BEME RAPID REVIEW

Problem-based learning in pre-clinical medical education: 22 years of outcome research

Lisa Hartling, MSc [1]
Carol Spooner, MSc [1]
Lisa Tjosvold, MLIS [1]
Anna Oswald, MD [2]

1 Alberta Research Center for Health Evidence, Department of Pediatrics, University of Alberta, Edmonton, Canada; 2 Division of Rheumatology, Department of Medicine, University of Alberta, Edmonton, Canada

Table and Figures

Figure 1. Flow of studies through the selection process

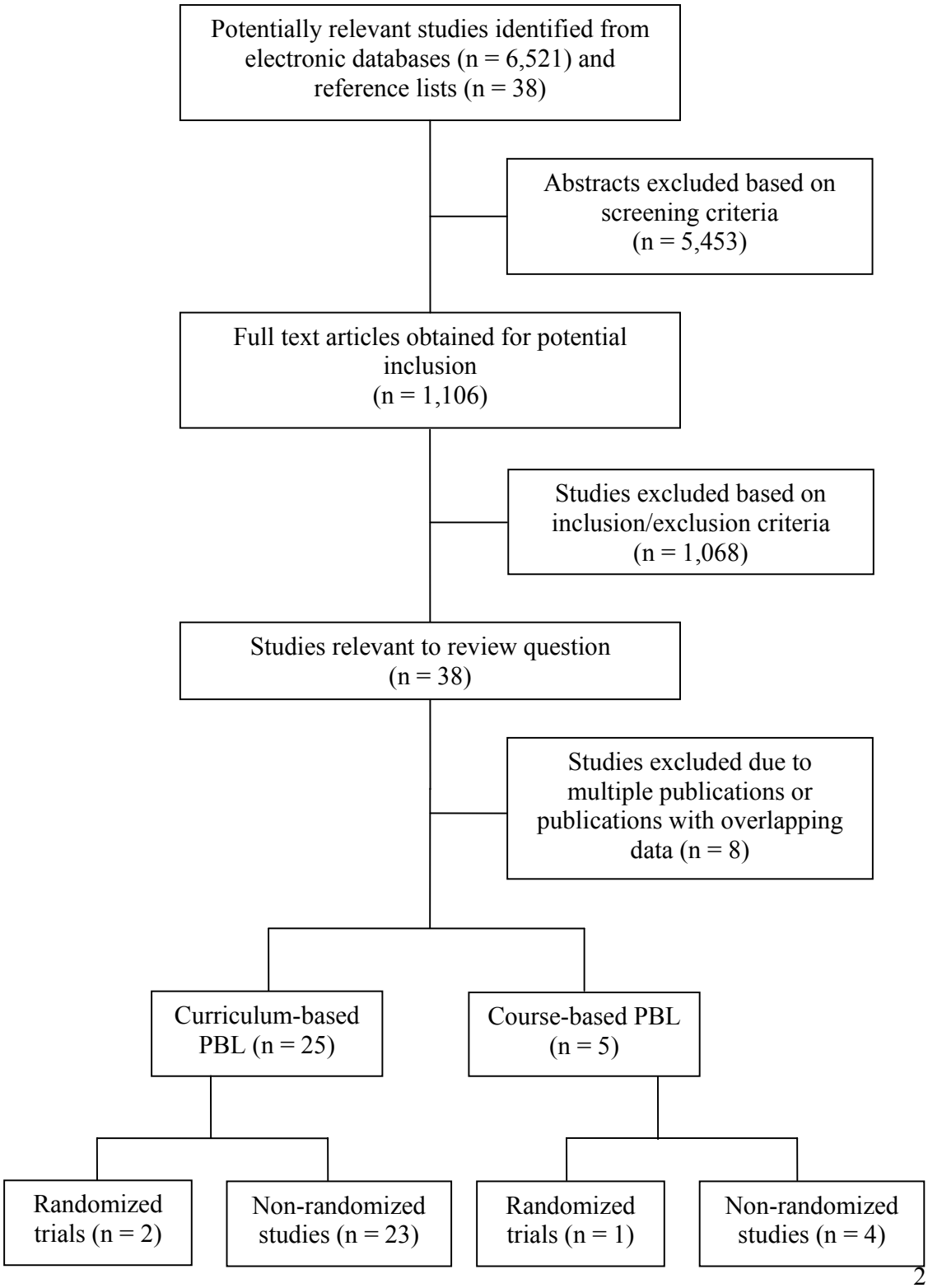

Table 1. Methodological quality of previous reviews as assessed by the Oxman and Guyatt Index{Oxman & Guyatt 1991 #7558}*

First author Year of publication	Search methods stated	Search comprehensive	Criteria for study inclusion reported	Bias in study selection avoided	Criteria for assessing internal validity reported	Validity assessed using appropriate criteria	Method of analysis reported	Studies combined/ analyzed appropriately	Conclusions supported by data	Summary Score (0 major flaws to 7 minimal flaws)
Albanese 1993	partial	no	no	unclear	no	no	yes	yes	partial	1
Colliver 2000	yes	no	partial	no	no	no	yes	yes	partial	1
Dochy 2003	yes	yes	yes	unclear	no	no	yes	yes	yes	2
Kalaian 1999	yes	partial	no	unclear	no	no	yes	yes	yes	2
Koh 2008	yes	yes	yes	unclear	yes	yes	yes	yes	yes	5
Newman 2003	yes	no	yes	yes	yes	yes	yes	yes	yes	3†
Vernon 1993	yes	partial	yes	unclear	no	no	yes	no	yes	2

* studies were assessed independently by two reviewers (LH, KB) and discrepancies were resolved through discussion

† study described as a pilot systematic review to assess the feasibility of conducting a systematic review and meta-analysis in this area; hence, the main limitation was the scope of the search

Table 2. Summary of Findings

Outcome	Findings	Study design
Knowledge (exam scores)	No difference between PBL and non-PBL for all or parts of exam	3 randomized trials and 12 observational studies
	Benefits favouring non-PBL for parts of exam or content-specific exam	1 randomized trial and 1 observational study (respectively)
	Benefits favouring PBL	2 observational studies and 1 interrupted time series
Recall	Benefits for non-PBL group for short-term recall	1 non-randomized trial
	No difference between PBL and non-PBL for long-term recall	1 non-randomized trial
Clerkship performance	Benefits for PBL group	3 observational studies
Intern performance	Better ratings for non-PBL group on majority of competencies	1 observational study
	No difference between PBL and non-PBL	1 observational study
Residency performance	Benefits favouring PBL (improved communication and collaboration skills, maturity, initiative)	1 observational study
	No difference between PBL and non-PBL	1 observational study
Post-graduate practice	No clinically or statistically significant benefits for PBL vs. non-PBL	2 observational studies
	Different practice patterns observed for PBL graduates	1 observational study
Diagnostic accuracy	Benefits favouring PBL students	3 observational studies
	Benefits favouring non-PBL students	1 observational study

Evidence Table 1. Description of studies evaluating curriculum-based PBL

First author, year of publication, country	Institution	Study design (n)	Description of PBL	Comparisons	Authors' primary outcome(s)
Albano 1996 Italy, Germany, The Netherlands	University of Limburg, Maastricht (The Netherlands); Universities of Bari, Torino Modena, Perugia (Italy); University of Witten/Herdecke (Germany)	cross-sectional (n=2,132*)	**PBL** (n=unclear): integrated, problem-based	**conventional curriculum** (n=unclear): discipline-oriented	knowledge as measured by test scores on Maastricht Progress Test
Boshuizen 1993 The Netherlands	University of Limburg (Maastricht) and one other Dutch medical school	cross-sectional (n=8†)	**PBL** (n=4): integration of biomedical and clinical sciences through problem-based learning	**traditional curriculum** (n=4): biomedical subjects studied first, followed by clinical sciences	knowledge integration (reflected in strategies applied to problem solving)
Distlehorst 2005 US	Southern Illinois University School of Medicine	retrospective cohort (n=648)	**PBL** (n=195): student-directed; patient-problem work; "only required activity other than tutor-group meetings was the weekly, half-day clinical experience"	**standard curriculum** (n=453): not described	Performance outcomes: USMLE 1 and 2; clerkship ratings; clerkship honors and remediation designation; CCX
Enarson 2001 US	Wake Forest University School of Medicine, Winston-Salem	retrospective cohort (n=1,229)	**PBL** (n=NS): "emphasized small-group, student-centered, self-directed learning of basic and clinical science material"	**LBL** (n=NS): organized by discipline with lectures as primary teaching method	USMLE Step 1 and 2 exams
Farquhar 1986 US	Michigan State University College of Human Medicine	retrospective cohort (matched for MCAT subtest	**PBL** (n=40): 344 hours of scheduled time over 50 weeks (average 7 hours/week); problem-solving exercises in small groups	**lecture-based** (n=40): 958 hours of scheduled time over 50 weeks (average 19 hours/week); discipline-based	NBME I scores

		scores) (n=80)			
Hmelo 1998 US	University of Pittsburgh (author's affiliation); study involved 2 schools which were not specified	prospective cohort (n=76)	**PBL** (n=39): students work in small groups with **case-based instruction**	**traditional curriculum** (n=37): lectures and laboratory exercises	accuracy of problem solving
Hoffman 2006 US	University of Missouri-Columbia School of Medicine (UMCSOM)	interrupted time series (3-4 years pre and 10 years post) (n=720)	**PBL** (n=96/year for 4 years): primary method of instruction is PBL through small groups supplemented by lectures and laboratories; structured experiences provide clinical exposure	**traditional curriculum** (n=112/year for 3 years): NS	USMLE Step 1 and 2 exams
Lycke 2006 Norway	University of Oslo	prospective cohort (n=150)	**PBL** (n=61): tutor-led groups and integration of disciplines	**traditional curriculum** (n=81): lecture-based, teacher-directed, some group work	test of medical knowledge given in final 12th semester (not specified as primary, but other outcomes were not evaluative)
Mennin 1993 US	University of New Mexico School of Medicine	randomized trial (and prospective cohort) (n=675 overall; 119 randomized to curricular tracks)	**Primary Care Curriculum** (**PCC**) (n=167; 85 randomized): community-oriented, **PBL**; students define own learning issues, "no formal lectures or summative content-based examinations;" "early, extended, primary care learning experience in a rural setting"	**conventional** track (n=508; 34 randomized): NS	NBME I, II, and III
Moore 1994 US	Harvard Medical School	randomized trial (n=121)	**New Pathway (NP)** curriculum (n=60): "students participated actively in and assumed responsibility for their own learning in small faculty-led tutorial groups, where they used clinical problems as the basis for their study"	**traditional** program (n=61): "lectures and syllabi in disciplinary courses"	student performance (knowledge, clinical reasoning, psychosocial skills and attitudes, learning styles and preferences)

Patel 2001 Canada	McGill and McMaster Universities	prospective cohort (n=21)	**PBL** (n=12): "curriculum is centered on a series of health care problems that require students to explore basic science and clinical subject matter simultaneously"	**conventional** (n=9): "based on clear distinction between the teaching of basic science and the clinical aspects of patient care"	problem-solving performance of housestaff (many facets of problem-solving, including accuracy of diagnoses)
Remmen 2001 Belgium and Netherlands	Universities of Antwerp and Ghent, Belgium; Universities of Groningen and Maastricht, Netherlands	cross-sectional (n=859)	**PBL** (n=118): problem-based curriculum + longitudinal skills training and assessment; "skills laboratory training with simulated patients" beginning in first year	**traditional lecture-based** (n=741)	written skills test
Richards 1996 US	Bowman Gray School of Medicine of Wake Forest University, Winston-Salem, NC	retrospective cohort (n=492)	**PBL** (n=91): "small group case discussions and independent, interdisciplinary learning…includes frequent contact with real or simulated patients"	**LBL** (n=401): 4-5 hours lecture/day; patient contact limited "to supervised encounters with small number of hospitalized patients as part of bedside tutoring groups"	clinical rating scale (first listed)
Rolfe 1995 Australia	Interns in New South Wales (NSW), Australia from a variety of institutions primarily Universities of Newcastle, NSW and Sydney	cross-sectional survey (n=486)	**PBL** (n=49): integrated curriculum	**traditional/conventional** (n=335)	supervisor ratings during internship
Santos-Gomez 1990 US	University of New Mexico School of Medicine	cross-sectional (n=185)	**Primary Care Curriculum** (PCC) (n=52): "small group, problem-based and self-directed learning in the first 2 years…also includes an early extended community-based preceptorship during the first year of medical school"	**conventional** track (n=133): "integrated teaching approach in the first and second year courses with emphasis on interdepartmental cooperation in teaching"	residency performance measured on 8 dimensions by Anchored Residency Evaluation Form

Schmidt 1996 Netherlands	University of Limberg (Maastricht), University of Amsterdam, University of Groningen, Netherlands	cross-sectional (n=612)	**PBL** (n=NS): "students meet twice a week for small-group discussion of problems…limited number of lectures, laboratory activities;" remainder of time is for self-directed learning	**conventional curriculum** (n=NS): "discipline-oriented and teacher-centered;" **integrated curriculum** (n=NS): "integrates the biomedical and clinical sciences around major organ systems. Students engage in patient demonstrations and small-group training sessions" with case-based learning	number of accurate diagnoses for 30 case histories
Shin 1993 Canada	McMaster University and University of Toronto	cross-sectional survey (n=96)	**PBL** (n=48): problem-based, self-directed	**traditional** (n=48): passive, didactic	post-graduate knowledge regarding management of hypertension
Tamblyn 2005 Canada	Sherbrooke University, Quebec; 3 traditional medical schools in Quebec (Laval, Montreal, McGill)	historical (retrospective) cohort (n=813)	community-oriented **PBL** curriculum (n=165): lectures replaced by problem-oriented tutorials in small groups; increase in community-based clinical experience during pre-clinical years and clerkship	**traditional** curricula (n=648)	annual performance in preventive care (mammography screening rate) (first listed)
Van Hessen 1990 Netherlands	University of Maastricht	cross-sectional (n=unknown)	**PBL** (number not provided): see Verhoeven below	**traditional** curricula (n=179 out of 1449 invited)	test scores on Maastricht Progress Test
Verhoeven 1998 Netherlands	University of Maastricht and University of Nijmegen, The Netherlands	prospective cohort (n=2,993)	**PBL** (n=1,904): "problem-based, student-centered and integrated curriculum that is organized in themes"	**conventional** (n=1,098): discipline-oriented educational method…(with) large group lectures and structured laboratory experiences"	test scores on "comprehensive final examination in medicine"
Verwijnen 1990 Netherlands	University of Maastricht and three other medical schools in the Netherlands	cross-sectional (number of unique subjects unclear)	**PBL:** as described for Verhoeven	**conventional**	test scores on various tests including the Maastricht Progress Test

Way 1999 US	Ohio State University College of Medicine and Public Health	retrospective cohort (n=778)	**PBL pathway** (n=NS): "emphasizes student-centered, self-directed learning…analysis and discussion of clinical cases during group meetings"	**lecture-discussion pathway** (n=NS): "large group lectures supplemented with small group tutorials;" **independent study pathway** (n=NS): independent learning with "highly structured reading materials, computer-based materials, and diagnostic practice examinations"	performance on USMLE Step 1
Whitfield 2002 US	Penn State College of Medicine	prospective cohort (n=617)	**PBL** (n=135): independent study and small group discussion of clinical cases; supplemental laboratory exercises and human simulator experiments; no scheduled lectures	**LBL** (n=482): 5 to 7 hr per day of didactic lectures and laboratories, with occasional clinical correlation lectures"	clerkship scores
Woodward 1990 Canada	McMaster University	cross-sectional survey (n=152)	**PBL** (n=105): solve problems in small groups lead by faculty tutor; basic and clinical sciences are integrated; self-directed learning; few lectures are optional	**non-PBL medical schools** (n=47)	performance ratings by internship supervisors
Woodward 1990 Canada	McMaster University and four other Ontario medical schools	cross-sectional (n=644) groups matched on year of graduation, sex, urban/rural practice, residence (billing district)	**PBL** (n=322): see Woodward above	**non-PBL medical schools** (n=322): emphasis on didactic, lecture-based teaching	billing patterns

* knowledge results based on n=299 from four schools

† total sample size was 12 which included two biochemists and two internists; their educational background was not reported, therefore the information presented here is restricted to the comparison of students from PBL vs. traditional curricula

PBL=problem-based learning; LBL=lecture-based learning; USMLE=United States Medical Licensing Examination; CCX=Clinical Competency Examination; NS=not specified; MCAT=Medical College Admission Test; NBME=National Board of Medical Examiners

Evidence Table 2. Results of studies evaluating curriculum-based PBL

First author, year of publication, country	Outcome	Quantitative results	Authors' conclusions	Stated limitations
Albano 1996 Italy, Germany, The Netherlands	test scores	No significant differences between Maastricht and three other schools for overall mean test scores among final (sixth) year students (range of means across four schools 41.5-44.4%). For specific domains, compared to other three schools, Maastricht students scored 5.0-7.5% higher on clinical sciences, 12.5-18.6% higher on public health/behavioural sciences, and 9.4-11.7% lower on basic sciences. Data presented on 'kinetics of knowledge acquisition' over six years of medical school.	Overall mean test scores for final year students from different schools were "surprisingly similar." Variations across domains were observed with Maastricht students performing better on domains of clinical and public health/behavioural sciences, but poorer on basic sciences.	PBL students familiar with test format; circumstances of participation varied across schools; cultural context may explain some variability particularly for public health/behavioural sciences domain
Boshuizen 1993 The Netherlands	accuracy of answers to a single clinical question	Comparisons between PBL and traditional students showed: no significant difference in the number of 'answer relevant terms' (data not presented); significant differences observed in 'acceptable answer components' and 'acceptable answers' favouring PBL. Data also presented on process (i.e., reasoning strategies).	PBL students were "better able to find an acceptable answer to the question."	NS
Distlehorst 2005 US	USMLE exam scores	No significant difference between PBL and standard curriculum for USMLE Step 1 (ES=-0.01, p=0.92) or USMLE Step 2 (ES 0.13, p=0.15)	"PBL students performed as well as, and in some cases better than, their STND counterparts."	not randomized; inability to measure some less tangible outcomes that make 'better' doctors
	overall clerkship performance	Significant difference favouring PBL vs. standard curriculum (p=0.008; ES=0.23); PBL had significantly higher scores in 2 of 6 clerkships		
	clerkship subcategories	PBL performed significantly better on all clerkship subcategories (ES ranged from 0.19 to 0.24); PBL students had significantly more honors; no significant difference in remediations		

	CCX	No significant difference overall (ES 0.06; p=0.52) or for 3 subcategories (ES ranged from -0.04 to 0.24 for ITT analysis)		
Enarson 2001 US	Exam scores	No significant differences between PBL and LBL for USMLE I (p=0.21) or USMLE II (p=0.30)	Standardized test performance was comparable. Traditional students had higher mean exam scores but no statistically significant difference.	student characteristics may explain outcomes (e.g., MCAT scores)
Farquhar 1986 US	NBME Part 1 overall	Results presented as difference in scores (LBL-PBL): overall difference of 5.88 (not significant); no significant differences on 6/7 subtests; microbiology subtest difference of 47.62 favoured LBL (p<0.05)	"Reduction of total scheduled and lecture time, at least when replaced by a guided, problem-solving program…is not detrimental to students' performance on an independent assessment." Longer-term studies needed.	small numbers comparing only 1 year of students
	Pass rate	No significant difference between PBL (90%) and LBL (87.5%)		
Hmelo 1998 US	diagnostic accuracy of problem solving	greater improvement for PBL vs. non-PBL students, p<0.05	PBL students improved more than nonPBL students with respect to accuracy of problem solving, coherence of explanations, use of hypotheses-driven reasoning strategies, justification of explanations, and use of science concepts to support explanations. Observations were consistent when previous problem-specific experience was taken into account.	NS
Hoffman 2006 US	USMLE Step 1 (basic sciences)	For PBL, mean scores for 6/10 classes significantly greater than national average (ES: -0.11 to 0.62) with an increase in numbers of students in 90th, 95th, and 99th percentile. For standard curriculum, 2/3 classes scored similar to national average, no trend in numbers of students who scored in the 90th, 95th, and 99th percentile.	PBL students showed enhanced performance on national licensing exams; improved performance was maintained over 10 years. "Analysis of residency performance data shows gains beyond knowledge, including PBL graduates' enhanced communication and collaboration skills, maturity, initiative, and the ability to project the qualities of a good physician."	Retrospective pre-post design; data collected (specifically from residency directors) was not designed for use in research; examines performance at one school only
	USMLE Step 2 (clinical medicine)	For PBL, mean scores of 6/9 classes significantly greater than national average (ES: 0.01 to 0.52); for standard curriculum, mean scores of last 4 classes near to or lower than average national scores		

	performance in first year residency	In all 17 comparisons PBL students scored higher than traditional students with 12/17 comparisons significant ($p<0.01$).		
Lycke 2006 Norway	test on medical knowledge during final 12th semester	No significant difference between PBL and traditional students (average 74% vs. 73% correct answers, p=0.79).	Both programmes resulted in a high level of medical knowledge. Authors noted differences in learning strategies and mental models of learning.	Tests of medical knowledge don't necessarily reflect differences in how knowledge is applied
Mennin 1993 US	NBME exam scores*	NBME I: Significant differences favouring conventional track (mean 521, SEM 13.4) over PBL (455, 8.5) NBME II: No significant difference between PBL (mean 472, SEM 10.2) and conventional track (485, 15.9) NBME III: No significant difference between PBL (mean 521, SEM 14.8) and conventional track (551, 20.7)	Structured curricula at the beginning of medical school may provide better learning environment for exam preparation.	NS
	failure rates*	NBME I: No significant difference between PBL (n=15, 18%) and conventional track (n=1, 3%) NBME II: No difference (0 failed in both groups)		
	overall NBME exam scores†	NBME I: Significant difference favouring conventional track (mean 504, SEM 3.44) over PBL (456, 6.07) NBME II: No significant difference between PBL (mean 469, SEM 7.02) and conventional track (460, 3.90) NBME III: Significant difference favouring PBL (mean 521, SEM 8.98) over conventional track (491, 5.13)		
Moore 1994 US	NBME Part 1	Overall, no significant difference between PBL (mean 0.06, SD 1.09) and LBL (0.07, 1.01); no difference between groups when tested in fourth year for recall of preventive medicine and biochemistry material taught in years 1 and 2	"The New Pathway curriculum led to a positive, but instructively stressful, educational experience; did no demonstrable harm; possibly enhanced humanism in the first two years; and, fostered students' development of self-directed learning skills."	complex intervention (susceptible to unanticipated confounders); selection bias; low participation and variable participation across outcomes; weak measurement tools
	clinical reasoning	No differences found on multiple measures		

Patel 2001 Canada	diagnostic accuracy	Significant differences favouring conventional curriculum vs. PBL: Case 1 (internal medicine): 91.7 vs. 82.0, p<0.05 Case 1 (family medicine): 83.0 vs. 60.0, p<0.05 Case 2 (internal medicine): 41.7 vs. 21.0, p<0.05 Case 2 (family medicine): 33.0 vs. 5.0, p<0.05	"Although all housestaff generated equal numbers of diagnostic hypotheses during the reasoning process, housestaff from the conventional curriculum generated a greater number of accurate hypotheses."	only tested 2 clinical cases; small sample limits generalizability; other confounders not ruled out
Remmen 2001 Belgium and Netherlands	written skills test	Confidence intervals don't overlap between PBL school and traditional schools for any years. Significant difference favoured PBL school over three traditional schools for each of four years.	PBL school prepared "students better for clerkships and probably enhanced learning during clerkships." PBL graduates are likely more skilled but conclusions are not definitive. More research needed to discern which elements of PBL create favourable effects.	low numbers from Dutch schools; cross-sectional; participation was voluntary therefore prone to selection bias; potential confounders not accounted for, most important may be "problem-based learning approach"
Richards 1996 US	clinical rating‡	Significant difference favouring PBL vs. conventional track for: -amount of factual knowledge: mean 3.11, SD 0.36 vs. 2.9, 0.42, p<0.0001 -take history and perform physical exam: mean 3.22, SD 0.34 vs. 3.05, 0.40, p=0.002 -derive differential diagnosis: mean 3.04, SD 0.35 vs. 2.86, 0.39, p=0.0005 -organize and express information: mean 3.49, SD 0.35 vs. 3.33, 0.41, p=0.004	After two pre-clinical years, PBL students had higher clinical ratings than LBL students (effect size of 0.5). Clerkship director concluded that "results were educationally meaningful;" however, authors state that the differences "were not large and may not be compelling enough, in and of themselves, to support the adoption of PBL."	authors controlled for differences in selection and rotation order through adjusted analyses
	NBME medicine shelf test	No difference between PBL (496.5, 95.2) and conventional track (489.4, 97.0), p=0.8		
Rolfe 1995 Australia	supervisor rating of interns	Analyses adjusted for age and gender showed PBL graduates significantly superior for 4/13 competencies (self-directed learning, reliability and dependability, relationships with patients and families, relationships with other professionals); graduates from a non-PBL school were significantly better on 3/13 competencies (understanding of basic mechanisms, diagnostic skills, teaching).	"The differences between graduates from problem-based and traditional NSW medical schools are relatively small."	potential for: inappropriate use of rating scale; overestimating real performance; measurement biases

Santos-Gomez 1990 US	residency performance - supervisor rating	PBL received higher average ratings by supervisors although not statistically significant (PBL: mean 4.12, SEM 0.07; conventional 3.95, 0.06; p=0.09)	Supervisors rated PBL students higher for 'attention to health care costs' and somewhat higher for 'communication with patients and patient education'. Trend for nurses to rate conventional students higher with respect to knowledge. The lack of differences on other performance measures may be attributable to the same clerkship experiences for the two groups in years 3 and 4.	Selection bias (PBL 75% response; conventional track 47% response) – response rates reported but not explicitly stated as a limitation
	residency performance – nurse rating	No significant difference for nurse ratings between PBL (mean 4.05, SEM 0.10) and conventional (4.12, 0.06); p=0.5		
Schmidt 1996 Netherlands	diagnostic performance	Significant interaction between curriculum type and student year: students from conventional program performed poorer than the other two programs; students from integrated curriculum performed significantly better than other two groups in years 2 and 3; year 5 and 6 students in PBL performed better than conventional but not integrated curriculum; students in integrated curriculum performed better than conventional curriculum in years 5 and 6	"Integration between basic and clinical sciences and an emphasis on patient problems may be the critical factors that determine superior diagnostic performance rather than whether a curriculum is self- or teacher-directed. PBL seems to live up to its expectations, but so does the integrated approach."	Selection bias (volunteers who received pay, represent 5% of total medical student population in The Netherlands – not explicitly listed as a limitation; cross-sectional design
Shin 1993 Canada	knowledge of management of hypertension	PBL graduates had higher mean scores overall vs. those from traditional school (68% vs. 62%, $p<0.01$); in multiple regression, medical school was the only variable that predicted overall score.	"The graduates of a problem-based, self-directed undergraduate curriculum are more up to date in knowledge of the management of hypertension than graduates of a traditional curriculum."	test scores may not reflect performance or patient outcomes; questionnaire was self-administered; difference was small and may not be clinically important; only one disorder; results may reflect different emphasis on the one disorder between schools; results could be attributed to admission process of PBL school

Tamblyn 2005 Canada	mammography screening rates	Rates changed significantly within PBL school pre- and post-curricular change (p=0.01); rates were significantly greater than one traditional school pre- (p=0.003) and post-change (p=0.002) but not significantly different for two other traditional schools	"Transition to a community oriented PBL curriculum was associated with significant improvements in preventive care and continuity of care and an improvement in indicators of diagnostic performance." "Although significant improvements were seen in three of four indicators, (authors) were not able to show that these changes were significantly larger than simultaneous changes in other schools, owing to a small number of graduates after transition period and low power."	use of administrative databases – could only assess "subset of outcomes that would be of importance in judging quality of primary care"; excluded salaried doctors; limited to one intervention school and first graduation class of new curriculum; unable to identify specific attributes of new curriculum that resulted in positive outcomes (e.g., school-level factors vs. curriculum)
	continuity of care ("proportion of all visits that were made to study doctor or generated by his or her referral among all visits made by his or her patients during the year")	Similar across schools pre-change; after transition, significant improvements were observed within the PBL school (p=0.007), but not compared to traditional schools (p=0.05; p=0.08; p=0.10)		
	prescribing disease-specific drugs vs. symptom-relief drugs	PBL school showed 32.2% (p=0.002) difference pre-post transition in rate of prescribing disease-specific minus symptom relief drugs; difference was significant for one of the three traditional schools (23.7%, p=0.001)		
	contraindicated prescribing	PBL school had higher rates of prescribing contraindicated drugs compared to others post-transition (significantly higher than one of the schools, p=0.01)		
Van Hessen 1990 Netherlands	test scores	Mean test scores significantly higher for recently graduated physicians vs. final year PBL students (difference 2.6%, $p<0.05$).	Mean scores were similar for recently graduated physicians from medical school with traditional curricula compared to final year students in a PBL curricula.	NS
Verhoeven 1998 Netherlands	total test scores	No significant differences overall. Mean scores also presented by year of program for PBL and non-PBL and for 2 sittings of exam; also for overall test and subtests (basic science, clinical science, social sciences); few differences overall (significant difference in year 6 for one sitting of exam and year 3 for another sitting of exam, both favouring non-PBL)	"No systematic differences were found on total test scores." Non-PBL tended to score better on basic sciences while PBL did better on social sciences, but this finding was inconsistent across test administrations. "The students master the same knowledge, it is only the moment in time that they learn it that differs."	non-PBL students were volunteers; test results for non-PBL students had no consequences but did for PBL students

Verwijnen 1990 Netherlands	test scores	Differences in mean scores between schools varied by test. By the end of year 6, differences in scores were very small and "all schools reached the level of scores obtained on all tests by the reference group of residents in general practice."	There are "no impressive differences" in factual knowledge between students from schools with PBL vs. conventional curricula. Authors noted that growth in knowledge appears constant through medical school, "irrespective of the structure of the curriculum, the educational principles employed, and possible school specificity of the tests."	some comparison groups were volunteers, not randomly selected; administration of tests at comparison schools was experimental and did not form part of the students' assessments; lack of standardization in administration of tests at different schools
Way 1999 US	USMLE Step 1 score	Results presented graphically; "when comparing student scores across pathways…there was essentially no difference above and beyond those attributable to admission variables" (predictors of performance: MCAT and GPA)	After controlling for predicted performance, students performed the same on the standardized exam.	potential contamination of groups (students from both tracks participate in some of the same courses including a board preparation course)

Whitfield 2002 US	fund of knowledge (FK) and clinical problem-solving skills (CPSS)	Scores always greater for PBL (significant in 4 of 6 specialties); multiple regression models adjusting for track, gender, race, year, age and GPA showed significant difference for FK (p=0.005) and CPSS (p=0.009); regression models adjusting for all variables as well as MCAT significant for FK (p=0.023) but not CPSS (p=0.06)	Effect sizes for fund of knowledge and clinical problem-solving skills were positive with varying magnitude (small to moderate) across years. Absolute true effect of PBL cannot be determined because of inability to control for all potential differences between self-selecting PBL and LBL students. Effect stronger in later years potentially due to improved program.	range of clerkship scores was small which makes "statistical and practical significance difficult to achieve"; clerkship scores are subjective and "only partly based on standardized testing"; students self-selected for curricular track; uncertain blinding of evaluators to student's curricular track
	USMLE 1	No significant difference between PBL (mean 208.06, SD 19.13) and LBL students (209.74, 20.79) (p=0.4)		
Woodward 1990 Canada	performance ratings by internship supervisors	PBL rated significantly better for overall performance ($p \leq 0.02$); proportion of interns rated weaker or much weaker than peers was lower for PBL (7% vs 11%); proportion of interns rated better or much better higher for PBL (70% vs 45%); for eight individual competencies direction favoured PBL but differences not always significant	Authors noted a "lack of distinctiveness in the profile of competencies."	NS
Woodward 1990 Canada	pattern of billings	Significant findings for: number of services provided (PBL mean 603.4, SD 407.8 vs. non-PBL 734.4, 518.2; $p \leq 0.01$); earnings (PBL mean \$10 195, SD \$5 688 vs. non PBL \$11 865, SD \$6 752; $p \leq 0.01$); number of patients seen (PBL mean 306, SD 187 vs. non-PBL 381, 219; $p \leq 0.01$); cost per patient (PBL mean \$45, SD 49 vs. non-PBL 35, 28; $p \leq 0.01$); number of services per patient (PBL mean 2.3, SD 1.8 vs non-PBL 2.0, 1.4; $p > 0.01 \leq 0.05$). No differences in controlled analyses for cost per service or number of encounters per patient. Significant differences found in four of fifteen service categories (PBL provided fewer emergency room services and minor assessments, but more psychotherapy and prenatal care, per 100 patients). No differences in frequency or cost of laboratory work, non-surgical diagnostic and therapeutic procedures, consultations from	PBL graduates spend "more time per patient encounter" and provide "more expensive services." Further, PBL graduates "provide somewhat more services per patient and tended to recall patients more frequently within a month. These differences appear to be best accounted for by their decreased use of minor assessments and increased provision of psychotherapy services to their patients."	multiple comparisons (authors chose lower cut-off for significance); limitations resulting from use of data collected for administrative purposes (e.g., unreliable diagnostic information, time spent in direct patient care); inability to confirm whether differences due medical school

	other physicians.	curricula, selection procedures or post-graduate training; no measures on quality or outcome of care

* results for randomized groups only
† results based on all students, not just randomized groups
‡ analyses adjusted for age group, gender, race, internal medicine clerkship rotation, internality, norm-favoring, self-realization, and achievement via Independence scales of the California Psychological Inventory, and total MCAT score
USMLE=United States Medical Licensing Examination; PBL=problem-based learning; SD=standard deviation; ES=effect size; CCX=Clinical Competency Examination; ITT=intention to treat; LBL=lecture-based learning; MCAT=Medical College Admission Test; NBME=National Board of Medical Examiners; NS=none stated; SEM=standard error of the mean; CI=confidence interval; GPA=grade point average; FK=fund of knowledge; CPSS=clinical problem-solving skills

Evidence Table 3. Description of studies evaluating course-based PBL

First author, year of publication, country	Study setting	Study design (n)	Active intervention	Passive comparison	Course content	Authors' primary outcome(s)
Casassus 1999 France	Bobigny Faculty of Medicine, Université Paris-Nord, France	cohort with concurrent and non-concurrent controls (NS)	**PBL** (n=68): students work through cases in small groups	**traditional methods** (n=NS)	hematology	knowledge
Dyke 2001 Australia	University of Western Australia Medical School	randomized trial (n=136)	**PBL** (n=40): students work in small groups with assigned tutor; groups meet 2x/week for up to 2 hours	**traditional** (n=96): 13 lectures (50 minutes each) and 10 tutorial sessions (max 90 minutes each)	epidemiology	academic assessment (first listed)
Eisenstaedt 1990 US	Temple University School of Medicine, Philadelphia, PA	prospective cohort (n=166)	**PBL** tutorial (n=59): focus on clinical vignettes; "some additional structure beyond that used in most PBL settings in the form of self-study questions and a bibliography. Each group met approximately 3 times per week for a 1- to 2-hour session, with additional time for independent study"	**traditional lecture-based** (n=107)	hematology-transfusion medicine segment of pathophysiology course	factors associated with participation in PBL
Hinduja 2005 England	two medical schools in England	prospective cohort (n=160)	**PBL** (n=80): no formal teaching or lectures; 120 minutes/week in dissection room; "integrated tutorials comprising pre-clinical and clinical disciplines last 60 minutes per week"	**traditional** dissection-based course (n=80): 90 minute lectures/week; 90 minute formal dissection teaching; small group tutorials x60 minutes/week	anatomy	knowledge on true/false questionnaire
Sivam 1995 US	Northwest Center for Medical Education,	**cohort with non-concurrent controls** (n=1674)	**PBL** (n=628): number of lectures is half of the traditional curriculum; case-based learning performed in small groups with tutorial sessions	**traditional** (n=1,046): "didactic lectures supplemented with some clinical correlation conferences and an essay-writing exercise on	pharmacology	performance on standard national board medical

	Indiana University School of Medicine, Gary, Indiana	and computer-based learning resources	a topic of recent advances in pharmacology”	licensure exams

PBL=problem-based learning; NS=not specified

Evidence Table 4. Results of studies evaluating course-based PBL

First author, year of publication, country	Outcome	Quantitative results	Authors' conclusions	Stated limitations
Casassus 1999 France	knowledge	No significant difference between PBL and traditional groups in mean % correct answers to multiple choice questions (35.6 vs. 34.5)	PBL students demonstrated better problem-solving skills but showed no difference in knowledge acquisition	comparison between heterogeneous groups who sat different exams; isolated experiment; potential for Hawthorne effect
	clinical and therapeutic files	PBL performed better compared to group taught concurrently with traditional methods as well as results obtained for previous remedial session (data presented graphically; $p<0.0013$)		
Dyke 2001 Australia	academic assessment – overall mark	No significant difference between PBL and conventional students: 73.5 (range 57,81) vs. 72.8 (45,85); 95% CI -1.58,3.08	Teaching method "does not affect conventional academic outcomes". PBL students had "richer learning experience."	students who didn't consent to participate (n=14) were put in the traditional group (not stated as limitation); possible contamination across groups
Eisenstaedt 1990 US	knowledge at conclusion of course	Significant difference favouring traditional (mean 80.21 [SE 1.65]) vs. PBL (66.53 [2.74]); $p<0.001$ (number of participants represents those completing exams at both conclusion of course and 2 years post-course)	Traditional students performed significantly better on multiple-choice exam at conclusion of course; retention of information at 2 years was no different between groups. Advantage from LBL may be short-lived and learning through PBL may result in better retention of information.	PBL one small segment of the curriculum, the rest remained the same; differential motivation between groups may partially explain results
	knowledge after 2 years	Significant decline in exam scores among control group from immediately after course to 2 year follow-up; no significant difference between groups at 2 year follow-up		
Hinduja 2005 England	mean score on knowledge test	Significant difference favouring traditional (mean 37.5/50, range 25-46, SD 3.85) vs. PBL (32.35/50, 18-45, 4.90); $p<0.001$	Students from traditional program performed better on multiple-choice test to assess anatomical knowledge. Greater range in results for PBL students suggests that stronger students may do equally well but weaker students may benefit from more structured learning.	generalizability; students were self-selected and may have been stronger students; confounders not considered such as amount of personal

				study time
Sivam 1995 US	mean score on NBME 1 or USMLE 1 compared to national average	"Two of the three traditional curriculum groups that took NBME I fared slightly below the national average in the pharmacology subject; the other group scored as well as the national average in the USMLE I. The latest three PBL curriculum groups that took USMLE I did as well as or better than the national average."	"PBL curriculum students are well prepared to perform the objective measure of evaluation in pharmacology."	NS

PBL=problem-based learning; CI=confidence interval; SE=standard error; SD=standard deviation; LBL=lecture-based learning; USMLE=United States Medical Licensing Examination; NBME=National Board of Medical Examiners; NS=none stated