

Appendix: Table 1

Reference number	Article	Simulated patient	Evaluation strategy	Relevant limitations stated by author	MMAT score
2	O'Leary 2013	Child	Feedback forms for simulation	<ul style="list-style-type: none"> • Single centre study • Doctors may perform differently in real patient encounters 	100%
			Performance evaluation: Time to event assessment		
8	Happel 2015	Not stated	Feedback forms for preceding simulation experience following live clinical events	<ul style="list-style-type: none"> • Pilot study with limited time frame • Measured effectiveness of simulation training relied on self-perception • Recall bias may have positively or negatively affected the perceived simulation training effectiveness • Trainees report a false sense of confidence when skills levels are low (Dunning-Kruger effect) • Response rate to questionnaire may have been affected by more engaged residents responding to questionnaire selectively 	100%
9	Patterson 2013	Not stated	Feedback forms for simulation	<ul style="list-style-type: none"> • Only 54% response rate with feedback forms 	75%
			Other: Identification of latent safety threats		
10	Zimmerman 2015	Infant Child	Feedback forms for simulation	<ul style="list-style-type: none"> • Single centre, observational study setting with a short study period • Did not assess the impact of the program in regard to patient outcome or performance in real life critical events 	100%
			Performance evaluation: TeamMonitor teamwork assessment tool following live critical events		
			Other: Identification of latent safety threats during live events and simulation		

11	Auerbach 2014	Infant	Performance evaluation: Trauma Simulation Analysis Tool	<ul style="list-style-type: none"> • Simulated cases were trauma patient only • The reliability of scoring could not be scored as the sessions were not video recorded 	75%
		Child Teenager	Feedback forms for simulation		
12	Cheng 2010	Infant Child	Feedback forms for simulation	<ul style="list-style-type: none"> • Did not assess patient level outcomes 	100%
13	Tsai 2006	Child	Feedback forms for simulation	<ul style="list-style-type: none"> • Did not assess for long term learning gains 	100%
			Performance evaluation: Score derived from combination of task specific technical skills, medication and behavioural scores		
14	Auerbach 2011	Infant	Feedback forms for simulation	<ul style="list-style-type: none"> • Did not assess patient level outcomes 	75%
15	Couto 2015	Not stated	Performance evaluation: Team Assessment Score	<ul style="list-style-type: none"> • Single centre, retrospective study • Did not assess interrater reliability 	100%
16	Greenberg 1999	Simulated parent	Feedback forms for simulation	<ul style="list-style-type: none"> • Did not assess for long term learning gains 	100%
			Performance evaluation: Task completion checklist		
17	Hunt 2006	Child	Performance evaluation: Task completion checklist	<ul style="list-style-type: none"> • Tool had not been validated • Doctors may perform differently in real patient encounters • Did not assess patient level outcomes 	100%
18	Schmutz 2015	Infant	Performance evaluation: Task completion checklist	<ul style="list-style-type: none"> • Complex checklist time consuming to complete • Did not consider non-technical aspects • Did not assess patient level outcomes 	100%

19	Donoghue 2010	Not stated	Performance evaluation: Task completion checklist	<ul style="list-style-type: none"> • Did not assess patient level outcomes • Clinical meaning of the resulting score – may solely reflect how common various conditions (i.e. trainee confidence with management.) • Doctors may perform differently in real patient encounters 	100%
20	Tofil 2014	Infant	Performance evaluation: Time to event assessment	<ul style="list-style-type: none"> • Did not assess for long term learning gains 	100%
21	O’Leary 2014	Infant Child	Other: Episodes of suboptimal care and their causation	<ul style="list-style-type: none"> • Incidents and causation factors identified are a function of the individual scenario 	100%
22	Katznelson 2014	Child	Other: Provider comfort scores	<ul style="list-style-type: none"> • 71% response rate to pre intervention surveys – did not capture all • Did not assess patient level outcomes 	100%

Table 1:

The table above summarises the simulated patient age, evaluation strategy, limitations as stated by the authors and the Mixed Methods Appraisal Tool (MMAT) scores for the articles which met the inclusion criteria for review.