio simili, le differenze nei protocolli impediscono di concludere che un singolo monitoraggio è stato ben valutato per la riproducibilità in diversi centri del sonno.

Al momento, esistono pochi dati sulla riproducibilità dei monitoraggi con sistema portatile. Questa rappresenta pertanto una importante area per futuri studi.

### 4.2.2. Analisi costo-beneficio dei sistemi di monitoraggio portatili

Solo un piccolo numero di studi paragonava i costi relativi al monitoraggio con sistema portatile e alla polisonnografia tradizionale. Poiché l'obiettivo di questi articoli era di effettuare un paragone con la polisonnografia in termini diagnostici, non venivano applicati metodi per effettuare l'analisi del rapporto costo-beneficio. Gli articoli facevano cenno, in modo molto variabile, ai benefici o ai costi relativi al monitoraggio con sistemi portatili o valutavano il potenziale risparmio economico di evitare la polisonnografia. Nell'effettuare questa ultima analisi, la maggior parte degli investigatori dava per scontato la necessità di effettuare il monitoraggio per due notti (uno diagnostico ed uno terapeutico) e non considerava il costo ridotto di una polisonnografia split-night.

Quando eseguita, l'analisi dei costi era limitata ai costi diretti, relativi al monitoraggio con sistema portatile o con polisonnografia. Non venivano invece considerati altri aspetti quali la difficoltà di accesso ai laboratori del sonno, i costi indiretti derivanti dalla mancata diagnosi e trattamento dell'apnea del sonno (es. incidenti automobilistici, incidenti sul lavoro, assenza dal lavoro); queste evenienze potevano verificarsi sia mancando la diagnosi con un risultato falso negativo ottenuto col sistema portatile sia per l'impossibilità del paziente di accedere alla polisonnografia convenzionale.

In molti degli studi si fanno delle deduzioni sui risparmi che si otterrebbero se i sistemi di monito-

raggio portatili rivelassero un buona sensibilità e specificità. Il paragrafo successivo prenderà in considerazione quegli studi che forniscono dati in termini di dollari, tempo ed esami addizionali risparmiati.

#### 4.2.2.1. Monitoraggio di Tipo 2

Dei 4 studi sul monitoraggio di tipo 2 solo 2 (di livello di evidenza IV e di livello di qualità "a" o "d") prendevano in considerazione i costi. In uno studio<sup>19</sup> si accennava a un progetto di riduzione dei costi basato sull'assunzione che le spese per il personale notturno e per il laboratorio del sonno avrebbero potuto essere ridotte utilizzando un sistema portatile, ma non venivano forniti dati numerici. Un secondo studio<sup>20</sup> indicava che il monitoraggio con sistema portatile costava la metà rispetto alla polisonnografia in laboratorio (pazienti che preferivano la polisonnografia in laboratorio, 48%; pazienti che preferivano il monitoraggio domiciliare, 28%; pazienti che non avevano preferenza, 24%) ma nello studio non venivano considerati i costi derivanti dal 20% di fallimenti dei sistemi portatili.

#### 4.2.2.2. Monitoraggio di Tipo 3

Sono stati revisionati 12 studi sul monitoraggio portatile di tipo 3, tuttavia, 6 di questi studi non facevano cenno ai costi, mentre 3 di questi<sup>24,29,60</sup> presumevano semplicemente che il monitoraggio con sistema portatile avrebbe determinato una riduzione dei costi per il minor tempo impiegato dal tecnico (set up e/o analisi). Uno studio basato sul monitoraggio in laboratorio<sup>23</sup> (livello di evidenza II e livello di qualità "a"), condotto in 104 pazienti, paragonava il tempo necessario per il set-up e per l'analisi dei dati con il sistema portatile e con la polisonnografia standard e riportava che il tempo per il set-up era pari a 20-25 minuti vs 60-75 minuti, rispettivamente, mentre il tempo per l'analisi dei dati era di 45-90 minuti vs 180 minuti, rispettivamente.

Due studi sul monitoraggio domiciliare riportavano i costi. Uno studio spagnolo<sup>26</sup> (89 pazienti, livello di evidenza II e livello di qualità "a") calcolava: (1) i costi diretti relativi al monitoraggio con sistemi portatili (che tenevano in considerazione l'uso singolo dell'apparecchiatura, il set-up a domicilio con e senza tecnico, e la percentuale di fallimenti che richiedevano la ripetizione del monitoraggio) e (2) i costi indiretti che includevano le spese di viaggio e riparazione. L'efficacia del monitoraggio con sistema portatile era pari al 92% (percentuale di falsi negativi, con un valore soglia di AHI di 10, pari a 8%) con il set-up domiciliare effettuato dal tecnico e pari all'80% con il set-up domiciliare effettuato dallo stesso paziente (ulteriore 12% di fallimenti). I costi riportati per il monitoraggio erano pari a circa un terzo dei costi della polisonnografia, indifferentemente dal fatto che il set-up fosse effettuato dal tecnico o meno.<sup>26</sup> Il secondo studio,<sup>30</sup> effettuato nel

Regno Unito (livello di evidenza IV, livello di qualità "b"; dati interpretabili in 123 su 150 pazienti) riportava che il costo del monitoraggio con sistema portatile era pari al 21,9% del costo della polisonnografia in laboratorio. Il tempo di attesa medio per il paziente per un monitoraggio con sistema portatile era di 18 giorni contro i 47 giorni di attesa per una polisonnografia standard ( $p < 0,001$ ). Quando si considerava l'intera popolazione di pazienti, inclusi quelli per i quali si sarebbe potuta evitare la polisonnografia e quelli che invece dovevano effettuare entrambi i test, il risparmio era di solo 3,17 sterline a paziente.

#### 4.2.2.3. Monitoraggio di Tipo 4

In questo caso l'analisi viene complicata dalla variabilità relativa ai vari sistemi utilizzati ed ai parametri monitorati, così come dalla scelta, in ciascun articolo, di focalizzare i costi o i benefici. Di 35 studi sul monitoraggio di tipo 4, 8 discutevano il problema dei costi-benefici. Cinque di questi erano studi sul monitoraggio portatile sorvegliato, 2 erano sul monitoraggio domiciliare non sorvegliato mentre in uno studio la metodologia era mista.

In uno studio condotto nel Regno Unito,<sup>56</sup> che si avvaleva dell'ossimetria domiciliare non sorvegliata (livello di evidenza IV livello di qualità "c"), si usava il criterio fornito dalla British Thoracic Society di  $\geq 15$  desaturazioni ( $> 4\%$ ) per ora spesa a letto. Poiché si otteneva una sensibilità bassa (31%) ed una specificità elevata (100%), gli autori concludevano che solo i pazienti sintomatici con una ossimetria domiciliare negativa dovevano essere sottoposti alla polisonnografia. Lo screening con l'ossimetria di 69 pazienti (costo stimato 10 sterline) avrebbe evitato l'esecuzione di 10 polisonnografie (costo stimato 500 sterline), con un risparmio complessivo di 62 sterline a paziente. L'altro studio canadese basato sull'ossimetria domiciliare non sorvegliata (livello di evidenza I e livello di qualità "d") riportava un'alta sensibilità (98%) ma una scarsa specificità (48%) e concludeva che 62 pazienti con ossimetria negativa (26%), su un totale di 240 pazienti, avrebbero potuto evitare la polisonnografia. Falsi negativi sono stati riportati in 2 dei 62 pazienti, e in 18 casi lo studio era stato ripetuto perché il paziente aveva dormito poco.

Due dei 6 studi che si avvalevano dell'ossimetria in un ambiente sorvegliato discutevano la possibilità di evitare la polisonnografia. Nello studio americano<sup>35</sup> (livello di evidenza II, e livello di qualità "d") si utilizzava un questionario predittivo insieme all'ossimetria e si concludeva che questo algoritmo combinato, in una popolazione ad alto rischio, riduceva il numero delle polisonnografie dell'8%-12% nei pazienti con ossimetria negativa, usando un valore soglia di RDI pari a 30. Gli autori affermavano anche che l'algoritmo rappresentava un mezzo per gli specialisti del sonno per identificare i pazienti ad

alto rischio che dovevano essere sottoposti subito a monitoraggio. L'altro articolo spagnolo<sup>45</sup> (livello di evidenza I, livello di qualità "b"), che si avvaleva dell'ossimetria in simultanea con la polisonnografia in laboratorio, concludeva che, i pazienti sintomatici con ossimetria negativa, dovrebbero essere sottoposti a polisonnografia, così come i pazienti con ossimetria positiva affetti da alterazione note della funzione respiratoria (presumibilmente affetti da patologie più complesse ma i dettagli non erano specificati). Lo studio suggeriva che, usando un livello di desaturazione del 4% e dei valori soglia di RDI di 5, 10, 15, se l'ossimetria veniva effettuata come esame iniziale, la polisonnografia avrebbe potuto essere evitata in 140 (51%), 119 (43%) e 105 (38%), rispettivamente, dei 275 pazienti analizzati.

Due studi sul monitoraggio sorvegliato, che si avvalevano della cannula nasale per misurare il flusso, riportavano i costi. Uno studio condotto nel Regno Unito<sup>14</sup> su 31 pazienti (livello di evidenza IV, livello di qualità "b"), includendo il costo di deprezzamento della strumentazione per 200 studi per letto per anno, considerava i costi pari a £14 per il monitoraggio portatile vs £126 per la polisonnografia. Considerando che i pazienti con AHI  $\geq 20$  dovevano poi essere sottoposti a ventilazione a pressione continua positiva (CPAP), il costo per paziente era di £35 per il monitoraggio portatile e di £154 per la polisonnografia. Un altro studio francese<sup>12</sup> (livello di evidenza II e livello di qualità "a"), che usava lo stesso sistema di monitoraggio portatile, non valutava direttamente i costi relativi al monitoraggio portatile, ma il risparmio derivante dalla conseguente riduzione nel numero di polisonnografie. Quando un algoritmo clinico, che valutava la probabilità pre-test di un paziente di avere un AHI  $\geq 15$ , indicava una probabilità  $> 80\%$  (46 pazienti), solo 5 pazienti, per i quali il monitoraggio portatile dava esito negativo, avrebbero richiesto l'esecuzione di una polisonnografia e ciò avrebbe consentito il risparmio di 41 studi del sonno. Dell'intera popolazione di 95 pazienti, solo 25 pazienti, per i quali il monitoraggio portatile dava esito negativo, avrebbero richiesto una polisonnografia, con un risparmio totale di 70 polisonnografie.

Altri due studi usavano sistemi di monitoraggio che associavano all'ossimetria canali addizionali quali la frequenza cardiaca ed il russamento in uno studio (livello di evidenza II e livello di qualità "a") e il flusso aereo misurato con il termistore in un altro. Il primo di questi studi,<sup>43</sup> svolto in Spagna, valutava 150 pazienti presso un laboratorio del sonno sorvegliato. L'algoritmo in questo studio includeva il monitoraggio portatile in tutti i pazienti (\$150 (in dollari americani)) e poi successivamente la polisonnografia (\$546) per coloro i quali i risultati erano dubbi (25%) e per coloro che richiedevano la ripetizione del test (10%). Gli autori concludevano che con questo approccio, nella loro popolazione di

pazienti con sospetta apnea del sonno, ci sarebbe stata una riduzione netta dei costi del 44%. L'ultimo studio<sup>33</sup> canadese, (livello di evidenza II, livello di qualità "a") considerava un algoritmo nel quale la polisonnografia era richiesta solo se il monitoraggio portatile aveva rivelato un RDI compreso tra 2 e 20. Secondo questi autori, il tempo dedicato dal tecnico alla polisonnografia era di 4 ore. Questi autori concludevano che il monitoraggio iniziale con sistema portatile (con una perdita dell'8% dei dati) eviterebbe 42 polisonnografie in 48 pazienti, con una riduzione netta totale del tempo di due ore per paziente (intervallo di confidenza 95%, da 1,6 a 2,5 ore).

#### 4.2.2.4. Conclusioni

L'uso dei sistemi portatili come strumento iniziale per la diagnosi di apnea del sonno, in pazienti selezionati, può ridurre i costi, riducendo l'utilizzo delle risorse sanitarie e permettendo ai pazienti di essere sottoposti direttamente allo studio per la titolazione della CPAP, se il monitoraggio dà esito positivo o di essere sottoposti ad ulteriori esami se l'esito è negativo. Poiché i risultati di questi studi non sono completamente generalizzabili, è opportuno considerarli con cautela, considerando che le conclusioni sono strettamente dipendenti dalla probabilità pre-test e dal livello di valore soglia scelto per la diagnosi di apnea del sonno. È chiaro che ulteriori studi dovranno prendere in considerazione l'analisi del costo/beneficio, paragonando il monitoraggio con i sistemi portatili con la polisonnografia split-night e valutando l'impatto sugli outcome importanti per il paziente.

#### 4.2.3. Percentuale di fallimenti relativi al monitoraggio con sistema portatile

##### 4.2.3.1. Monitoraggio di tipo 2

La quantità di dati persi veniva riportata in 3 dei 4 studi. Uno studio<sup>19</sup> effettuato a domicilio in 77 pazienti, non riportava la perdita di alcun monitoraggio, ma riportava la perdita dei dati relativi alla saturazione di ossigeno nel 4% dei pazienti. Un altro studio<sup>20</sup> domiciliare non sorvegliato riportava la perdita del 20% dei dati.

##### 4.2.3.2. Monitoraggio di tipo 3

Sei studi riportavano chiaramente la percentuale di monitoraggi portatili per i quali i dati raccolti non erano sufficienti. Questa percentuale oscillava dal 3%<sup>24,29</sup> al 18%<sup>30</sup> negli studi effettuati a domicilio senza sorveglianza. Nei due studi effettuati in laboratorio la percentuale dei dati persi era pari al 3,3% dei pazienti in uno studio<sup>21</sup> e al 9% nell'altro.<sup>18</sup>

##### 4.2.3.3. Monitoraggi di tipo 4

Quattordici studi riportavano la percentuale di monitoraggi portatili per i quali i dati raccolti non erano sufficienti. Uno studio<sup>33</sup> riportava che, dei

monitoraggi eseguiti a domicilio ed in laboratorio, che usavano il termistore per misurare il flusso e la saturazione di ossigeno, si perdevano circa l'8% dei dati. Altri 3 studi,<sup>41,42,53</sup> a domicilio senza sorveglianza, riportavano una perdita di dati dal 7% al 10%. Tutti questi dati si basavano sulla saturazione dell'ossigeno per identificare gli eventi. Nei 10 studi effettuati in laboratorio sotto sorveglianza la perdita di dati variava dal 2%<sup>36</sup> al 14%,<sup>48</sup> per i sistemi che usavano la saturazione di ossigeno, e dal 10% al 14% per i 4 studi che usavano la cannula nasale per misurare il flusso.<sup>10,13,15,17</sup> Uno studio basato sulla variazione della frequenza cardiaca riportava la perdita del segnale dal 5% all'11% dei casi.<sup>52,58</sup>

#### 4.2.4. Popolazioni di pazienti studiati e possibilità di generalizzare i dati

##### 4.2.4.1. Severità della patologia

###### 4.2.4.1.1. Monitoraggio di Tipo 2

Tre dei 4 studi sul monitoraggio di tipo 2<sup>20,25,51</sup> riportavano dati sulla distribuzione dei valori di AHI nella popolazione di pazienti studiati. Questi studi includevano pazienti con un ampio range di valori di AHI (AHI medio > 20), tra i quali pazienti che non avevano l'apnea del sonno (AHI < 5) e pazienti con apnea di grado severo (AHI > 40).

###### 4.2.4.1.2. Monitoraggio di tipo 3

Tutti e dodici gli studi sul monitoraggio di tipo 3 includevano pazienti che non avevano l'apnea del sonno (AHI < 5). Undici dei 12 studi includevano pazienti con apnea del sonno di grado severo e, mentre il range di valori di AHI non era riportato nel 12 studio,<sup>22</sup> 3 studi<sup>23,27,31</sup> riportavano un AHI medio < 20. Uno studio<sup>22</sup> non riportava la distribuzione dell'AHI nei pazienti studiati.

###### 4.2.4.1.3. Monitoraggio di tipo 4

Trentadue dei 35 studi sul monitoraggio di tipo 4 includevano un ampio range di pazienti, dai soggetti sani ai pazienti affetti da apnea di grado severo. Di questi 32 studi 6 riportavano un AHI medio più basso (< 20)<sup>13,16,17,33,47,57</sup> rispetto agli altri. Il range dei valori di AHI non veniva riportato in 3 studi.<sup>40,48,54</sup>

###### 4.2.4.1.4. Conclusioni

Le caratteristiche operative di un sistema portatile dovrebbero essere testate in pazienti con e senza apnea del sonno e in un campione di pazienti con un range ampio di valori di AHI. La maggior parte degli studi considerati includeva pazienti con un range ampio di valori di AHI; ciò è indicato sia dalla presenza di pazienti con valori elevati di AHI sia da un valore medio di AHI > 20. Ciò non è sorprendente, considerato che la maggior parte degli studi includeva pazienti inviati alle cliniche o ai laboratori del sonno per la diagnosi polisonnografica di apnea del sonno (vedi paragrafo 4.2.4.4). Poiché nella mag-

gior parte degli studi l'AHI medio indicava un range di severità da moderato a severo (AHI > 20), i risultati di questi studi potrebbero non essere applicabili a popolazioni con una probabilità pre-test più bassa (es. popolazione generale, donne, pazienti afferenti agli ambulatori di medicina generale).

##### 4.2.4.2. Presenza di patologie concomitanti

###### 4.2.4.2.1. Monitoraggio di tipo 2

Solo uno studio<sup>20</sup> riportava la presenza di comorbidità [es. bronchite cronica (criteri non specificati)]. Uno studio<sup>25</sup> riportava altri disturbi del sonno, la cui diagnosi veniva fatta dopo la conclusione di tutti gli esami.

In uno studio si specificava che i pazienti inclusi venivano inviati per la presenza di "disturbi del sonno" (non solo apnea del sonno), ma le diagnosi non venivano poi specificate. Negli altri due studi non si faceva cenno ad altri disordini del sonno.

###### 4.2.4.2.2. Monitoraggio di tipo 3

Dei 12 studi che si avvalevano del monitoraggio di tipo 3, in 9 non si accennava alla presenza di comorbidità, incluse le patologie respiratorie. Uno studio<sup>18</sup> includeva i pazienti con patologie polmonari ostruttive e restrittive. In due studi<sup>31,60</sup> venivano esclusi i pazienti con patologie polmonari "importanti", anche se non venivano specificati i criteri di esclusione. In uno studio<sup>31</sup> si escludevano i pazienti con scompenso cardiaco congestizio in classe NYHA superiore a 2.

Tre studi<sup>24,28,30</sup> specificavano che i pazienti per i quali c'era il sospetto di un disturbo del sonno diverso dall'apnea venivano esclusi. In uno studio<sup>23</sup> venivano inclusi i pazienti con disturbi del sonno diversi dall'apnea e veniva fornita la diagnosi finale. Tutti gli altri studi includevano solo pazienti con sospetta apnea del sonno.

###### 4.2.4.2.3. Monitoraggio di tipo 4

Dei 35 studi che si avvalevano del monitoraggio di tipo 4, 23 non facevano cenno alle comorbidità, incluse le patologie respiratorie. Tre studi specificavano che i pazienti con BPCO,<sup>41,43</sup> o con BPCO e patologie restrittive, venivano inclusi nello studio. Due studi<sup>46,50</sup> si avvalevano dei pazienti con patologie respiratorie come gruppo di controllo. In molti studi si escludevano i pazienti con patologie respiratorie significative, definite da un  $FEV_1 < 50%$  del predetto,<sup>36</sup> dall'ipossiemia diurna (saturazione di ossigeno < 90%<sup>40,56</sup> o < 7.3 kPa<sup>36</sup>) o dalla pratica di ossigenoterapia.<sup>35,49</sup> Uno studio<sup>13</sup> escludeva i pazienti con BPCO e ipossemia diurna, ma non specificava i valori di cut-off. Uno studio<sup>58</sup> escludeva i pazienti con scompenso cardiaco congestizio o diabete mellito insulino-dipendente, ma non quelli con patologie respiratorie.

Solo 3 studi<sup>16,40,54</sup> accennavano ad altre diagnosi relative al sonno, fatte dopo la conclusione degli

esami. Nessuno studio menzionava in modo specifico l'esclusione dei pazienti con disturbi del sonno diversi dall'apnea.

#### 4.2.4.2.4. Conclusioni

In molti pazienti che si presentano al centro/laboratorio del sonno per il sospetto di apnea del sonno sono presenti comorbidità quali asma, BPCO, e sindrome da ipoventilazione da obesità. La BPCO e la sindrome da ipoventilazione sono associate a desaturazioni notturne. La presenza di queste patologie potrebbe pertanto influenzare la performance dei sistemi di monitoraggio portatili, particolarmente di quelli che utilizzano l'ossimetria come parametro principale. Lo scompenso cardiaco congestizio è anche una patologia frequente nei pazienti con sospetta apnea del sonno e potrebbe influenzare i risultati del monitoraggio se il paziente durante il sonno presenta respiro di Cheyne-Stokes o apnee centrali. Nella maggior parte degli studi inclusi in questa revisione non si specificava se i soggetti inclusi erano affetti da comorbidità importanti (patologie respiratorie e cardiache) che potevano influenzare la performance dei sistemi portatili e gli outcome. Pertanto, al momento, i risultati di questa revisione dovrebbero essere applicati solo a pazienti privi di importanti comorbidità respiratorie o cardiache.

Inoltre nei centri del sonno vengono valutati altri disturbi del sonno quali la narcolessia e la sindrome dei movimenti periodici delle gambe. L'inclusione di pazienti con questi disturbi in uno studio sui sistemi di monitoraggio portatile, riduce la prevalenza dell'apnea del sonno nella popolazione studiata. La maggior parte degli studi includeva solo soggetti per i quali c'era il sospetto clinico di apnea del sonno e non forniva diagnosi alternative per quei soggetti che non avevano l'apnea del sonno. Pertanto i risultati di questa revisione possono essere applicati solo ai pazienti con sospetto clinico di apnea del sonno e non necessariamente ai pazienti con altre possibili patologie del sonno inviati ad un centro specializzato.

#### 4.2.4.3. Sesso/Razza

##### 4.2.4.3.1. Monitoraggio di tipo 2

Tre studi riportavano il rapporto uomo/donna nella loro popolazione di pazienti con una percentuale di uomini > 70% in due di essi. Nessuno studio riportava la presenza di soggetti di razza non caucasica.

##### 4.2.4.3.2. Monitoraggio di tipo 3

Undici di 12 studi riportavano il rapporto uomo/donna nella loro popolazione di pazienti, con una percentuale di uomini superiore al 70% in 10 di questi studi e pari al 56% nell'altro studio.<sup>27</sup> Nessuno studio riportava la presenza di soggetti di razza non caucasica.

##### 4.2.4.3.3. Monitoraggio di Tipo 4

Dei 35 studi sul monitoraggio di tipo 4, 27 riporta-

vano il numero di uomini e donne studiati. In 24 di questi studi gli uomini rappresentavano > 70% dei soggetti studiati. Uno studio<sup>35</sup> riportava anche le razze incluse nella popolazione studiata (bianchi 72%; neri, 24%; altri 4%).

#### 4.2.4.3.4. Conclusioni

È stato dimostrato che la prevalenza di apnea del sonno<sup>2</sup> è minore tra le donne, nelle quali la patologia si manifesta principalmente con ipoapnee.<sup>32</sup> Uno studio<sup>33</sup> ha dimostrato che la prevalenza di apnea del sonno negli afroamericani è simile a quella dei bianchi, ma che in questi l'AHI è più elevato, dopo la correzione per l'indice di massa corporea e l'età. La maggior parte degli studi in questa review includeva principalmente soggetti di sesso maschile, come è tipico delle popolazioni inviate ai centri del sonno. Pertanto applicare questi risultati ad altre popolazioni, nelle quali esiste una percentuale più alta di soggetti di sesso femminile, potrebbe essere inappropriato. Inoltre, poiché solo uno studio indicava l'inclusione di soggetti di razza non caucasica, non è chiaro se i risultati possono essere applicati a pazienti di razza non caucasica.

Non ci sono dati in letteratura che suggeriscono che la frequenza degli eventi respiratori durante il sonno è influenzata dal sesso o dalla razza. Sono necessari ulteriori studi per stabilire l'utilità dei sistemi portatili per il monitoraggio di popolazioni di razza non caucasica e di quelle popolazioni in cui è presente una percentuale più alta di donne.

#### 4.2.4.4. Popolazioni di pazienti reclutati

##### 4.2.4.4.1. Monitoraggio di tipo 2

Solo in uno studio<sup>19</sup> si dichiarava che il processo di reclutamento era iniziato in una clinica del sonno. Tuttavia il monitoraggio portatile veniva effettuato solo nei pazienti per i quali si pensava fosse necessaria la polisomnografia. Gli altri 2 studi erano effettuati su pazienti inviati al laboratorio del sonno e non specificavano secondo quale criterio erano stati inviati.

##### 4.2.4.4.2. Monitoraggio di tipo 3

Dei 12 studi sul monitoraggio portatile di tipo 3 solo uno studio<sup>27</sup> era stato eseguito in un sottogruppo di pazienti inclusi in una popolazione per lo studio dell'apnea del sonno. In 6 studi<sup>21-23,25,30,31</sup>, i pazienti venivano reclutati da una clinica del sonno. Tuttavia, in 3 di questi studi,<sup>23,28,30</sup> il monitoraggio con il sistema portatile era effettuato solo in quei pazienti per i quali la polisomnografia era ritenuta necessaria. I 6 studi rimanenti erano condotti in un laboratorio del sonno, ma i criteri di inclusione non venivano specificati.

##### 4.2.4.4.3. Monitoraggio di tipo 4

Dei 35 studi sul monitoraggio portatile di tipo 4, 22 includevano solo i pazienti con sospetta apnea

del sonno inviati al laboratorio del sonno per la polisonnografia. Nessuno studio specificava le ragioni per le quali questi pazienti erano stati inviati al laboratorio del sonno. Tuttavia 3 studi<sup>39,45,49</sup> riportavano che i pazienti erano inviati da specialisti vari, inclusi medici di medicina generale, ospedalieri, pneumologi e otorinolaringoiatri.

Tredici studi<sup>34-36,40-42,45,51-53,56-58</sup> dichiaravano, in modo specifico, che il reclutamento dei pazienti era cominciato nella clinica del sonno. Tuttavia, solo in sette studi<sup>36,38,40,45,49,53,56</sup> si dichiarava che i pazienti reclutati erano consecutivi, o rappresentavano un campione random di tutti i pazienti inviati alla clinica per il sospetto di apnea del sonno. In questi 7 studi si effettuavano i test diagnostici su tutti i pazienti inviati, e non solo sui pazienti in cui c'era il sospetto di apnea del sonno dopo una valutazione clinica.

#### 4.2.4.4.4. Conclusioni

La popolazione nella quale si valuta un test diagnostico è fondamentale per l'interpretazione delle caratteristiche operative del test quali sensibilità, specificità ed LR. Per l'apnea del sonno la probabilità pre-test è bassa nella popolazione generale, più alta in una clinica del sonno e più alta ancora in un laboratorio del sonno. In questa revisione la maggior parte degli studi (30 di 51) includeva pazienti inviati al laboratorio del sonno per il sospetto di apnea del sonno. I criteri esatti adoperati dai clinici per avviare i pazienti alla polisonnografia non erano specificati. Il fatto di avvalersi di una popolazione di pazienti pre-selezionata da certi specialisti per studiare i sistemi di monitoraggio portatili rende difficile l'applicazione dei risultati a pazienti che non sono stati valutati dalla stessa categoria di specialisti. Poiché esiste un solo studio<sup>27</sup> effettuato in una popolazione generale, i dati disponibili non sono sufficienti a stabilire se i risultati degli altri 50 studi sono applicabili a popolazioni diverse da quelle afferenti alle cliniche/laboratori del sonno.

### 4.3. Conclusioni

#### 4.3.1. Valutazione generale e limiti dello studio

Nelle revisioni sistematiche su argomenti importanti, quali l'uso dei sistemi portatili per la diagnosi dell'apnea del sonno, l'AHRQ suggerisce di basare le raccomandazioni sulla quantità, qualità e consistenza delle evidenze disponibili.<sup>66</sup> Il gruppo di studio dell'ATS/ACCP/AASM, che ha revisionato gli studi, ha ritenuto impossibile effettuare una meta-analisi di questi studi con curve ROC, a causa della marcata eterogeneità dei metodi, delle definizioni e dei tipi di monitoraggio. In totale sono stati revisionati 51 studi. Il numero di studi per un singolo sistema di monitoraggio variava da 1 a 8. Tra gli studi che si avvalevano dell'ossimetria c'era una certa variabilità, perché in alcuni studi l'ossimetria veniva usata come segnale unico, mentre in altri

veniva abbinata ad un altro segnale (es. frequenza cardiaca o russamento). Inoltre ciascuno studio usava un metodo diverso per l'analisi del segnale. Pertanto non è giustificabile una raccomandazione su un qualsiasi singolo monitoraggio, perché la quantità e la consistenza dei risultati, al momento, non possono essere valutate adeguatamente.

La qualità degli studi sui sistemi portatili per la diagnosi dell'apnea del sonno era variabile, anche se la maggior parte di questi era di elevato livello di evidenza e qualità (livello di evidenza I o II e livello di qualità "a o b"). Solo alcuni degli studi erano di livello I, perché la maggior parte degli autori non valutava popolazioni di pazienti consecutivi. La maggior parte degli studi includeva pazienti che erano stati inviati ad un laboratorio del sonno e ciò aumentava la probabilità di errore di selezione. La maggior parte degli studi di livello II soddisfaceva la maggior parte degli altri criteri di qualità (livello di qualità "a o b"). Le cause più comuni del basso livello di qualità erano la mancata descrizione dei sistemi usati, la mancata definizione degli eventi, la mancata esecuzione in cieco dell'analisi e l'impossibilità di effettuare la revisione manuale dell'analisi automatica.

Diversi limiti degli studi pubblicati fanno sì che i risultati non siano ampiamente applicabili, così come segue:

1. Tutti gli studi venivano effettuati in centri/laboratori del sonno nei quali la probabilità pre-test (prevalenza) di apnea del sonno era elevata. In questo setting, il numero di falsi negativi aumenta e si riduce il valore predittivo negativo del test.
2. Gli studi includevano principalmente pazienti di sesso maschile, senza comorbilità importanti, e pertanto l'applicabilità dei risultati alle donne e ai pazienti con scompenso cardiaco congestizio o BPCO è sconosciuta.
3. Il numero di pazienti negli studi revisionati variava ampiamente (da 20 a 359) e ciò influenza direttamente l'accuratezza dei risultati e la loro riproducibilità.
4. La maggior parte degli studi paragonava i monitoraggi portatili con la polisonnografia in un setting sorvegliato, cosicché l'utilità dei sistemi portatili in un setting domiciliare non è bene stabilita.
5. Avvalersi del tempo trascorso a letto come riferimento per i monitoraggi di tipo 3 e 4 può produrre valori di RDI più bassi rispetto ai valori di AHI ottenuti con la polisonnografia.
6. Avvalersi di due valori di soglia di RDI (uno basso per escludere la patologia ed uno alto per confermarne la presenza) fa sì che un certo numero di pazienti non possa essere classificato.
7. La frequenza di campionamento dell'ossimetro potrebbe non essere sufficiente ad individuare tutte le desaturazioni.
8. I protocolli sui monitoraggi portatili non sorvegliati non permettevano di aggiustare i risultati

del confronto della polisonnografia in base alla variabilità da notte a notte dovuta alla diversa posizione del corpo (supina vs laterale) e alle fasi del sonno (REM vs non REM).

9. I risultati di uno qualsiasi degli studi considerati sono completamente dipendenti, e pertanto limitati, dalla tecnologia disponibile nel momento in cui il protocollo è stato effettuato. Da quando gli studi sono stati pubblicati, molti sistemi di monitoraggio sono stati sottoposti a modifiche che potrebbero influenzare i risultati di studi futuri.
10. I sistemi di monitoraggio di tipo 3 e 4 non identificano i risvegli. Poiché questi sono importanti per la definizione degli eventi, ciò potrebbe essere visto come un limite di questi tipi di monitoraggio.

#### 4.3.2. Questioni primarie e tipi di monitoraggio

In questo paragrafo vengono riassunti gli studi in base ai tipi di monitoraggio e vengono discusse tre questioni principali. Poiché per la maggior parte dei monitoraggi sono stati pubblicati solo uno o due studi (che erano quelli con i protocolli più variabili), i commenti sono limitati agli studi con elevato livello di evidenza e di qualità (livello di evidenza I, II e livello di qualità "a o b") che rispondevano in modo specifico alle questioni poste, rispetto al numero totale di studi di alta qualità (consistenza). Saranno discussi anche gli studi con un basso numero di falsi negativi o falsi positivi ( $\leq 10\%$ ).

##### 4.3.2.1. Monitoraggio di tipo 2

I monitoraggi di tipo 2 permettono di stadiare il sonno e dovrebbero essere quelli che concordano meglio con la polisonnografia, per la possibilità di calcolare un AHI. Tuttavia c'erano solo 4 studi<sup>19,20,25,32</sup> pubblicati sul monitoraggio di tipo 2, ed in uno di questi studi<sup>19</sup> mancavano i dati relativi alla sensibilità e specificità. Due dei tre studi rimanenti avevano un basso livello di evidenza (livello IV). Per la piccola quantità e bassa qualità di questi studi, non possono essere tratte conclusioni sulla loro consistenza e pertanto essi non saranno discussi nei paragrafi successivi.

**Sinopsi:** Mancano evidenze convincenti a indicare che i monitoraggi di tipo 2 possono essere utilizzati in un setting sorvegliato o non sorvegliato.

##### 4.3.2.2. Evidenza che i sistemi portatili possono essere usati per ridurre la probabilità che un paziente abbia un AHI anormale (LR < 0,2)

###### 4.3.2.2.1. Monitoraggio di tipo 3

Dodici studi hanno effettuato 13 confronti tra il monitoraggio portatile e la polisonnografia. In tutti gli studi si usava il termistore per definire le apnee e le ipoapnee nel polisonnogramma (la maggior parte includeva anche un criterio di desaturazione), e per definire gli eventi respiratori relativi al monitoraggio

con sistemi portatili (in uno studio<sup>60</sup> non veniva specificato come erano definiti gli eventi) (Tabella 1). Otto dei 9 studi sul monitoraggio sorvegliato erano di alto livello di evidenza e qualità (3 studi erano di livello di evidenza I e di livello di qualità "a") e tutti riportavano un basso LR ( $< 0,2$ ) (tabella 3). Sette degli 8 studi<sup>21-23,27-29,31</sup> riportavano una percentuale bassa di falsi negativi. Due dei 4 studi in setting non sorvegliato<sup>26,29</sup> erano di livello di evidenza e livello di qualità più alti (entrambi di livello di evidenza II e livello di qualità "a o b"). Entrambi gli studi riportavano bassi valori di LR ed una percentuale relativamente alta di falsi negativi.

**Sinopsi:** Diversi studi di elevato livello di evidenza e qualità dimostrano, in modo consistente, che alcuni sistemi di monitoraggio sono utili a ridurre la probabilità che un paziente abbia l'apnea del sonno in un setting sorvegliato. I risultati relativi al setting non sorvegliato devono essere considerati preliminari e suggeriscono che questi sistemi potrebbero essere utili, ma la loro reale utilità dovrà essere confermata da ulteriori studi.

###### 4.3.2.2.2. Monitoraggio di tipo 4

Sette confronti (6 studi<sup>33,34,39,43,44,57</sup>) combinavano l'ossimetria con almeno un altro parametro, come la frequenza cardiaca o il russamento, per lo scoring degli eventi respiratori. Cinque studi<sup>33,34,43,44,57</sup> venivano effettuati in un setting sorvegliato; tutti erano di livello di evidenza II e di livello di qualità "a" e riportavano un LR basso. Uno studio<sup>57</sup> riportava un'alta percentuale di falsi negativi. Uno<sup>33</sup> dei due studi in setting non sorvegliato<sup>33,39</sup> era di livello di evidenza più alto (II-a), ma entrambi riportavano percentuali relativamente alte di falsi negativi.

Diciannove confronti (18 studi) valutavano come riferimento la sola ossimetria. Tredici di questi erano studi in setting sorvegliato, con 6 di più alto livello di evidenza e qualità<sup>36,37,45,52,54,55</sup> e 4 che riportavano un basso LR ed una percentuale accettabile di falsi negativi.<sup>36,37,52,55</sup>

Dei 6 studi domiciliari non sorvegliati solo uno era di alto livello di evidenza e qualità.<sup>53</sup> Questo studio riportava un basso LR ed una bassa percentuale di falsi negativi.

In 7 studi, tutti in setting sorvegliato, veniva utilizzata la cannula nasale. Di questi, solo 3 studi<sup>11-13</sup> erano di elevato livello di evidenza e qualità. Tutti e tre gli studi riportavano un basso LR e due di questi<sup>11,13</sup> una bassa percentuale di falsi negativi.

**Sinopsi:** L'ossimetria da sola può ridurre la probabilità che un paziente abbia l'apnea del sonno in un setting sorvegliato e non sorvegliato. Tuttavia, nell'ultimo caso, i risultati relativi al setting non sorvegliato devono essere considerati preliminari. L'aggiunta di un secondo segnale produceva risultati simili a quelli ottenuti con la sola ossimetria (anche

se gli studi valutati erano di meno), e pertanto valgono le stesse conclusioni. La cannula nasale può essere utile in un setting sorvegliato, ma non possono essere tratte conclusioni sul suo utilizzo in un setting non sorvegliato.

#### **4.3.2.3. Evidenza che i monitoraggi portatili possono essere usati per aumentare la probabilità che un paziente abbia un AHI anormale (LR > 5)**

##### 4.3.2.3.1. Monitoraggi di tipo 3

Otto di 9 studi in un setting non sorvegliato erano di alta qualità e riportavano un alto LR. Tuttavia i valori di LR andavano da 6,4 all'infinito (Tabella 4). Cinque di 8 studi<sup>18,21,22,28,31</sup> riportavano una bassa percentuale (< 10%) di falsi positivi. Due dei 4 studi<sup>26,29</sup> effettuati in setting non sorvegliato erano di qualità più elevata, entrambi riportavano valori elevati LR (da 5 a 9) ed uno studio<sup>29</sup> riportava un'alta percentuale di falsi positivi.

**Sinopsi:** *Diversi studi dimostrano che il monitoraggio di tipo 3 può aumentare la probabilità che un paziente abbia l'apnea del sonno in un setting sorvegliato. I dati sul setting non sorvegliato sono tuttavia limitati.*

##### 4.3.2.3.2. Monitoraggio di tipo 4

Dei 6 confronti fatti in 5 studi<sup>33,34,43,44,57</sup> sui monitoraggi di tipo 4 che usavano l'ossimetria combinata ad un secondo parametro e che erano di alto livello di evidenza e di qualità, 5 confronti (4 studi<sup>33,34,43,57</sup>) riportavano un LR alto (da 10,3 a 45). Due studi<sup>33,44</sup> riportavano una percentuale elevata (> 10%) di falsi positivi.

Degli 8 studi<sup>35-37,45,48,52,54,55</sup> con più alto livello di evidenza e qualità, che usavano la sola ossimetria in un setting sorvegliato, 5<sup>35,36,45,48,54</sup> riportavano alti valori di LR (da 8,2<sup>36</sup> a 18<sup>35</sup>). Uno dei 5 studi<sup>36</sup> riportava un'alta percentuale di falsi positivi. Dei 3 studi<sup>40,51,53</sup> con livello di evidenza e di qualità più alto, effettuati in un setting non sorvegliato, due<sup>40,41</sup> riportavano valori alti di LR con percentuali accettabili di falsi positivi.

Solo due<sup>12,13</sup> dei tre studi<sup>11-13</sup> che usavano la cannula nasale e che erano di livello di evidenza e qualità più alto, riportavano un alto LR, ma solo uno<sup>12</sup> riportava una bassa percentuale di falsi positivi nel setting sorvegliato. Non c'erano studi nel setting non sorvegliato.

**Sinopsi:** *L'ossimetria da sola può aumentare la probabilità che il paziente abbia l'apnea del sonno in un setting sorvegliato e non sorvegliato. Tuttavia, nel setting non sorvegliato l'utilità sembra inferiore rispetto al setting sorvegliato. Quando veniva aggiunto un secondo segnale i risultati erano simili a quelli ottenuti con la sola ossimetria. Tuttavia, manca l'evidenza che questo tipo di combinazione di segnali possa*

*essere usata in un setting non sorvegliato. I risultati relativi all'uso della cannula nasale nel setting sorvegliato dovrebbero essere considerati preliminari, ma suggeriscono che l'utilità di questo approccio è ancora dubbia. Non possono essere tratte conclusioni sul suo utilizzo in un setting non sorvegliato.*

#### **4.3.2.4. Evidenza che un singolo monitoraggio con sistema portatile può essere usato per ridurre ed aumentare la probabilità che un paziente abbia un AHI anormale (LR negativo, ≤ 0,2; LR positivo, > 5)**

Ventotto confronti da 27 studi, hanno riportato dati che indicano che il monitoraggio con sistema portatile può sia ridurre (basso LR per un test negativo) che aumentare (alto LR per un test positivo) la probabilità che la popolazione studiata sia affetta da apnea del sonno (Tabella 5). La maggior parte di questi studi era di livello di evidenza e qualità più alto (livello di evidenza I o II e livello di qualità "a" o "b" per 18 di 27 studi). La configurazione del sistema portatile era molto variabile, così come il numero di pazienti studiati, la severità dell'apnea del sonno e la definizione di anormalità. La prevalenza di apnea del sonno era elevata nella maggior parte degli studi, ma non in tutti. Alcuni studi si basavano sull'analisi automatica, altri sull'analisi manuale mentre altri su entrambi i metodi (Tabella 1).

La sensibilità e la specificità spesso erano calcolate per vari livelli di RDI, e ciò produceva risultati che potevano escludere o confermare la diagnosi di apnea del sonno. In 15 di questi studi le migliori sensibilità e specificità erano ottenute a uno stesso livello di RDI, mentre in 13 studi la migliore sensibilità e la migliore specificità erano a livelli di RDI diversi. In questo ultimo caso, per alcuni pazienti il risultato del test non era diagnostico (penultima colonna in Tabella 5), e sarà utilizzato come indicatore di utilità nei paragrafi successivi.

##### 4.3.2.4.1 Monitoraggio di tipo 3

Nel setting sorvegliato tutti gli 8 studi sul monitoraggio di tipo 3 che avevano un livello di evidenza e qualità più alto, riportavano valori di LR sia alti che bassi, indicando che il monitoraggio portatile poteva essere usato sia per confermare che per escludere la diagnosi di apnea del sonno. Uno studio<sup>29</sup> usava valori soglia multipli e riportava un'alta percentuale di pazienti nei quali il test non era né positivo né negativo. Due dei 4 studi in setting non sorvegliato<sup>26,29</sup> sul monitoraggio di tipo 3 con livello di evidenza e di qualità elevato, riportavano valori di LR sia alti che bassi. Questi studi usavano valori soglia diversi di RDI ed in entrambi la percentuale di pazienti per il quale il test non era né positivo né negativo era alta.

**Sinopsi:** *I monitoraggi di tipo 3 sono utili a ridurre ed aumentare la probabilità che un paziente possa*

*avere l'apnea del sonno in un setting non sorvegliato. L'utilità in un setting non sorvegliato è ancora dubbia.*

#### 4.3.2.4.2. Monitoraggio di tipo 4

Quattro studi<sup>33,34,43,57</sup> sul monitoraggio di tipo 4, che si avvalevano dell'ossimetria combinata ad almeno un altro canale, tutti eseguiti in setting sorvegliato, riportavano valori di LR sia alti che bassi. Tutti erano del livello di qualità più alto. Tre studi<sup>33,43,57</sup> usavano valori soglia diversi per ottenere un LR alto e basso e riportavano un'alta percentuale di pazienti per il quale il test effettuato con sistema portatile non era diagnostico.

Un solo studio condotto a domicilio senza sorveglianza,<sup>40</sup> che usava solo l'ossimetria, riportava valori di LR sia alti che bassi. In questo studio la percentuale di test non diagnostici era elevata. Inoltre, 3 di 9 studi<sup>34-36</sup> che usavano solo l'ossimetria, di livello di evidenza e qualità più alto, condotti in un setting non sorvegliato, riportavano valori di LR sia alti che bassi. In due di questi studi<sup>35,36</sup> la percentuale di pazienti nei quali il test non era diagnostico era elevata. Due studi di qualità più alta, che si avvalevano della cannula nasale (entrambi effettuati in un setting sorvegliato), riportavano valori di LR sia alti che bassi, e in uno di questi, la percentuale di test non diagnostici era elevata.

**Sinopsi:** *L'utilità di usare l'ossimetria da sola per aumentare e ridurre la probabilità di apnea del sonno in un setting sorvegliato o non sorvegliato, con o senza un canale addizionale è ancora dubbia. Al momento, i risultati dovrebbero essere considerati preliminari. L'utilità della cannula nasale per aumentare e diminuire la probabilità che un paziente abbia l'apnea del sonno in un setting sorvegliato o non sorvegliato, non è stata dimostrata.*

#### 4.3.3. Considerazioni conclusive

I motivi per i quali i clinici si interessano al monitoraggio con i sistemi portatili sono diversi. Alcuni li vorrebbero usare come sistema di screening per escludere la presenza della patologia, mentre altri per decidere se iniziare il trattamento. Alcuni sistemi si sono mostrati promettenti nell'escludere la patologia (Tabella 3), altri nel confermarne l'esistenza (Tabella 4) e alcuni promettenti in entrambi i casi (Tabella 5). I clinici interessati all'uso dei sistemi portatili per il monitoraggio dei pazienti con sospetta apnea del sonno dovrebbero valutare attentamente ciò che si aspettano dal monitoraggio, i segnali che vorrebbero registrare, la qualità degli studi pubblicati sull'argomento e se il setting e le popolazioni di questi studi sono simili ai loro; ciò renderebbe ragionevole l'accettazione dei dati pubblicati. Al momento, poiché non è possibile basare una raccomandazione sulla quantità e sulla consistenza dei risultati, il gruppo di studio dell'ATS/ACCP/AASM raccomanda che i cli-

nici considerino la validazione dei sistemi portatili che vogliono usare nel setting in cui essi operano. Se il sistema di monitoraggio deve entrare a far parte della pratica clinica, è auspicabile che il clinico, nel momento in cui sceglie un particolare sistema di monitoraggio, si ponga le seguenti tre domande:

1. Quale è lo scopo del monitoraggio? La revisione del gruppo di studio dell'ATS/ACCP/AASM ha evidenziato che esistono più studi di elevata qualità che dimostrano l'utilità del sistema portatile per escludere o confermare la diagnosi di apnea del sonno piuttosto che per spiegare entrambe le funzioni, particolarmente ad un dato valore soglia di RDI.
2. In che setting dovrebbe essere usato il sistema di monitoraggio? L'uso ideale dei sistemi di monitoraggio portatili sarebbe a domicilio del paziente, senza sorveglianza. I vantaggi teorici di questo setting consistono nella diminuzione dei costi e nella possibilità di studiare più pazienti contemporaneamente. Tuttavia, la maggior parte degli studi sui sistemi portatili, con livello di evidenza e qualità elevato, che riportano i valori di sensibilità e specificità, sono stati effettuati in un setting sorvegliato (livello di evidenza I o II e livello di qualità "a" o "b", 23 di 36 studi in setting sorvegliato vs 4 di 13 studi in setting non sorvegliato), e mancano i dati che dimostrano l'effettiva diminuzione dei costi con i sistemi portatili. Pertanto l'utilità dei sistemi portatili in un setting non sorvegliato non è stata ancora ben stabilita. Ci sono tuttavia potenziali vantaggi nell'utilizzo dei sistemi portatili in un setting non sorvegliato, consistenti principalmente nella riduzione nel tempo per il set up e lo scoring, che potrebbero spingere il clinico ad usare questi sistemi piuttosto che la polisonnografia convenzionale.
3. Che tipo di sistema dovrebbe essere usato? I sistemi disponibili vanno dai più completi, che effettuano la polisonnografia completa, a quelli con un limitato numero di canali (flusso, sforzo respiratorio e ossimetria), fino alla semplice ossimetria. La polisonnografia completa ha il vantaggio potenziale di misurare i parametri relativi al sonno. Il monitoraggio con sistema a numero limitato di canali è più semplice da effettuare ed ha il vantaggio di misurare il flusso e l'ossimetria, come avviene per la polisonnografia completa. L'ossimetria ha il vantaggio di essere semplice. Questa review ha trovato un elevato numero di studi di alta qualità relativi ai monitoraggi di tipo 3 e 4.

In generale i dati più consistenti e di alta qualità erano quelli relativi al monitoraggio di tipo 3 nel setting sorvegliato, nel quale esso poteva confermare o escludere la presenza di apnea del sonno. Il numero

di risultati falsi riportato da questi studi era basso e la maggior parte di questi riusciva ad identificare un unico valore soglia di RDI per distinguere i pazienti con e senza apnea del sonno.

## 5.0. DIRETTIVE PER GLI STUDI FUTURI

Questa review ha evidenziato diversi aspetti che dovrebbero essere approfonditi da studi futuri (paragrafi 1.1.3, 1.3, 1.4, 1.5, 2.3, 2.4, 3.4, e 4.2.4) e che sono stati discussi nel paragrafo 4.3.

### 5.1. Popolazioni di pazienti non ancora studiate

Al momento quasi tutti gli studi sul monitoraggio portatile sono stati effettuati su popolazioni di pazienti reclutati nelle cliniche/laboratori del sonno, con un'alta probabilità pre-test, composte principalmente da uomini di razza bianca, con poche o nes-

suna patologia concomitante; pertanto non è possibile generalizzare i risultati di questi studi.

Gli studi futuri dovrebbero ovviare a questo problema ed includere diverse popolazioni di pazienti così come segue:

1. Pazienti che afferiscono agli ambulatori di medicina di base;
2. Pazienti con importanti comorbidità;
  - BPCO
  - Scompenso cardiaco
  - Ictus
  - Ipertensione severa
3. Razze diverse da quella bianca
4. Donne

Inoltre gli studi futuri dovrebbero considerare l'utilità degli algoritmi predittivi clinici in combinazione con il monitoraggio portatile, aspetto che fino ad oggi è stato valutato da un solo studio.<sup>35</sup>

**Tabella 11—Metodi raccomandati per il monitoraggio con i sistemi portatili per la diagnosi di apnea del sonno**

Caratteristiche dello studio	Punti chiave	Aspetti importanti
Studio-generale	<ol style="list-style-type: none"> <li>1. Fonti di finanziamento</li> <li>2. Possibile conflitto d'interesse</li> </ol>	
Popolazione studiata	<ol style="list-style-type: none"> <li>1. Campione di pazienti consecutivi (prospettico), non soggetto ad errori di campionamento</li> <li>2. Ampio campione di pazienti:               <ul style="list-style-type: none"> <li>Ampio range di probabilità pre-test</li> <li>Rappresentazione adeguata di entrambi i sessi</li> <li>Descrizione dell'etnia</li> <li>Descrizione chiara delle comorbidità per permettere l'analisi separata</li> </ul> </li> </ol>	<ol style="list-style-type: none"> <li>1. Descrizione del setting dello studio</li> <li>2. Criteri di elegibilità chiaramente definiti</li> <li>3. Descrizione chiara del numero di pazienti che:               <ul style="list-style-type: none"> <li>Erano elegibili</li> <li>Si sono rifiutati di partecipare</li> <li>Venivano esclusi dopo l'inizio dello studio</li> <li>Hanno completato lo studio</li> <li>Avevano dati incompleti o mancanti</li> </ul> </li> </ol>
Polisonnografia	<ol style="list-style-type: none"> <li>1. Analisi in cieco dei risultati del monitoraggio e dei dati clinici</li> <li>2. La decisione di effettuare la polisonnografia non dipende dai risultati del monitoraggio portatile</li> <li>3. Preselezione dei criteri di positività del test (evitare l'analisi <i>post hoc</i>)</li> </ol>	<ol style="list-style-type: none"> <li>1. Seguire le misurazione e i metodi raccomandati</li> <li>2. Descrizione chiara dello scoring degli eventi respiratori</li> <li>3. Frequenza di campionamento dell'ossimetria chiaramente dichiarata</li> <li>4. Assegnazione random dei pazienti alla polisonnografia o al monitoraggio portatile per il primo test (per studi non simultanei)</li> <li>5. Prevalenza dell'apnea per qualsiasi soglia definita chiaramente riportata</li> <li>6. Chiara descrizione degli apparecchi e delle ditte produttrici</li> </ol>
Monitoraggio con sistema portatile	<ol style="list-style-type: none"> <li>1. Analisi in cieco dei risultati del monitoraggio e dei dati clinici dei pazienti</li> <li>2. La decisione di effettuare il monitoraggio con sistema portatile non dipende dai risultati della polisonnografia</li> <li>3. Preselezione dei criteri di positività del test (evitare l'analisi <i>post hoc</i>)</li> </ol>	<ol style="list-style-type: none"> <li>1. Esecuzione in contemporanea il monitoraggio portatile e la polisonnografia</li> <li>2. Descrizione chiara dello scoring degli eventi respiratori</li> <li>3. Se viene eseguita l'analisi manuale riportare la variabilità</li> <li>4. La ripetibilità da notte a notte deve essere riportata come la media delle differenze e i limiti di concordanza<sup>29</sup></li> <li>5. Devono essere riportati gli intervalli di confidenza 95% della sensibilità, specificità ed LR.</li> <li>6. Devono essere dichiarati i criteri di inclusione per un risultato positivo o negativo, il numero di test test non diagnostici la percentuale di falsi negativi</li> <li>7. Chiara descrizione degli apparecchi e delle ditte produttrici.</li> <li>8. Percentuale di fallimenti o percentuale di studi che dovevano essere ripetuti per la scarsa qualità o per problemi tecnici</li> </ol>

## 5.2. Raccomandazioni per i metodi di studio

Quattro aspetti principali dovrebbero essere considerati dagli investigatori nel paragonare il monitoraggio portatile con la polisonnografia standard (Tabella 11). Gli investigatori dovrebbero documentare ogni possibile conflitto di interessi se una delle ditte produttrici di sistemi portatili ha pagato per una loro consulenza. Il reclutamento deve includere pazienti consecutivi, da un gruppo che non è stato precedentemente selezionato dagli investigatori. Se sono presenti comorbidità, i pazienti dovrebbero essere chiaramente definiti e stratificati in gruppi separati. Il campione di pazienti dovrebbe essere sufficientemente grande da far sì che i risultati siano rappresentativi della popolazione e il power dello studio, basato su una ipotesi chiaramente dichiarata, dovrebbe essere calcolato prima di iniziare lo studio. La popolazione dello studio dovrebbe essere descritta abbastanza bene da permettere ai lettori di capire se la popolazione studiata è simile alla loro popolazione di pazienti, in modo da giustificare l'applicazione dei risultati nella pratica clinica. I dettagli sull'analisi del polisonnogramma e del monitoraggio portatile dovrebbero essere sufficienti a permettere al lettore di ripetere lo studio e ad assicurare che siano stati controllati importanti fonti di errore. In Tabella 11 sono elencati i punti importanti e fondamentali che gli editori, i revisori, e i lettori degli studi clinici sui sistemi di monitoraggio portatili dovrebbero considerare.

Al momento i sistemi di analisi della polisonnografia variano da un centro ad un altro. Le raccomandazioni dell'AASM possono essere usate come una guida.<sup>1</sup> La frequenza di campionamento dell'ossimetro dovrebbe essere sufficiente a registrare accuratamente il segnale di saturazione di ossigeno e dovrebbe essere sempre dichiarata. Per misurare il flusso oronasale dovrebbe essere usato uno dei metodi raccomandati. È stato riconosciuto che negli studi revisionati i dati relativi al flusso oronasale erano quelli che più comunemente andavano persi. La definizione di ipoapnea dovrebbe basarsi sulle raccomandazioni dell'AASM per la polisonnografia. Anche se il catetere esofageo è il metodo preferito per misurare i risvegli correlati allo sforzo respiratorio, questo è generalmente poco confortevole e non adatto ai sistemi portatili. I risvegli correlati allo sforzo respiratorio sono inclusi nella definizione di AHI e sono anche caratteristici della sindrome dell'aumentata resistenza delle vie aeree superiori.<sup>5,6</sup> La relazione tra questi approcci standardizzati e gli outcome clinici è al momento sconosciuta, così come lo è la rilevanza clinica del valore soglia dell'AHI.

Il sistema di analisi del monitoraggio portatile dovrebbe anche essere standardizzato e dichiarato chiaramente. La relazione tra il monitoraggio portatile e la polisonnografia dovrebbe essere stabilita attraverso uno studio in simultanea presso il laboratorio del sonno. Questo rappresenterà la base attra-

verso la quale giudicare i risultati ottenuti in un setting sorvegliato. I valori soglia di RDI, per il monitoraggio portatile, e di AHI, per la polisonnografia, dovrebbero essere selezionati prima dello studio come l'outcome primario variabile sul quale basare la sensibilità e la specificità. Il criterio raccomandato dall'AASM di un AHI  $\geq 5$  è appropriato per il setting clinico. Se uno studio viene effettuato per determinare il valore soglia di RDI, un secondo studio dovrebbe essere effettuato per validare l'accuratezza di questo valore soglia. Gli studi dovrebbero essere in cieco e i metodi dovrebbero essere descritti in dettaglio.

## 5.3. Raccomandazioni per i protocolli di studio

### 5.3.1. Confronto tra il monitoraggio portatile e la polisonnografia

Se si esegue un protocollo di studio nel quale i risultati del monitoraggio portatile vengono confrontati con quelli della polisonnografia, si raccomanda di considerare gli aspetti descritti nella tabella precedente. Inoltre si dovrebbero aggiustare i risultati del monitoraggio portatile non sorvegliato, paragonati ad una polisonnografia eseguita in una notte diversa, per fattori quali la variabilità da notte a notte.<sup>84,85</sup> I fattori che contribuiscono alla variabilità da notte a notte sono solo in parte conosciuti,<sup>84</sup> ma alcuni sono spesso difficili da identificare, come per esempio la posizione del corpo e le fasi del sonno, particolarmente il sonno REM. Anche se non esiste un modo assoluto per aggiustare i risultati per questi fattori, un approccio possibile sarebbe quello di stabilire la variabilità da notte a notte del polisonnogramma, del polisonnogramma sorvegliato e del monitoraggio portatile non sorvegliato in un determinato gruppo di pazienti. Questo potrebbe essere ottenuto effettuando gli studi in tre notti diverse, usando la polisonnografia in due e il monitoraggio portatile non sorvegliato, non in simultanea con il polisonnogramma, in un'altra notte. Gli intervalli di tempo tra gli studi dovrebbero essere simili. La variabilità tra il polisonnogramma della prima notte e quello della seconda notte dovrebbe essere confrontata con la variabilità tra il monitoraggio portatile non sorvegliato e ciascuno dei due polisonnogrammi, nello stesso paziente. Idealmente, dovrebbe essere condotto un secondo studio con monitoraggio portatile non sorvegliato per valutare la variabilità da notte a notte di questo tipo di monitoraggio. Se la variabilità relativa ai monitoraggi portatili non sorvegliati e alle polisonnografie non sorvegliate è simile, allora è probabile che le differenze tra il monitoraggio portatile e il polisonnogramma siano dovute alla variabilità da notte a notte. Se c'è una discrepanza, la grandezza della discrepanza rappresenta la perdita di sensibilità e/o specificità che deriva dall'uso dei sistemi portatili. Questo sistema fornirebbe inoltre informazioni sulla perdita di dati relativi al monitoraggio portatile non sorvegliato rispetto alla polisonnografia.

L'ordine in cui i pazienti vengono sottoposti a polisonnografia e a monitoraggio portatile dovrebbe essere random, per evitare un effetto dell'ordine stesso (l'effetto prima notte si osserva nei pazienti sottoposti a polisonnografia). Inoltre è fortemente raccomandato che tutti gli studi si avvalgano del sensore di posizione per calcolare i valori di AHI e RDI nella posizione supina e non supina. Questo aiuterebbe a stabilire quanto della variabilità da notte a notte sia da attribuire alla diversa posizione durante il sonno.

### 5.3.2. Protocolli di studio alternativi

A parte gli aspetti tecnici discussi, l'utilità di un approccio metodologico si basa sull'outcome della diagnosi. Se con l'utilizzo del sistema portatile gli outcome sono simili a quelli della polisonnografia, lo scambio è giustificabile, anche se la sensibilità e la specificità sono ridotte. Determinare gli outcome da misurare è tuttavia difficile. Sintomi quali la sonnolenza sono difficili da misurare oggettivamente. La valutazione soggettiva (con questionari sulla sonnolenza) fornisce risultati spuri se il soggetto si rende conto che l'intervento è in placebo. Per questo ai pazienti dovrebbe essere chiesto se si considerano in un gruppo placebo o di intervento. Outcome come la riduzione degli incidenti del traffico o dell'occorrenza di ictus ed infarto del miocardio possono richiedere anni per essere valutati e richiedono popolazioni di studio ampie. L'utilizzo delle risorse sanitarie e lo stato di salute sono misure difficili, applicabili a gruppi di soggetti, ma difficili da applicare nei singoli pazienti.

Un altro approccio utile per la diagnosi e le strategie di trattamento è quello di usare algoritmi che includono sia la polisonnografia che il monitoraggio portatile. Ciascuna strategia potrebbe essere studiata per valutare se outcome come i sintomi, lo stato di salute, l'utilizzo di risorse sanitarie o il costo/beneficio sono paragonabili per ciascun braccio di un dato algoritmo e ciò potrebbe rappresentare una convalida delle linee guida sulla pratica clinica. Per esempio, i bracci dello studio potrebbero essere i seguenti: 1) studio split-night; 2) monitoraggio portatile sorvegliato seguito da auto-titolazione non sorvegliata per determinare la pressione efficace della CPAP; 3) monitoraggio portatile sorvegliato split-night; 4) pazienti con alta probabilità pre-test che non vengono sottoposti a studio diagnostico, ma avviati direttamente alla titolazione della CPAP in laboratorio o in setting non sorvegliato con un sistema di auto-titolazione. Il numero di approcci possibili e i bracci decisionali possono essere molteplici, e, ai pazienti che non migliorano ad un determinato punto, vengono offerte valutazioni alternative. In questo tipo di approccio il sistema di monitoraggio portatile non è valutato come un test singolo ma nell'ambito di una strategia più ampia. Quei

bracci che determinano outcome simili potrebbero essere considerati equivalenti.

È necessario stabilire delle priorità di ricerca. La richiesta di test clinici per studiare i disturbi respiratori del sonno sta crescendo.<sup>86</sup> Come si evince da questo report la necessità più impellente è quella di studi di alta qualità per valutare la performance dei sistemi portatili nel setting non sorvegliato. Infatti l'aspetto più importante relativo ai sistemi di monitoraggio portatili è quello di comprendere la loro utilità nel setting nel quale dovrebbero essere usati. Gli aspetti più cruciali della ricerca futura sui sistemi portatili sono la scelta di outcome appropriati e di metodi con la massima efficienza diagnostica e terapeutica.

### BIBLIOGRAFIA

- 1 American Academy of Sleep Medicine. Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research; the report of an American Academy of Sleep Medicine task force. *Sleep* 1999; 22:667-689
- 2 Young T, Palta M, Dempsey J, et al. The occurrence of sleep-disordered breathing among middle-aged adults. *N Engl J Med* 1993; 328:1230-1235
- 3 Young T, Evans L, Finn L, et al. Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women. *Sleep* 1997; 20:705-706
- 4 Ferber RA, Millman RP, Coppola MP, et al. ASDA standards of practice: portable recording in the assessment of obstructive sleep apnea. *Sleep* 1994; 17:378-392
- 5 American Sleep Disorders Association. Practice parameters for the indications for polysomnography and related procedures: Polysomnography Task Force, American Sleep Disorders Association Standards of Practice Committee. *Sleep* 1997; 20:406-422
- 6 Chesson AL, Ferber RA, Fry JM, et al. The indications for polysomnography and related procedures. *Sleep* 1997; 20:423-487
- 7 Ross SD, Allen IE, Harrison KJ, et al. Systematic review of the literature regarding the diagnosis of sleep apnea: evidence report/technology assessment No. 1. Rockville, MD: Agency for Health Care Policy and Research; February 1999; AHCPR Publication No. 99-002
- 8 Ross SD, Sheinbait IA, Harrison KJ, et al. Systematic review and meta-analysis of the literature regarding the diagnosis of sleep apnea. *Sleep* 2000; 23:519-532
- 9 Sackett DL, Strauss SE, Richardson WS, et al. Evidence-based medicine: how to practice and teach EBM. 2nd ed. Edinburgh, Scotland, UK: Churchill Livingstone, 2000
- 10 Rees K, Wraith PK, Berthon-Jones M, et al. Detection of apnoeas, hypopnoeas and arousals by the AutoSet in the sleep apnoea/hypopnoea syndrome. *Eur Respir J* 1998; 12:764-769
- 11 Gugger M. Comparison of ResMed AutoSet (version 3.03) with polysomnography in the diagnosis of the sleep apnoea/hypopnoea syndrome. *Eur Respir J* 1997; 10:587-591
- 12 Mayer P, Meurice J-C, Philip-Joet F, et al. Simultaneous laboratory-based comparison of ResMed Autoset™ with polysomnography in the diagnosis of sleep apnoea/hypopnoea syndrome. *Eur Respir J* 1998; 12:770-775
- 13 Kiely J, Delahunty C, Matthews S, et al. Comparison of a limited computerized diagnostic system (ResCare Autoset™) with polysomnography in the diagnosis of obstructive sleep apnoea syndrome. *Eur Respir J* 1996; 9:2360-2364

- 14 Bradley P, Mortimore I, Douglas N. Comparison of polysomnography with ResCare Autoset in the diagnosis of the sleep apnoea/hypopnoea syndrome. *Thorax* 1995; 50:1201-1203
- 15 Bagnato M, Nery L, Moura S, et al. Comparison of AutoSet and polysomnography for the detection of apnea-hypopnea events. *Braz J Med Biol Res* 2000; 33:515-519
- 16 Gugger M, Mathis J, Bassetti C. Accuracy of an intelligent CPAP machine with in-built diagnostic abilities in detecting apneas: a comparison with polysomnography. *Thorax* 1995; 50:1199-1201
- 17 Fleury B, Rakotonanahary D, Hausser-Hauw C, et al. A laboratory validation study of the diagnostic mode of the Autoset™ system for sleep-related respiratory disorders. *Sleep* 1996; 19:502-505
- 18 Redline S, Tosteson T, Boucher MA, et al. Measurement of sleep-related breathing disturbances in epidemiologic studies: assessment of the validity and reproducibility of a portable monitoring device. *Chest* 1991; 100:1281-1286
- 19 Fry J, DiPhillipo MA, Curran K, et al. Full polysomnography in the home. *Sleep* 1998; 21:635-642
- 20 Portier F, Portmann A, Czernichow P, et al. Evaluation of home versus laboratory polysomnography in the diagnosis of sleep apnea syndrome. *Am J Respir Crit Care Med* 2000; 162:814-818
- 21 Zucconi M, Ferini-Strambi L, Castronovo V, et al. An unattended device for sleep-related breathing disorders: validation study in suspected obstructive sleep apnoea syndrome. *Eur Respir J* 1996; 9:1251-1256
- 22 Emsellem H, Corson W, Rappaport B, et al. Verification of sleep apnea using a portable sleep apnea screening device. *South Med J* 1990; 83:748-752
- 23 Man G, Kang B. Validation of a portable sleep apnea monitoring device. *Chest* 1995; 108:388-393
- 24 Ancoli-Israel S, Mason W, Coy T, et al. Evaluation of sleep disordered breathing with unattended recording: the Nightwatch System. *J Med Eng Technol* 1997; 21:10-14
- 25 Mykityn IJ, Sajkov D, Neill AM, et al. Portable computerized polysomnography in attended and unattended settings. *Chest* 1999; 115:114-122
- 26 Parra O, Garcia-Escelans N, Montserrat JM, et al. Should patients with sleep apnoea/hypopnoea syndrome be diagnosed and managed on the basis of home sleep studies? *Eur Respir J* 1997; 10:1720-1724
- 27 Ballester E, Solans M, Vila X, et al. Evaluation of a portable respirator recording device for detecting apneas and hypopnoeas in subjects from a general population. *Eur Respir J* 2000; 16:123-127
- 28 Ficker JH, Wiest GH, Wilpert J, et al. Evaluation of a portable recording device (Somnocheck®) for use in patients with suspected obstructive sleep apnoea. *Respiration* 2001; 68:307-312
- 29 White D, Gibb T, Wall J, et al. Assessment of accuracy and analysis time of a novel device to monitor sleep and breathing in the home. *Sleep* 1995; 18:115-126
- 30 Whittle AT, Finch SP, Mortimore IL, et al. Use of home sleep studies for diagnosis of the sleep apnoea/hypopnoea syndrome. *Thorax* 1997; 52:1068-1073
- 31 Verse T, Pirsig W, Junge-Hulsing B, et al. Validation of the POLY-MESAM seven-channel ambulatory recording unit. *Chest* 2000; 117:1613-1618
- 32 Orr WC, Eiken T, Pegram V, et al. A laboratory validation study of a portable system for remote recording of sleep-related respiratory disorders. *Chest* 1994; 105:160-162
- 33 Baltzan M, Verschelden P, Al-Jahdali H, et al. Accuracy of oximetry with thermistor (OxiFlow) for diagnosis of obstructive sleep apnea and hypopnea. *Sleep* 2000; 23:61-69
- 34 Stoohs R, Guilleminault C. MESAM 4: an ambulatory device for the detection of patients at risk for obstructive sleep apnea syndrome (OSAS). *Chest* 1992; 101:1221-1227
- 35 Gurubhagavatula I, Maislin G, Pack AI. An algorithm to stratify sleep apnea risk in a sleep disorders clinic population. *Am J Respir Crit Care Med* 2001; 164:1904-1909
- 36 Vazquez J, Tsai W, Flemons W, et al. Automated analysis of digital oximetry in the diagnosis of obstructive sleep apnoea. *Thorax* 2000; 55:302-307
- 37 Yamashiro Y, Kryger MH. Nocturnal oximetry: is it a screening tool for sleep disorders? *Sleep* 1995; 18:167-171
- 38 Wiltshire N, Kendrick A, Catterall J. Home oximetry studies for diagnosis of sleep apnea/hypopnea syndrome: limitation of memory storage capabilities. *Chest* 2001; 120:384-389
- 39 Schafer H, Ewig S, Hasper E, et al. Predictive diagnostic value of clinical assessment and nonlaboratory monitoring system recordings in patients with symptoms suggestive of obstructive sleep apnea syndrome. *Respiration* 1997; 64:194-199
- 40 Gyulay S, Olson LG, Hensley MJ, et al. A comparison of clinical assessment and home oximetry in the diagnosis of obstructive sleep apnea. *Am Rev Respir Dis* 1993; 147:50-53
- 41 Golpe R, Jimenez A, Carpizo R, et al. Utility of home oximetry as a screening test for patients with moderate to severe symptoms of obstructive sleep apnea. *Sleep* 1999; 22:932-937
- 42 Williams AJ, Yu G, Santiago S, et al. Screening for sleep apnea using pulse oximetry and a clinical score. *Chest* 1991; 100:631-635
- 43 Esnaola S, Duran J, Infante-Rivard C, et al. Diagnostic accuracy of a portable recording device (MESAM IV) in suspected obstructive sleep apnea. *Eur Respir J* 1996; 9:2597-2605
- 44 Koziej M, Cieslicki J, Gorzelak K, et al. Hand-scoring of MESAM 4 recordings is more accurate than automatic analysis in screening for obstructive sleep apnoea. *Eur Respir J* 1994; 7:1771-1775
- 45 Chiner E, Signes-Costa J, Arriero J, et al. Nocturnal oximetry for the diagnosis of the sleep apnoea syndrome: a method to reduce the number of polysomnographies? *Thorax* 1999; 54:968-971
- 46 Bonsignore G, Marrone O, Macaluso C, et al. Validation of oximetry as a screening test for obstructive sleep apnea syndrome. *Eur Respir J* 1990; 11(suppl):542S-544S
- 47 Rauscher H, Popp W, Zwick H. Quantification of sleep disordered breathing by computerized analysis of oximetry, heart rate and snoring. *Eur Respir J* 1991; 4:655-659
- 48 Cooper B, Veale D, Griffiths C, et al. Value of nocturnal oxygen saturation as a screening test for sleep apnea. *Thorax* 1991; 46:586-588
- 49 Levy P, Pepin J, Deschaux-Blanc C, et al. Accuracy of oximetry for detection of respiratory disturbances in sleep apnea syndrome. *Chest* 1996; 109:395-399
- 50 Pepin J, Levy P, Lepaulle B, et al. Does oximetry contribute to the detection apneic events?: mathematical processing of the Sao<sub>2</sub> signal. *Chest* 1991; 99:1151-1157
- 51 Olson LG, Ambrogetti A, Gyulay SG. Prediction of sleep-disordered breathing by unattended overnight oximetry. *J Sleep Res* 1999; 8:51-55
- 52 Zamarron C, Romero P, Rodriguez J, et al. Oximetry spectral analysis in the diagnosis of obstructive sleep apnoea. *Clin Sci* 1999; 97:467-473
- 53 Series F, Marc I, Cormier Y, et al. Utility of nocturnal home oximetry for case finding in patients with suspected sleep apnea hypopnea syndrome. *Ann Intern Med* 1993; 119:449-453

- 54 Douglas N, Thomas S, Jan M. Clinical value of polysomnography. *Lancet* 1992; 339:347–350
- 55 Rauscher H, Popp W, Zwick H. Model for investigating snorers with suspected sleep apnea. *Thorax* 1993; 48:275–279
- 56 Ryan PJ, Hilton MF, Boldy DAR, et al. Validation of British Thoracic Society guidelines for the diagnosis of sleep apnoea/hypopnea syndrome: can polysomnography be avoided? *Thorax* 1995; 50:972–975
- 57 Issa F, Morrison D, Hadjuk E, et al. Digital monitoring of sleep-disordered breathing using snoring sound and arterial oxygen saturation. *Am Rev Respir Dis* 1993; 148:1023–1029
- 58 Zamarron C, Romero PV, Gude F, et al. Screening of obstructive sleep apnoea: heart rate spectral analysis of nocturnal pulse oximetric recording. *Respir Med* 2001; 95: 759–765
- 59 Reda M, Gibson GJ, Wilson JA. Pharyngoesophageal pressure monitoring in sleep apnea syndrome. *Otolaryngol Head Neck Surg* 2001; 125:324–331
- 60 Claman D, Murr A, Trotter K. Clinical validation of the Bedbugg™ in detection of obstructive sleep apnea. *Otolaryngol Head Neck Surg* 2001; 125:227–230
- 61 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1:307–310
- 62 Streiner DL, Norman GR. Health measurement scales: a practical guide to their development. 2nd ed. New York, NY: Oxford University Press, 1995
- 63 Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999; 282:1061–1066
- 64 Jaeschke R, Guyatt G, Sackett DL. Users' guide to the medical literature: how to use an article about a diagnostic test; A. Are the results of the study valid? *JAMA* 1994; 271:389–391
- 65 Jaeschke R, Guyatt G, Sackett DL. Users' guide to the medical literature: how to use an article about a diagnostic test; B. What are the results and will they help me care for my patients? *JAMA* 1994; 271:703–707
- 66 West S, King V, Carey TS, et al. Systems to rate the strength of scientific evidence: evidence report/technology assessment No. 47. Rockville, MD: Agency for Healthcare Research and Quality; April 2002; AHRQ Publication No. 02-E016
- 67 Simel DL, Samsa GP, Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. *J Clin Epidemiol* 1991; 44:763–770
- 68 Acebo C, Watson R, Bakos L, et al. Sleep and apnea in the elderly: reliability and validity of 24-hour readings in the home. *Sleep* 1991; 14:56–64
- 69 Ancoli-Israel S, Kripke D, Klauber M, et al. Sleep-disordered breathing in community-dwelling elderly. *Sleep* 1991; 6:486–495
- 70 Farney R, Walker L, Jensen R, et al. Ear oximetry to detect apnea and differentiate rapid eye movement (REM) and non-REM (NREM) sleep: screening for the sleep apnea syndrome. *Chest* 1986; 89:533–539
- 71 Gyulay S, Gould D, Sawyer B, et al. Evaluation of a micro-processor-based portable home monitoring system to measure breathing during sleep. *Sleep* 1987; 10:130–142
- 72 Hida W, Shindoh C, Miki H, et al. Prevalence of sleep apnea among Japanese industrial workers determined by a portable sleep monitoring system. *Respiration* 1993; 60:332–337
- 73 Lojander J, Salmi T, Maasilta P. Reproducibility of oximetry with a static charge-sensitive bed in evaluation of obstructive sleep apnoea. *Clin Physiol* 1998; 18:225–233
- 74 Polo O, Brissaud I, Sales B, et al. The validity of the static charge sensitive bed in detecting obstructive sleep apnea. *Eur Respir J* 1988; 1:330–336
- 75 Roche F, Gaspoz J, Court-Fortune I, et al. Screening of obstructive sleep apnea syndrome by heart rate variability analysis. *Circulation* 1999; 100:1411–1415
- 76 Rollheim J, Osnes T, Miljeteig H. The sites of obstruction in OSA, identified by continuous measurements of airway pressure and flow during sleep: ambulatory versus in-hospital recordings. *Clin Otolaryngol* 1999; 24:502–506
- 77 Salmi T, Partinen M, Hyyppa M, et al. Automatic analysis of static charge sensitive bed (SCSB) recordings in the evaluation of sleep-related apneas. *Acta Neurol Scand* 1986; 74: 360–364
- 78 Stoohs R, Guilleminault C. Investigations of an automatic screening device (MESAM) for obstructive sleep apnoea. *Eur Respir J* 1990; 3:823–829
- 79 Svanborg E, Larsson H, Carlsson-Nordlander B, et al. A limited diagnostic investigation for obstructive sleep apnea syndrome: oximetry and static charge sensitive bed. *Chest* 1990; 98:1341–1345
- 80 Tvinnereim M, Mateika S, Cole P, et al. Diagnosis of obstructive sleep apnea using a portable transducer catheter. *Am J Respir Crit Care Med* 1995; 152:775–779
- 81 Chediak AD, Acevedo-Crespo JC, Seiden DJ, et al. Nightly variability in the indices of sleep-disordered breathing in men being evaluated for impotence with consecutive night polysomnograms. *Sleep* 1996; 19:589–592
- 82 Tashkandi YN, Rowley JA, Badr MS. Effect of gender on PSG indices of sleep-disordered breathing [abstract]. *Am J Respir Crit Care Med* 2000; 161:A680
- 83 Redline S, Tishler PV, Hans MG, et al. Racial differences in sleep-disordered breathing in African-Americans and Caucasians. *Am J Respir Crit Care Med* 1997; 155:186–192
- 84 Le Bon O, Hoffmann G, Tecco J, et al. Mild to moderate sleep respiratory events: one negative night may not be enough. *Chest* 2000; 118:353–359
- 85 Littner MR. The diagnosis of OSA: where do we draw the line [editorial]? *Chest* 2000; 118:286–288
- 86 Namen AM, Dunagan DP, Fleischer A, et al. Increased physician-reported sleep apnea: the National Ambulatory Medical Care Survey. *Chest* 2002; 121:1741–1747