

A Review of Longitudinal Community and Hospital Placements in Medical Education

TOPIC REVIEW GROUP

Professor JE Thistlethwaite, Professor of Medical Education, Centre for Medical Education Research and Scholarship, University of Queensland

Dr Emma Bartle, Lecturer in Medical Education, Centre for Medical Education Research and Scholarship, University of Queensland

Amy Ai Ling Chong, medical student, School of Medicine, the University of Queensland

Dr Marie-Louise Dick, Associate Professor in General Practice, The University of Queensland

Dr David King, Senior Lecturer, Discipline of General Practice, The University of Queensland

Dr Sarah Mahoney, Senior Lecturer, Flinders University, South Australia

Dr Tracey Papinczak, Faculty of Engineering, Architecture and Information Technology, The University of Queensland

Dr George Tucker, Clinical Director, Rural Clinical Schools, The University of Queensland

REVIEW UNDERTAKEN BETWEEN NOVEMBER 2011 AND OCTOBER 2012

CONTENTS

Abstract	2
Glossary of terminology and acronyms	3
Executive summary	4
Introduction	6
Review goals and questions	8
Method	8
Search strategies	8
Inclusion and exclusion criteria	9
Coding the papers	11
Data analysis.....	11
Realist evaluation	12
Findings	12
Classification of papers.....	12
Case studies.....	16
Findings answering the research questions	17
How are longitudinal placements defined? What are the lengths of such placements?.....	17
What learning outcomes/objectives are defined?	18
Are the placements effective? In what ways are these placements effective?	19
<i>Change in attitudes</i>	20
What factors influence effectiveness?	20
<i>Continuity of patients care</i>	21
<i>Participation</i>	21
<i>Continuity of mentorship</i>	22
<i>The learning environment and support (continuity of location)</i>	23
What other impacts do the placements have on clinicians, students, future career choice and patients?.....	23
<i>Clinicians</i>	23
<i>Students</i>	24
<i>Choice of career</i>	24
<i>Patients' feedback and impact</i>	25
How do longitudinal placements promote learning?.....	25
<i>Factors inhibiting learning</i>	25
<i>Student anxieties and concerns</i>	26
Discussion	27
Conclusion	28
Practice points	28
Recommendations for medical education	29
Limitations of the review	29
Recommendations for further research	29
Appendix 1: search syntax	30
Appendix 2: flow diagram of literature search and paper selection	31
Appendix 3: realist approach as applied to this review	32
Appendix 4: case studies	33
Appendix 5: table of papers	40
References	44
References in review	45

ABSTRACT

Background

Traditionally, clinical learning for medical students consists of short-term and opportunistic encounters with primarily acute-care patients, supervised by an array of clinician preceptors. In response to educational concerns, some medical schools have developed longitudinal placements rather than short-term rotations. Many of these longitudinal placements are also integrated across the core clinical disciplines, are commonly termed longitudinal integrated clerkships (LICs) and often situated in rural locations. This review aimed to explore, analyse and synthesise evidence relating to the effectiveness of longitudinal placements, for medical students in particular to determine which aspects are most critical to successful outcomes.

Method

Extensive search of the literature resulted in 1679 papers and abstracts being considered, with 53 papers ultimately being included for review. The review group coded these 53 papers according to standard BEME review guidelines. Specific information extracted included: data relating to effectiveness, the location of the study, number of students involved, format, length and description of placement, the learning outcomes, research design, the impact level for evaluation and the main evaluation methods and findings. We applied a realist approach to consider what works well for whom and under what circumstances.

Findings

The early LICs were all community-based immersion programs, situated in general practice and predominantly in rural settings. More recent LIC innovations were situated in tertiary-level specialist ambulatory care in urban settings. Not all placements were integrated across medical disciplines but were longitudinal in relation to location, patient base and/or supervision. Twenty-four papers focussed on one of four programs from different viewpoints. Most evaluations were student opinion (survey, interview, focus group) and/or student assessment results. Placements varied from one half day per week for 6 months through to full time immersion for more than 12 months.

The predominant mechanism relating to factors influencing effectiveness was continuity of one or more of: patient care, supervision and mentorship, peer group and location. The success of LICs and participation satisfaction depended on the preparation of both students and clinical supervisors, and the level of support each received from their academic institutions.

Conclusion

Longitudinal placements, including longitudinal integrated placements, are gaining in popularity as an alternative to traditional block rotations. Although relatively few established LICs currently exist, medical schools may look for ways to incorporate some of the principles of LICs more generally in their clinical education programmes. Further research is required to ascertain the optimum length of time for placements depending on the defined learning outcomes and timing within the programme, which students are most likely to benefit and the effects of context such as location and type of integration.

GLOSSARY OF TERMINOLOGY AND ACRONYMS

Term	Acronym	Definition
Attachment		Location, supervisor or team, and time for an activity, which a medical student attends for learning purposes. Equivalent to 'placement'
Block clerkship	BC	Time spent by medical students in a placement or attachment focussing on one specialty discipline, and usually followed sequentially by similar placements in other disciplines. Usually not longer than 8 weeks duration.
Block rotation	BR	Same as block clerkship
Clerkship		Time spent by medical students in a placement or attachment for experiential clinical learning
Continuity		(In the context of this paper) used more or less synonymously with 'longitudinal'. Students have repeated, on-going contact with a clinician or clinicians /clinical team and their patient bases
Community Partnership Program	CPP	Longitudinal experience in rural Tennessee
Harvard Medical School	HMS	
in-training evaluation reports	ITER	Reports providing assessment of student progress during clinical training
Immersion		Deep experience to exclusion of most other activities
Integrated		(In clinical education) Combining components of the medical course so that they are learned concurrently. Most often refers to either: biomedical sciences being integrated with clinical learning; or, clinical learning of all core specialist disciplines
Longitudinal		Involves a regular, recurrent placement in the same setting with the same preceptor (and thus with access to the same patient base) over a period of time.
Longitudinal integrated clerkship	LIC	Time spent by medical students in a regular, recurrent placement in the same setting with the same preceptor for experiential clinical learning of all core specialist disciplines concurrently
Mentor		(Often used interchangeably with supervisor or preceptor) person who guides, oversees and teaches medical students, particularly in clinical settings.
Objective structured clinical examination	OSCE	Practical examination of clinical skills widely used for formal assessment purposes in medical schools
Placement		Location, supervisor or team, and time for an activity, which a medical student attends for learning purposes. Equivalent to 'attachment'
Patient practitioner orientation scale	PPOS	A tool used to measure attitudes to patient-centred care
Parallel rural community curriculum	PRCC	Rural LIC program at Flinders University South Australia
Preceptor		(Often used interchangeably with supervisor or mentor) person who guides, oversees and teaches medical students, particularly in clinical settings
Rotation		Time spent by medical students in a placement or attachment focussing on one specialty discipline, and usually followed sequentially by similar placements in other disciplines (hence the rotation of placements).

		Usually not longer than 8 weeks duration (often used synonymously with block clerkship / block rotation)
Rotation-based clerkship	RBC	Time spent by medical students in a placement or attachment for experiential clinical learning, structured as rotations
Rural physician associate program	RPAP	Rural LIC at University of Minnesota
Rural clinical school	RCS	Clinical school (for training in health professions) based in a rural area
Supervisor		(Often used interchangeably with mentor or preceptor) person who guides, oversees and teaches medical students, particularly in clinical settings
Tasks of medicine scale	TOMS	Tool used to assess prioritization of biomedical and psychosocial tasks in a patient encounter
Traditional track		Refers to the rotation-based clerkship which has long been, and still is, the commonest structure for clinical learning in medical schools
Topic review group	TRG	The group of academics who undertook this BEME review
Tutor		(Often used interchangeably with mentor, supervisor or preceptor) person who guides, oversees and teaches medical students, particularly in clinical settings. Often refers to person who works with a small group of students rather than one-to-one

EXECUTIVE SUMMARY

Background

Longitudinal integrated clerkships (LICs) have existed since at least 1971 when the Rural Physician Associate Program (RPAP) began in Minnesota in USA. Although a number of medical schools in different parts of the world have since developed longitudinal integrated clerkships in one form or another, they have usually consisted of small programs, separate from, rather than integral to, mainstream medical education. Traditionally, clinical learning for medical students consists of short-term and opportunistic encounters with primarily acute-care patients, supervised by an array of clinician preceptors. This model is increasingly being criticised as poorly aligned to current educational theory and society's current and predicted health-care needs (Hauer et al., 2009; Hirsh et al., 2007). In response to these concerns some medical schools have developed longitudinal placements that are often integrated placements rather than short-term discipline-specific rotations.

This review aimed to explore, analyse and synthesise evidence relating to the effectiveness of longitudinal placements, including LICs, and to summarise the information currently available from reports and research into longitudinal clinical education. The review examined the nature and scope of programs and factors associated with successful outcomes.

Method

Extensive search of the literature resulted in 1679 papers and abstracts being considered, with 53 papers that met the inclusion criteria ultimately being included for review.

Inclusion criteria were: that the placements described were for medical students, had a duration of at least 13 weeks, the student had continuity of patient contact in same location and/or continuity of preceptor, that evaluation data was presented, and that the paper was in English. Commentaries or opinion pieces were not included.

Specific information extracted included: data relating to effectiveness, the location of the study, number of students involved, format, length and description of placement, the learning outcomes, research design, the impact level for evaluation according to Kirkpatrick's levels (Kirkpatrick, 1967) and the main evaluation methods and findings. Reviewers rated the strength of findings and scored their overall impressions as is standard on the BEME coding sheet. We applied a realist approach to consider what worked well for whom and under what circumstances.

Main findings

There were a number of different models for longitudinal placements. Early longitudinal placements were all community-based, integrated, immersion programs, situated in general practice and predominantly in rural settings. Some more recent innovations were situated in tertiary-level ambulatory care in urban settings.

- The majority of papers came from either North America (34) or Australasia (14).
- Twenty-four papers focussed on one of four programs from different viewpoints.
- Kirkpatrick levels were rated predominantly 1 (student reaction) or 2b (change in knowledge).
- Strength of findings was rated 3 (conclusions can probably be based on the results) in 31 cases, 4 (results are clear and very likely to be true) in 15 cases and the remainder (7) were rated 2 (results weak/ambiguous but there appears to be a trend).
- Four papers were published prior to 1990, eight were published between 1990-1999, and the remaining 41 were published since 2000.
- Thirty-three of the papers involved evaluation of fewer than 50 students, and the total number of students accommodated by most of the full-time, year-long programs in any one year was generally small (8 – 10 students). Some of these programs were able to grow by having multiple small student cohorts in geographically diverse settings.
- Most evaluations were student opinion (survey, interview, focus group) and/or student assessment results. The majority (37) were conducted post placement.
- Placements varied from one half-day per week for 6 months through to full-time immersion for more than 12 months.
- Senior, experienced clinicians provided the clinical supervision in almost all cases.
- Most of the placements were longitudinal integrated clerkships. A small number of longitudinal placements were discipline-specific.
- Outcomes assessed included humanism in patient care and doctor-patient relationships, understanding of needs of rural and underserved communities, and orientation to rural and primary care career choices.
- A number of studies demonstrated approximate equivalence in overall assessment performance for LIC students compared to students on traditional rotations.

The predominant mechanism reported relating to factors influencing effectiveness was continuity in some form. The ideal was continuity of patient care, but it was recognized that continuity of supervision and mentorship, peer group and location were also significant factors. Senior, experienced clinicians provided the clinical supervision in almost all cases. The papers

highlighted other factors such as participation and authenticity of student role as a member of the team as influencing effectiveness for student learning. However some students were less well suited to the self-directed style of learning required in these programs, and some felt that the workload was too high. Students also had some anxieties about perceived gaps in learning opportunities, particularly when comparing their experiences to that of their peers in traditional block rotations.

Conclusion

Longitudinal placements, including LICs, are gaining in popularity as an alternative to traditional block rotations in medical education. Although there are relatively few LICs, and most full-time LICs cater for only small numbers of students, particularly in urban areas, many medical schools are looking for ways to incorporate some of the principles of LICs more generally in their clinical education programs. The underlying mechanism for the benefits attributed to longitudinal placements is overwhelmingly that of continuity in its varying forms of patient, mentor and location longitudinal exposure. This results in the development of trust, and long-term professional relationships between preceptors and students, and between patients and students. Students become members of a community of practice with attendant responsibility depending on their competency and seniority. They gain an enhanced understanding of the impact of health and illness on individuals and on families and communities, and develop rapport, compassion and caring approaches.

The success of longitudinal placements depends on the preparation of both students and clinical supervisors, and the level of support each receives from their academic institutions.

Introduction

A core component of medical education is development of clinical experience. Students are expected to learn through practical integration and application of university-acquired knowledge and skills to professional medical practice during interactions with patients in clinical settings. This experiential learning is traditionally organised as clinical clerkships or core clinical rotations involving short-term placements (typically full-time for four to eight weeks) through various specialties in hospital-based and general practice settings.

This traditional educational paradigm of short-term and opportunistic encounters with primarily acute-care patients, with students being supervised by an array of clinician preceptors, is increasingly being criticised as poorly aligned to modern educational theory and society's current and predicted health-care needs (Hauer et al., 2009; Hirsh et al., 2007). In response to these concerns some medical schools have developed longitudinal placements rather than short-term rotations. In countries with a large rural and remote population base, such longitudinal attachments are commonly in rural centres and involve rural hospitals, general practices and/or other community settings, and are integrated across the core clinical disciplines. There have been suggestions, dating back several decades as well as more recently, that more of the medical curriculum should be embedded in the community in this manner (Hart, 1985; Oswald, 1989; Thistlethwaite et al., 2007). Early patient contact is also predominantly community-based (Dornan et al., 2005) and may be longitudinal over several months. Such contact is often not opportunistic in that patients are invited by GPs/family physicians to interact with students for specific activities. Globally various longer attachments have been piloted, and some sustained for more than a decade such as in Minnesota and South Australia (detailed in appendix 4).

These community-based programs have shown great promise in providing opportunities for student learning in relation to health promotion and disease prevention in community settings, as well as demonstrating the value of involvement in continuity of care and the management of patients with chronic illness. They expand clinical placements thus helping to address the challenge of managing increasing student numbers (Thistlethwaite et al., 2007). As early as 1971 the rural physician associate program (RPAP) was established in Minnesota (USA) as a 36-week, longitudinal immersion learning experience for third year medical students (Zink et al., 2008). In 1993 the Cambridge Community-based Clinical Course (UK) piloted a scheme in which medical students undertook a 15-month general practice immersion placement (Oswald et al., 1995; Oswald et al., 2001), while in Australia, a 12-month parallel rural community curriculum has been in place for some years at Flinders University (Worley et al., 2000).

During 2009 a survey of 'longitudinal integrated clinical (LIC) clerkships' was conducted in the USA, with responses from 15 institutions that had active LIC clerkships in place (Norris, et al., 2009); these clinical placements included community-based clerkships and hospital ambulatory rotations of five months or longer. Only one program was hospital in-patient-based while the majority were situated in rural environments. These clerkships are also referred to as continuity-based, with the provision that: 'continuity-based clinical medical education requires that the student stay in one place, with one set of faculty members and one groups of patients, for an extended period of time' (Norris et al., 2009, p. 902).

The only previous review of longitudinal placements focussed on ambulatory care rotations with a search of the literature from 1996-2000 (Ogrinc et al., 2002). The authors included only seven papers, all originating from the USA. Two of these related to placements that did not fit our inclusion criteria for length of experience while the remaining five did and are included in this review.

At the start of the BEME review in November 2011 there had not been any other systematic review of the number and effectiveness of these programs across different locations. As the University of Queensland School of Medicine was introducing a pilot longitudinal general practice-based rotation for year 3 graduate entry medical students we decided to look at the different models of longitudinal placements and the evidence for effectiveness. However, just after submission of this review to BEME in October 2012, Walters et al. (2012) published a review specifically of the outcomes of longitudinal integrated clerkships. We compare our findings to this work but have not included the paper in the systematic review itself due to the timing.

The nomenclature and definitions of 'longitudinal', 'integrated' and 'continuity' vary in the literature and therefore one of the aims of our review is to clarify the terminology. There is no agreed duration that is required for a program to be termed 'longitudinal'. A program may be called longitudinal if it involves a regular, recurrent placement in the same setting with the same preceptor (and thus with access to the same patient base) over a period of time. A longitudinal program may also be an 'immersion' experience particularly when it involves full time placement. 'Immersion' often implies exclusive engagement in a setting or activity, but the term has also been used in relation to intermittent but recurrent, long-term attachments.

'Continuity' is used more or less synonymously with 'longitudinal' and the intent of 'longitudinal' programs is that they afford 'continuity' of some type. It is likely that, since all the early longitudinal programs were based in rural general practice, one of the early visible attributes was the opportunity for students to see the same patient population over the period of the

program, and therefore potentially derive benefits in learning that paralleled the known benefits to patients of continuity of care. As longitudinal clinical programs have developed further, other 'continuities' have become apparent, namely continuity of relationships with preceptors, and continuity with a small group of co-learners. These continuities occur almost by default in rural settings, since opportunities for more diverse relationships are limited by geography. Continuity has a number of meanings depending on its application. Hadac et al. (1979) in their paper 'Can continuity of medical care be taught' quote 'one of the most frequently cited definitions of continuity' (p. 525), that of Shortell: '...the extent to which medical care services are received as a coordinated and uninterrupted succession of events consistent with the medical care needs of the patients' (Shortell, 1976, as quoted in Hadac et al., 1979, p. 526).

Some longitudinal programs are also 'integrated'. In this context, integrated is usually used to describe the learning of all core clinical disciplines simultaneously. Thus in Walters et al.'s (2012) review, a definition of LICs developed by the International Consortium of Longitudinal Integrated Clerkships (CLIC) was that students in LICs: participate in the provision of comprehensive care of patients over time; participate in continuing learning relationships with these patients' clinicians; meet the majority of core clinical competencies across multiple disciplines simultaneously through these experiences. Walters' review excluded programs of less than 6 months, implying that LICs by definition must also have a duration of at least six months.

In some contexts, however, particularly in programs placed in the early years of medical education, 'integrated' may be used to describe learning that integrates the basic medical sciences with clinical experience and understanding.

Review goals and questions

The goals of this review were to explore, analyse and synthesise evidence relating to the effectiveness of longitudinal placements, including LICs, as a means of achieving learning outcomes in medical student pre-qualification training programmes and, if effective, how these placements achieve this result.

Questions

- How are longitudinal placements defined?
- What are the lengths of such placements?
- What learning outcomes/objectives are defined?
- Are the placements effective?
- In what ways are these placements effective?
- What factors influence effectiveness?
- What other impact do these placements have on clinicians, students and patients?
- How do they promote learning?

Method

Search strategies

AC, a second year medical student (graduate entry), under the supervision of an information scientist in the University of Queensland School of Medicine library carried out the PubMed search between 22 November and 13 December 2011. CINAHL, EMBASE, Medline and Web of Knowledge (WoK) were searched over January and February 2012 (see appendix 1 for search

syntax). The initial search was run over Medline and PubMed concurrently, and then on the remaining databases. The citations and abstracts (n= 1679) were imported into EndNote X4.

After defining the topic, summarising it into one sentence, identifying the key concepts, listing the possible synonyms or related terms for those concepts, we then combined the key terms or concepts in our search using Boolean searching (AND and OR). PubMed uses MeSH (Medical Subject Headings) database, a controlled vocabulary used for indexing articles in PubMed which provides a consistent way to retrieve information that may use different terminology for the same concepts. We used PubMed as our main database and adapted to other databases by using the same search terms with or without the phrase “ “. Use of phrase searching turned off the intelligent search engine, which automatically mapped our term to MeSH. Only Medline did not provide phrase searching, while PubMed, CINAHL, Embase and WoK allowed phrase searching. Both controlled and uncontrolled vocabularies were brainstormed together as a group.

One member of the TRG also scanned the table of contents of three of the major medical education journals (Academic Medicine, Medical Education and Medical Teacher) for all of 2011 (to check the accuracy of the search strategy) and for the first six editions of the 2012 volumes (to ensure the search was up-to-date).

Inclusion and exclusion criteria

Prior to reading the abstracts, the TRG (topic review group) defined the inclusion and exclusion criteria for the study (table 1). There was considerable discussion at this stage about the definition of longitudinal and this was further refined during our reading of papers and subsequent coding. We eventually defined longitudinal for the purposes of this review as any student activity that takes place for more than 13 consecutive weeks and involves either the student interacting with the same patient population and/or with the same clinical mentor/preceptor to ensure continuity of experience. We nominated the 13-week duration cut-off point as this duration made a clear differentiation from the traditional single-discipline rotations (usually approximately 8 weeks in duration); it also allowed for the incorporation of the vast majority of papers reporting longitudinal placements. By same patient population we mean that students were able to see the same patients on different occasions if the patients were followed-up in either hospital or community settings. Frequency of clinical activities in the placement during this time may range from daily to monthly.

We excluded such learning experiences as family studies which focussed only on one patient or family over a length of time. These experiences, while often very valuable, are not longitudinal placements in that there is no relationship between the unique learning task and the overall clinical learning program, but are rather stand-alone activities surrounded by whatever program the student is undertaking.

Table 1: Inclusion and exclusion criteria

INCLUSION	EXCLUSION
<ul style="list-style-type: none"> • Medical students • Placements longer than 13 weeks • Access to and continuity of patient population in same location and/or continuity of tutor/preceptor/mentor • Evaluation data relating to effectiveness of placements • In English 	<ul style="list-style-type: none"> • Health professional students other than medical • Placements of less than 13 weeks • No continuity of patients and location and/or tutors • No evaluation data • Not in English • Commentaries or opinion pieces

AC and one other TRG member reviewed all the abstracts independently and decided on relevance. Each member was allocated a proportion of the abstracts to ensure a team effort. After discussion and reaching a consensus, the reviewers agreed to keep 85 abstracts (for full retrieval) and to discard the remainder. Each full paper was read independently by two TRG members and then discussed to reach agreement about whether the paper should be retained and coded. The reference list in each retained full paper was scanned for other papers that might be relevant; these were read and included as appropriate (listed as journal scan in table 2). The number of citations for each database is shown in table 2 and a flow chart of the literature search and paper selection is provided in appendix 2. Apart from one paper that was included twice in PubMed, the duplicates are papers found in more than one database with PubMed having been searched first as the index database searched first. Of the 16 journal scan papers, 13 were found from the reference lists of other included papers and three from scanning the 2012 journals. There were no additional papers from 2011 that had not already been identified by our search strategy.

Table 2: Citations from databases

Database	Total number of abstracts	Number of abstracts excluded	Duplicates	Number of full papers retrieved	Number of full papers excluded	Number of full papers coded
PubMed	441	369	1	71	39	32
Medline	201	151	50	0	0	0
CINAHL	97	78	15	4	3	1
EMBASE	512	444	63	5	3	2
WoK (Web of Knowledge)	428	349	72	7	5	2
TOTAL	1679	1391	201	87	50	37
Journal scan						16
TOTAL PAPERS IN REVIEW						53

The main reason for exclusion at the abstract stage was that the papers were opinion, description or minor reviews without empirical data. At the full paper stage, papers were rejected because they did not fit the inclusion criteria of a student placement spanning more than 13 weeks duration with any of: continuity of patient population, continuity of supervision or both, and at the same location. This ruled out papers reporting on rural clinical schools where students undertook discrete specialty-specific rotations without integration across those specialties, with the majority of such rotations lasting eight weeks or fewer. The final 53 papers include data relating to effectiveness.

Coding the papers

All full papers for inclusion were entered into the coding stage and filed under *coded* in the relevant EndNote library. We modified the standard BEME coding sheet to suit our research questions. After two papers were coded by all members of the TRG group independently and discussed, the coding sheet was further amended with agreement on a preliminary definition of longitudinal. In version 1, the standard fields included: title, date and authors of paper, the coder and date of coding, compliance with the inclusion criteria of focusing on longitudinal community and/or hospital placements for medical students with data relating to effectiveness, the location of the study, number of students involved, format, length and description of placement, the learning outcomes, research design, the impact level for evaluation according to Kirkpatrick's hierarchy (Kirkpatrick, 1967) and the main evaluation findings. Coders were asked to answer the review questions listed above from the content of the paper. Each paper was also assessed on its evaluation methods, including: strength of findings (1 to 5) and overall impression (poor to excellent) as is standard on the BEME coding sheet. In version 2 we refined the inclusion criteria of longitudinal to mean facilitation and/or access to the same patient population (i.e. continuity) for more than 13 weeks consecutive (but not necessarily full) weeks.

Criteria for judging the papers were more global than specific but included:

- Number of participants
- Number of cohorts (i.e. year groups)
- Whether there was any comparison of cohorts including historical, Kirkpatrick level 2 or above
- Nature of outcome data
- Inclusion of pre and post intervention data
- Any attempt at exploring how the placements were effective
- Clear description of analytical method.

As a preliminary inter-rater agreement exercise, all eight members of the TRG independently coded three papers. There was consistency in the reviewers' assessment of strength of findings and overall impression. Paper 1: all reviewers except one rated this as 1/poor, with one reviewer rating it as not filling the inclusion criteria. Paper 2: four reviewers rated this as either 1 or 2 and poor, while three reviewers rated it as 1, 2 or 3 and acceptable. Unanimous consensus after discussion the rated this paper as poor, 2. Paper 3: all reviewers rejected this as having no evaluation data.

Papers were shared out and pairs of TRG members coded each paper independently and any discrepancy in coding was discussed between the two to reach a consensus. The final number of coded papers for inclusion was 53.

Data analysis

A summary of each completed coding form was entered into an Excel spread sheet. We used the outcomes-focussed evaluation of the Kirkpatrick framework (Kirkpatrick, 1967), as modified by the Joint Evaluation Team (Barr et al., 2005), to begin with. This framework helps compare differences between learners, which is particularly important to indicate whether a new educational activity has any disadvantages in relation to assessment results, as students and staff may be particularly concerned about innovations that are not applied across a whole year cohort. However outcomes-based evaluation using Kirkpatrick's levels has limitations (Yardley & Dornan, 2012) and many of the papers included here have also included process evaluation, mainly through qualitative research approaches (see table 7). Moreover not all papers include comparison data (see appendix 5 for more details).

Realist evaluation

Our analysis of the data was further informed by the realist evaluation approach of Pawson & Tilley (1997); an approach that aims to answer the question ‘what works for whom in what circumstances?’ (p. 85). This approach moves beyond the outcomes-based evaluation of the Kirkpatrick framework, which considers whether an intervention is effective, to explore the factors impacting on any effectiveness. Furthermore, when used as a method of synthesising research, it focuses on explanation rather than judgment, recognising that interventions are usually complex rather than linear (Pawson et al., 2005). In the case of longitudinal placements, similar interventions may have varying degrees of efficacy of learning depending on the context: the learning and organisational environment, the length of placement, and the people (learners, tutors, patients etc.) involved. The context is thus about whether the right conditions are in place for learning. The mechanism by which change occurs (and in relation to this BEME review, the change is mainly about learning) may include: learners’ and teachers/preceptors’ attitudes, values and motivation; the nature and type of patient interactions; the capabilities/competencies the students derive from their participation, and/or resources. Thus: context + mechanism = outcome (also abbreviated to C-M-O). The realist approach does not aim to compare the outcomes of learners, who have had an educational intervention, with a control group who have not, but rather to explore the reasons for any changes in terms of the learning process. A realist review as defined by Pawson et al. (2005) has five steps and we consider these and how this BEME review has followed the steps in appendix 3.

None of the papers we reviewed adopted a realist approach explicitly in the evaluation though there was an underlying theme through the majority that continuity of learning experience would have positive effects on learners. To achieve such continuity was, of course, the reason for developing the placements. However by our systematic review through the BEME approach, we may theorise about possible mechanisms affecting the effectiveness of longitudinal placements and how these are influenced (or triggered) by the context in which the learners are learning. Such theoretical explanations are described as middle-range theories that permit further empirical testing. A realist review, such as this one, is ‘an interpretive theory-driven narrative summary which applies realist philosophy to the synthesis of findings from primary studies’ (Wong et al., 2012, p.93) to answer the defined research questions.

Findings

Classifications of papers

Based on information presented in the included papers, the following are presented in tables: the number of papers by categorisation, including: Kirkpatrick level (many papers included evaluation data at more than one level), strength of findings and overall impression of the paper as coded by two reviewers after consensus was reached (table 3); geographical location of the study (table 4); year of publication (table 5); number of students in the studies who undertook the placement and were included in the evaluation – this does not include any comparison group (table 6); types of evaluation data collection (table 7) and study design (table 8).

Table 3: Coding of included papers

Outcome evaluation	Rating level	Number of papers
Kirkpatrick level	1 - student reaction	37
	1/2a - student career intention	4
	2a - change in attitude	6
	2b - change in knowledge	24
	3 - change in behaviour	0
	4 - patient impact	1
	(Preceptor reaction)	1
Strength of findings	1: No clear conclusions can be drawn. Not significant	0
	2: Results weak/ambiguous but there appears to be a trend	7
	3: Conclusions can probably be based on the results.	31
	4: Results are clear and very likely to be true.	15
	5: Results are unequivocal	0
Overall impression	Poor	3
	Acceptable	27
	Good	19
	Excellent	4

Table 4: Geographical location

LOCATION	NUMBER (n = 53)
UK	2
Other Europe	1
Asia	1
North America	34
N America + Australia	1
Africa	0
Australasia	14

Table 5: Year of publication

YEAR	NUMBER (n = 53)
2012	3
2011	6
2010	5
2009	5
2008	7
2007	3
2006	3
2005	2
2004	3
2003	0
2002	0
2001	2
2000	2
1990-1999	8
Pre-1990	4

The number of students in the papers varied markedly and could include students undertaking the longitudinal placement, then a subset of these who were involved in the evaluation (not necessarily the response rate as not all students were asked to take part in interviews and focus groups) and, for some studies, a comparison group or groups (see appendix 5). The comparison groups were either students undertaking more traditional specialty attachments at the same medical school, or students from a different medical school. As none of the studies involved any randomisation, we are referring to these groups as comparisons rather than controls. In table 6 the number of students refers to the number for whom evaluation data are given rather than the number who undertook the placements. Also, some institutions have published several papers relating to the same longitudinal program, focusing on different process and outcome variables, and different numbers / subsets / cohorts of students.

Table 6: Number of student participants in papers

TOTAL NUMBER undertaking placements and involved in evaluation	NUMBER of papers (n =52)*	PAPERS
Fewer than 50	33	Bell et al., 2008; Berger & Schaffer, 1986; Couper & Worley, 2010; Denz-Penhey et al., 2004; Denz-Penhey & Murdoch, 2008; Fratarelli & Kamemoto, 2004; Geyman et al., 1984; Hadac et al., 1979; Hamilton et al., 1998; Hauer et al., 2012; Hirsh et al., 2012; Konkin & Suddards, 2011; Lewin et al., 1999; McLaughlin et al., 2011; Mihalynuk et al. 2008; O'Brien et al., 2012; Ogur et al., 2007; Ogur & Hirsh, 2009; Oswald et al., 2001; Poncelet et al., 2011; Poole et al. 2010; Power et al., 2006; Ringdahl et al., 2009; Schauer & Schieve, 2006; Stagg & al, 2009; Walmsley et al., 2009; Walters et al., 2011; Wilson & Cleland, 2008; Worley & Lines, 1999; Worley et al., 2000; 2004; 2006; Zink et al., 2010b
51-100	8	Couper et al., 2011; Denz-Penhey & Murdoch, 2009; Denz-Penhey et al., 2004; Florence et al., 2007; Henley et al., 2000; Herold et al., 1993; von Below et al., 2008; Zink et al., 2008
101-200	4	Carney et al., 2005; Deterding et al., 1999; Lubetkin et al. 1999; Peters et al., 2001
Over 200	6	Denz-Penhey et al. 2010; Halaas, et al., 2008; Prislín et al., 1998 (across 5 institutions); Verby et al., 1988; Wee et al., 2011; Zink et al., 2010a
Not given/unclear	1	Lemon et al., 1995

Note: Teherani et al. (2009) includes data from preceptors only

Table 7: Methods of evaluation data collection

TYPE	NUMBER (n =53)	PAPERS
Student surveys	22	Bell et al., 2008; Berger & Schaffer, 1986; Denz-Penhey et al., 2004; Hadac et al., 1979; Hamilton et al., 1998; Henley et al., 2000; Hirsh et al., 2012; Lemon et al., 1995; Lewin et al., 1999; Lubetkin et al., 1999; Ogur et al., 2007; Peters et al., 2001; Poncelet et al., 2011; Poole et al. 2010; Prislín et al., 1998; Ringdahl et al., 2009; Schauer & Schieve, 2006; Stagg & al, 2009; Verby 1998; von Below et al., 2008; Wee et al., 2011; Wilson & Cleland, 2008
Student 1:1 interviews	14	Couper et al., 2011; Denz-Penhey et al., 2004, 2005; 2008; 2009; 2010; Hamilton et al., 1998; Hauer et al., 2012; Konkin & Suddards, 2011; Oswald et al., 2001; Walmsley et al., 2009; Walters et al., 2011; Worley et al., 2000; Worley et al., 2006
Change in students' attitudes	6	Bell et al., 2008; Florence et al., 2007; Hirsh et al., 2012; Konkin & Suddards, 2011; Ogur & Hirsh, 2009; Ogur et al., 2007
Student focus/small groups	9	Bell et al., 2008; Couper et al., 2010, 2011; Lemon et al., 1995; Mihalynuk et al., 2011; Ogur et al., 2007; Peters et al., 2001; Poncelet et al., 2011; Wilson & Cleland, 2008
Direct observation of students in clinical environment	2	O'Brien et al., 2012; Peters et al., 2001
Student narratives/essays/case reports	3	Geyman et al., 1984; Ogur & Hirsh, 2009; Zink et al., 2008
Analysis of student performance during OSCEs	1	Zink et al., 2010b
Graduate surveys or career choice data	4	Florence et al., 2007; Halaas et al., 2008; Herold et al., 1993; Stagg & al, 2009
Staff interviews	8	Couper et al. 2010; 2011; Denz-Penhey et al., 2004; Denz-Penhey & Murdoch, 2009; Peters et al., 2001; Tehereni et al., 2009; Walters et al., 2011; Worley et al., 2000;
Staff surveys	6	Lemon et al., 1995; Lubetkin et al., 1999; Ogur et al., 2007; Poole et al. 2010; Power et al., 2006; von Below et al., 2008;
Patient surveys/interviews	2	Berger & Schaffer, 1986; Wee et al., 2011;
Student assessments (not necessarily comparison data)	24	Bell et al., 2008, Carney et al., 2005; Denz-Penhey & Murdoch, 2010; Deterding et al., 1999; Fratarelli & Kamemoto, 2004; Hirsh et al., 2012; Lewin et al., 1999; Lubetkin et al., 1999; McLaughlin et al., 2011; Ogur et al., 2007; Oswald et al., 2001; Poncelet et al., 2011; Poole et al. 2010; Power et al., 2006; Ringdahl et al., 2009; Schauer & Schieve, 2006; Verby 1998; Wee et al., 2011; Wilson & Cleland, 2008; Worley & Lines, 1999; Worley et al., 2000; Worley et al., 2004; Zink et al., 2010a, 2010b

Table 8: Research design

Papers with comparison data	24	Bell et al., 2008; Denz-Penhey & Murdoch, 2010; Florence et al., 2007; Fratarelli & Kamemoto, 2004; Hauer et al., 2012; Herold et al., 1993; Hirsh et al., 2012; Lewin et al., 1999; McLaughlin et al., 2011; Mihalyuk et al., 2008; Ogur et al., 2007; Oswald et al., 2001; Poncelet et al., 2011; Poole et al. 2010; Power et al., 2006; Ringdahl et al., 2009; Schauer & Schieve, 2006; Verby, 1998; Wilson & Cleland, 2008; Worley & Lines, 1999; Worley et al., 2000; 2004; 2006; Zink et al., 2010a
Data collected pre and post placements	3	Bell et al., 2008; Hirsh et al., 2012; Wilson & Cleland, 2008
Data collected during/mid and post placements	13	Denz-Penhey et al., 2004, 2005; Denz-Penhey & Murdoch, 2009; Hauer et al., 2012; Lubetkin et al., 1999; Mihalyuk et al., 2008; O'Brien et al., 2012; Ogur et al., 2007; Oswald et al., 2001; Poncelet et al., 2011; Walters et al., 2011; Worley et al., 2000, 2006
Data collected post placements only	37	Berger & Schaffer, 1986; Carney et al., 2005; Couper et al., 2011; Couper & Worley, 2010; Denz-Penhey & Murdoch, 2008, 2010; Deterding, et al., 1999; Florence et al., 2007; Fratarelli & Kamemoto, 2004; Geyman et al., 1984; Hadac et al., 1979; Halaas et al., 2008; Hamilton et al., 1998; Henley et al., 2000; Herold et al., 1993; Konkin & Suddards, 2011; Lemon et al., 1995; Lewin et al., 1999; McLaughlin et al., 2011; Ogur & Hirsh, 2009; Peters et al., 2001; Poole et al. 2010; Power et al., 2006; Prislun et al., 1998; Ringdahl et al., 2009; Schauer & Schieve, 2006; Stagg & al, 2009; Tehereni et al., 2009; Verby, 1998; von Below et al., 2008; Walmsley et al., 2009; Wee et al., 2011; Worley & Lines, 1999; Worley et al., 2004; Zink et al., 2008; 2010a; 2010b
Data collected more than 3 months following the placements	9	Denz-Penhey & Murdoch, 2008; Florence et al., 2007; Halaas et al., 2008; Herold et al., 1993; Mihalyuk et al., 2008; Ringdahl et al., 2009; Stagg & al, 2009; Verby, 1998; Worley et al., 2006

Of note is that except for the placements described in Carney et al. (2005), Deterding et al. (1999), Hamilton et al. (1998), Henley et al. (2000), Lubetkin et al. (1999), Peters et al. (2001), von Below et al. (2008) and Walmsley et al. (2009), students undertaking these longitudinal placements are volunteers, though most have to go through a selection process if the programme is oversubscribed. Poncelet et al. (2011) use a lottery to select from volunteers.

Case studies

Of the 53 papers, 24 related to one of four programs. These four clusters of articles focused on the same longitudinal attachments, but reported on different cohorts, using different methods of evaluation (e.g., interview, focus groups), or presenting findings at different Kirkpatrick levels; therefore in the above tables above they have been treated separately. However we have also grouped them together into case studies: table 9 gives the case study name and relevant papers. All of the four are members of the Consortium of Longitudinal Integrated Clerkships as detailed by Norris et al. (2009). The case studies are presented in detail in appendix 4.

Table 9: Case studies (placements with multiple papers)

CASE STUDY NAME	NUMBER OF PAPERS	REFERENCES
UWA: University of Western Australia rural clinical school	5	Denz-Penhey et al., 2004 Denz-Penhey et al., 2005 Denz-Penhey & Murdoch, 2008 Denz-Penhey & Murdoch, 2009 Denz-Penhey & Murdoch, 2010
HMS: Harvard Medical School - includes the HMS-Cambridge Integrated Clerkship (HMS-CIC) & Beth Israel Deaconess Medical Center (BIDMC) sites	5	Bell et al., 2008 Hirsh et al., 2012 Ogur et al., 2007 Ogur & Hirsh, 2009 Peters et al., 2001
PRCC: Parallel Rural Community Curriculum, Flinders University, Australia	8	Couper et al., 2010 Couper & Worley, 2011 Stagg et al., 2009 Walters et al., 2011 Worley & Lines, 1999 Worley et al., 2000 Worley et al., 2004 Worley et al., 2006
RPAP: Rural Physician Associate Program, Minnesota	6	Halaas et al., 2008 Power et al., 2006 Verby et al., 1988 Zink et al. 2008 Zink et al., 2010a Zink et al., 2010b

Findings answering the research questions

How are longitudinal placements defined? What are the lengths of such placements?

During the initial review of abstracts, the TRG had to decide what period of time would be used since there is no existing criterion for a program to be termed 'longitudinal'. The enormous variety of program structures added to the difficulty that the TRG faced with this definition. After much discussion, the TRG defined longitudinal for the purposes of the inclusion criteria for this review as lasting over 13 weeks with students interacting with the same patient population and/or the same faculty member/preceptor/clinician over this time period. This time period was selected as traditional specialty-based rotations are usually of eight weeks duration and therefore, in our opinion, a longitudinal placement should exceed this amount of time. Longitudinal activities during this time ranged from intervals of daily to monthly.

Studies then incorporated a number of different models and definitions of longitudinal (see table 10). The most common format was a full year in the same location (model G). For the longer full-time placements (models F, G, H) to be longitudinal, in addition to the length of time, there had to be continuity of patients and/or teachers as agreed in the inclusion criteria. Many placements were also integrated; that is the students did not rotate through blocks of specialty training but learnt about the specialty either opportunistically because of patients seen, or through attachments to different specialties throughout the week. Thus, a student at a specific rural clinical school for one year who rotated through eight-week specialty blocks would not be classified as doing a longitudinal placement. In contrast, a student at a specific rural clinical school for one year undertaking integrated specialties and/or with a named preceptor would be classified as undertaking a longitudinal placement. Model F was also an integrated placement but for a shorter time.

Models A – E were longitudinal because students were attached to a preceptor in a defined location for a proportion of the week for more than three months, but for the rest of the week would either be doing non-integrated specialty rotations or combined multispecialty rotations. The locations included general practices or community-based offices, and hospital clinics.

Table 10: Format and length of placements

MODEL	FORMAT	EXAMPLES
A.	From between 1 evening or ½ day/week to 2 ½ days per week for between 6months to 1 year	Berger & Schaffer, 1986; Fratarelli et al., 2004; Geyman et al., 1984; Hadac et al., 1979; Hamilton et al., 1998; Henley et al., 2000; Lubetkin et al., 1999; Ringdahl et al., 2007; Wee Liang et al., 2011
B.	About 4 hours/week for 2 years	Carney et al., 2005;
C.	One day per month for 2 years	von Below et al., 2008
D.	½ to one day/week for 3 weeks per month or weekly for 3 years	Deterding et al., 1999; Florence et al., 2007; Lemon et al., 1995;
E.	½ day/week for 4 years	Herold et al., 1993
F.	Full time for 7 to 9 months	HMS ; Konkin et al., 2011; McLaughlin et al., 2011*; Poole et al., 2010; RPAP ; Schauer et al., 2006;
G.	Full-time for 1 year	UWA ; Hauer et al., 2012; Lewin et al., 1999; McLaughlin et al., 2011*; Mihalynuk et al., 2008; O’Brien et al., 2012; Poncelet et al., 2011; Tehereni et al., 2009; Walmsley et al., 2009; Wilson et al., 2008.
H.	Full-time for more than 1 year	Oswald et al., 2001; Prislin et al., 1998**

*McLaughlin et al. (2011): includes three medical schools with slightly different formats

** Prislin et al. (1998) includes five medical schools with different formats

What learning outcomes/objectives are defined?

Not all the papers defined specific and explicit learning outcomes, however there was an implicit or explicit (e.g. Wilson & Cleland, 2008) understanding that the learning outcomes were the same as those of students undertaking traditional block rotations if such activities were running in parallel. Some outcomes were very broad and more related to program objectives rather than assessable outcomes, for example: integration of basic and clinical science (Bell, 2008), humanism in patient care (Bell et al., 2008; Couper & Worley, 2010), concepts relating to the doctor-patient relationship (Carney et al., 2005; Lewin et al., 1999), primary care needs of rural, underserved communities (Florence et al., 2007; Halaas et al., 2008), development of insights into community-based medicine (Frattarelli & Kamemoto, 2004); continuity of care (Bell, 2008; Frattarelli & Kamemoto, 2004; Hadac et al. 1979; Geyman et al., 1984; Halaas et al., 2008; Hirsh et al., 2012; Poncelet et al., 2011; Schauer & Schieve, 2006); to experience the lifestyle and work of rural family doctors and their families (Verby, 1998); to work effectively within organisations (Lewin et al., 1999) and to understand a physician’s professional role (von Below et al., 2008).

Other outcomes, while still broad and generic, could be assessed in line with overall student outcomes for the medical program or rotation, for example: fundamentals of interviewing and physical examination (Carney et al., 2005), disease prevention & health promotion (Florence et al., 2007; Lewin et al., 1999), and primary care practice and problems including long term conditions

(Geyman et al., 1984; Hamilton et al., 1998; Power et al., 2006; Schauer & Schieve, 2006; Verby, 1988).

For some rotations there were specific longer-term and policy outcomes for the initiative itself, for example: to encourage graduates to practise in under-served rural, remote and outer metropolitan regions (Couper & Worley, 2010; Denz-Penhey et al., 2005; Poole et al., 2010; Worley et al., 2000) and retention and performance in family medicine residency (Ringdahl et al., 2007).

Are the placements effective? In what ways are these placements effective?

Effectiveness in this respect has a number of meanings. In educational terms, effectiveness is related to whether students achieved the defined learning outcomes or objectives of the placements. In terms of evaluating a new intervention, the aim is more to establish whether the new placement has the same effects on learning as the more traditional program – is it worse, the same or better? Both of these outcomes relate to Kirkpatrick levels 2b and 3, with student self-report of satisfaction and/or change relating to level 1. Some of the rurally-based programs were also established with the goal of increasing the number of graduates subsequently choosing to work in rural locations. We define this as level 1/2a in table 6 as it was not always stated in the papers if individual students had changed career choice but rather it stated total numbers from a program choosing rural careers compared with non-longitudinal students.

Across all the papers the majority of students were highly satisfied with the longitudinal placements in all formats. In papers reporting assessment data (n = 24) all placements but two were effective in that students did as well as, if not better than, their peers when compared by the same assessment methods (Bell et al., 2008; Denz-Penhey & Murdoch, 2010; Hirsh et al., 2012; Lewin et al., 1999; Ogur et al. 2007; Oswald et al., 2001; Poncelet et al., 2011; Poole et al. 2010; Power et al., 2006; Ringdahl et al., 2009; Schauer & Schieve, 2006; Verby 1998; Wilson & Cleland, 2008; Worley & Lines, 1999; Worley et al., 2004; Zink et al., 2010a). Verby's (1998) paper reporting on 16 years of the RPAP includes a comparison of RPAP students' confidence compared to their non-RPAP peers. The RPAP cohort self-reported higher scores in several areas including treatment and professional skills; in other areas, there was no difference. Zink et al. (2010b) compared the performance during an OSCE of eight RPAP students and eight of their peers and found that the longitudinal students were better at building rapport with patients and had a better structure for their interviews, although there was no significant difference in overall communication or decision making.

Ringdahl et al. (2009) compared their participants' in-training assessments over time with other graduates on family medicine residencies. This was because the program at the University of Kentucky was specifically developed for fourth-year students who had applied for family medicine residency training after graduation and who were subsequently attached to a family practice for up to one day per week during their final year. The program is called the integrated residency (IR). IR residents outperformed traditional residents in examinations for all three years of residency (Ringdahl et al., 2009).

There were two exceptions in relation to overall assessment performance. The first of these involved comparison of the obstetrics and gynaecology final examination outcomes for students in a longitudinal multispecialty clerkship versus those doing traditional block rotations. The program comprised a six-month longitudinal community-orientated clerkship at the University of Hawaii, which included weekly sessions with a community preceptor. While students on the longitudinal program and those on the traditional block program showed no differences in clinical performance as rated by their tutors, and no statistically significant difference in student

final clerkship grade in obstetrics and gynaecology, those in the block program scored significantly higher in the written component of their final examinations (Fratarelli & Kamemoto, 2004). The authors stated possible reasons for this outcome including student self-selection to the different programs, and failure of the examination process to capture elements where longitudinal students may have performed better (for example, interpersonal and clinical skills). Although not discussed in the paper, the difference may also have been related to the written assessment focusing more on hospital in-patients than the predominantly ambulatory patients seen by the longitudinal students. The second exception was students at three medical schools in Canada undertaking the longitudinal integrated clerkship (LIC). In year 3, these students had significantly lower OSCE scores than rotation-based clerkship (RBC) students, although they performed better in an assessment of clinical decision-making (McLaughlin et al., 2011). However, this may have been due to the RBC faculty having more input into the design of the OSCE stations.

In terms of perception, rather than hard assessment data, students and preceptors surveyed for the community-based first year program at the University of New York felt that the weekly office session helped students achieve communication and socialisation skills better than practical skills (Lubetkin et al., 1999). This study also highlighted the fact that students and preceptors may not agree on whether students have met defined program goals, with preceptors being more satisfied that goals had been met compared to students.

Change in attitudes

Only three of the papers included the use of validated scales to evaluate changes in students' attitudes and all of these were from Harvard Medical School (HMS). The validated PPOS (patient practitioner orientation scale), a measure of attitudes to patient-centred care, was administered to pilot integrated (eight) and traditional track (18) students pre and post the nine-month attachment at HMS. While there were no differences between the two sets of students before the clerkships, the pilot students' patient-centredness did not change, whereas the traditional students showed a decline (Bell et al., 2008). The scale was used in a later evaluation of 27 integrated clerkship students and 45 traditional track students with similar results (Hirsh et al., 2012). Ogur et al. (2007) administered the TOMS (tasks of medicine scale) at the beginning and end of the clinical year. TOMS assesses students' prioritization of biomedical and psychosocial tasks in a patient encounter (Krupat et al., 1999). By the end of the year the integrated students' scores had increased compared to the traditional students, suggesting that there had been no ethical erosion in the former group.

Other papers included students' self-reports of changes in attitudes and approach to patients but these are included at Kirkpatrick level 1. We make three exceptions for inclusion at level 2a. Florence et al. (2007) asked specifically about students' attitudes to professional preparation and found that CPP (Community Partnership Program – a longitudinal experience in rural Tennessee) students rated their preparedness to practise in interdisciplinary teams higher than traditional students. The CPP students were also more likely to work in a volunteer capacity in the community. Konkin & Suddards (2011) and Ogur & Hirsh (2009) adopted a narrative approach to student evaluation. Their in-depth analysis of students' stories suggested that students developed compassion and a deeper connection with patients during their placements.

What factors influence effectiveness?

The major factors, or mechanisms, appear to be continuity of patient care, quality of participation, continuity of preceptor/mentor, learning environment and, where there are

different types of placements within the same medical school, students feeling that there is equity in learning and assessment.

The preceptors' discipline/specialty in early longitudinal placements also appears to affect student learning and performance later in their programs. At the University of Colorado, students undertook a primary care curriculum during the first three years (of the four year course). This entailed being attached to the same preceptor in his/her community clinic for a half day per week for three weeks every month. Preceptors were primary care internal physicians, paediatricians or family physicians. Students, who had paediatric preceptors, outperformed their peers in subsequent final year paediatric examinations, even though all students had subsequently undertaken a hospital-based paediatrics rotation (Deterding et al., 1999).

Continuity of patient care

Evaluation data showed that students valued the opportunity to work and learn with a defined group of patients over time (Poncelet et al., 2011). Longer rotations enabled the development of personal relationships and rapport with patients (Geyman et al., 1984; Walmsley et al., 2009) and the ability to follow patients over time; this allowed students to see how illnesses develop, change and either resolve or worsen with time (Couper & Worley, 2010; Oswald et al. 2001). Students were involved with, and learned about, the 'whole life cycle of health and disease' (Couper et al., 2011, online p.4). They gained an enhanced understanding of the psychosocial impact of chronic disease and its management (Walmsley et al., 2009). They also began to develop a physician identity grounded in caring (Konkin & Suddards, 2011), with the realisation that continuity of care generates a trusting relationship between patient and doctor that may help to prevent exacerbations of chronic illness (Berger & Schaffer, 1986).

O'Brien et al. (2012) used an ethnographic approach to compare students on a LIC with block clerkship students (BC). While LIC students followed up 34% of patients, BC students only saw 5% of their patients more than once. The University of Washington's (Seattle) initiative focussed specifically on extended patient contact (one half-day per week with a family physician for one year) and students' attitudes towards the values of this continuity were evaluated (Hadac et al., 1979). The authors of this study defined continuity of care in their introduction and specifically state that students 'were also asked to pay particular attention to such aspects of longitudinal comprehensive care as preventative measures, long-term counselling, use of community health resources and the effects of family dynamics on patient management (Hadac et al., 1979, p. 528). The students reported that the continuity experience was valuable and should be included in the medical curriculum for all students. (Note: while this study surveyed attitudes, it did not explore any change in attitude and therefore is included as student reaction.)

Participation

In comparison to students on traditional short rotations, students on longitudinal placements were self-reportedly less frustrated and less marginalised, being more integrated into clinical life and work (Bell et al., 2008; Ogur et al. 2007). They did less 'scut work' compared to other rotations (Prislin et al., 1998, p. 683) and felt more valued (Worley et al., 2006). These students took on responsibility, which built over time, whereas students on block placements had to 'start over' for each new rotation (Mihalynuk et al., 2008, p. 730). Students undertook increasingly complex tasks as the doctor-student relationship matured and trust and respect developed (Walters et al., 2011). They described their experience as 'hands-on' (Worley et al.,

2006). LIC students spent significantly longer time carrying out direct patient care than BC (block clerkship) students who were more likely to be observing care (O'Brien et al., 2012). This could have the paradoxical result that LIC students saw fewer patients in out-patient departments than BC students because they were taking greater responsibility for the patient and interacting with them for longer (O'Brien et al., 2012). Not surprisingly, LIC students also reported more optimal mentorship (Bell et al., 2008) and less competition for patients than their peers in the teaching hospitals (Worley et al., 2006; Zink et al., 2008); such competition was described as 'draining' (Worley et al., 2000, p.563).

Continuity of mentorship

In the papers are a number of terms relating to clinicians who interact in an educational capacity with students: mentor, supervisor, tutor and preceptor. These terms are not usually defined though the nature of the relationship is usually apparent from the text. When citing papers we use the term that the authors use, while acknowledging that there are similarities and differences in the usage.

Continuity of mentorship helped students' personal and professional growth (Couper & Worley, 2010) and enabled them to have greater responsibility for patient care as trust developed (Couper et al., 2011; Denz-Penhey et al., 2005; Fratarelli & Kamemoto, 2004). Students and mentors developed close relationships, which students perceived provided personal interactions missing from classroom and clerkship activities (Lemon et al., 1995). Preceptors were important role models for students (Walmsley et al., 2009), reporting that they regarded students as future colleagues (Walters et al., 2011).

The continuity of the student-preceptor relationship and the motivation of the preceptor appeared more important in one study than the placement location or status of the clinician: Carney et al. (2005) compared 155 student placements over the first two years (of a graduate entry four year program) in community clinics, academic medical centres and out-patient offices and found no difference in OSCE scores. Students at the University of Illinois rated family nurse practitioners (FNPs) as effective mentors in their stable locations in ways that were complementary to rather than interchangeable with those of their family physician preceptors. The FNPs were evaluated highly on their psychosocial approach and communication skills, dealing well with sensitive issues, and being supportive of the students' learning process; they scored lower on 'subject-matter expertise and problem-solving emphasis' (Henley et al., 2000, p. 493). The stable location or home base of the mentor was also preferable to moving between locations to enable wider clinical experience (Denz-Penhey et al., 2005). Such continuity has also been described as an apprenticeship model (Fratarelli & Kamemoto, 2004).

Another possible advantage of the mentorship relationship was demonstrated by McLaughlin et al. (2011), who looked at the completion of in-training evaluation reports (ITER) at three Canadian medical schools. Students undertaking the longitudinal integrated clerkship (LIC) in third-year had their ITERs completed by the same mentor for each of the six mandatory clinical disciplines, whereas the rotation-based clerkship students (RBC) had a different assessor for each rotation. The mean ITER rating was higher for the LIC students compared to the RBC for both clinical skills and professional attributes. The authors in considering this finding speculated whether it reflected the 'longer and more contextually rich observation period' (p. S27). On the other hand, it might be symptomatic of a leniency effect with mentors finding it difficult to give lower scores to students with whom they have had a supportive relationship.

Of interest are the findings of another study comparing longitudinal with block placements (Mihalynuk et al., 2008). While only involving six students on each type of rotation, the difference in continuity experiences and definition are notable. Longitudinal students appreciated their year-long attachment to primary care and the continuity of educational mentorship; block students identified benefit from seeing different patients with similar problems repeatedly over their short placement time (Mihalynuk et al., 2008).

The learning environment and support (continuity of location)

Students found that being in the same learning and clinical environment (i.e. the same clinical team, rather than purely geographical location) for a prolonged period made them feel comfortable as well as useful (Walmsley et al., 2009), describing the environment as nurturing (Zink et al., 2008). Notably, Ogur & Hirsh (2009) when reporting on the HMS-CIC stated that the 'students' site of learning was not place-, team- or specialty-specific but rather resided wherever their patients' needs arose' (p. 848). This then resulted in students not being as heavily exposed to the adverse effects of the hidden curriculum and not developing the defence mechanisms that health professionals adopt in intense clinical environments (Ogur & Hirsh, 2009).

Coordinators-administrators of placements also had an essential role in rural areas (Denz-Penhey et al., 2004) and it was important that they acted on past feedback to make changes to improve placements for subsequent years (Denz-Penhey & Murdoch, 2009). Students needed to feel confident that, during the longitudinal placements, they were undertaking the same curriculum as their peers doing more traditional rotations and that they were learning what they required to pass their assessments (Denz-Penhey et al., 2004).

One goal of the HMS-CIC was that students should be taught more often by senior doctors (attendings) rather than house officers, and this goal was achieved for the eight students as reported by Ogur et al. (2007). Continuity of location also helped students appreciate and develop teamwork capabilities (Berger & Schaffer, 1986; Couper et al., 2011; Denz-Penhey & Murdoch, 2008; Wee et al., 2011) while feeling better prepared for interdisciplinary practice than non-longitudinal graduates (Florence et al., 2007).

What other impact do the placements have on clinicians, students, future career choice, and patients?

Clinicians

Some longitudinal placements in rural areas have recruited GPs new to teaching. New tutors could be initially wary of the additional workload and concerned about students moving from being observers to participants in practice (Denz-Penhey et al., 2004). Preceptors reported a need to develop alternate teaching skills due to their greater responsibility for student developmental learning (Tehereni et al., 2009). However, while there was a higher time commitment at the beginning of longitudinal placements, this decreased as students became more skilled, in contrast to the time needed at the beginning of each traditional student block for orientation (Poncelet et al., 2011). Some preceptors expanded their working hours to fit in time for teaching (Tehereni et al., 2009) and overall felt that LIC students took more time because of individualised mentoring and disruption to clinic hours (Tehereni et al., 2009). Rural preceptors were motivated by the desire to increase rural recruitment (Walters et al., 2011).

Practitioners have reported in interviews that their teaching role has given new meaning to their practice (Couper & Worley, 2010) and provided 'professional enrichment' and variety from routine consulting (Walters et al., 2011, p.457). The quality of students' written reports helped preceptors become more enthusiastic about arranging patients for students to follow (Hadac et al., 1979). Clinicians were stimulated by having students in their workplaces and reported that students became assets to the practice (Couper et al., 2011) as they progressed in the amount of their authentic clinical participation (Walters et al., 2011). In the HMS study, faculty satisfaction was high with the majority (82.6%) finding their professional lives more satisfying (Ogur et al., 2007). Teaching of longitudinal students was rated as positive and rewarding (Poncelet et al., 2011; Tehereni et al., 2009), bringing kudos and recognition (Walters et al., 2011). However some community preceptors were concerned that students did not see a wide diversity of patients, as they were dependent on the opportunistic nature of primary care attendances (Tehereni et al., 2009). Others felt they needed greater support from their specialist colleagues in the clinics (von Below et al., 2008).

Some preceptors reduced the number of patients seen to accommodate students depending on feasibility (Tehereni et al., 2009), which could impact on patients' waiting times due to less availability of appointment times. They also reported feeling pressure because of the frequently competing needs of their patients and students (Walters et al., 2011).

Students

Students perceived that their relationships with patients fostered through continuity emphasised the importance of patient care not only from a medical perspective but also a 'life perspective' (Couper & Worley, 2010, p. 35). Thus, it enabled them to see the patient 'behind the illness' (Denz-Penhey & Murdoch, 2008, p.783) as well as the importance of family dynamics and the interplay between a patient's health and their family (Geyman et al., 1984; Hadac et al., 1979). Social context and its importance to patient care were highlighted (Hirsh et al., 2012; Ogur et al., 2007) enabling students to gain a broader perspective on the experience of illness (Ogur & Hirsh, 2009). Students formed therapeutic relationships with patients (Konkin & Suddards, 2011; Ogur & Hirsh, 2009) and felt that their experience had better prepared them to be caring (Ogur et al., 2007). As particular placements became bedded into curricula, students began to expect better learning opportunities and teaching in rural locations than in the bigger centres (Denz-Penhey & Murdoch, 2009).

Integration of biomedical and clinical learning was a feature of two programs. One of the initiatives involving early patient contact in the first years of the program in Gothenberg, Sweden, provided increased motivation for students as they studied biomedical sciences in the classroom (von Below et al., 2008), while students in Singapore were able to apply their science knowledge in the clinical setting (Wee et al., 2011).

In terms of patient care, students were able to correct faulty or misleading information in patient records (Hadac et al., 1979) and bridge gaps in communication between multiple providers (Ogur & Hirsh, 2009). They were also able to share information about patients across health care settings if they had allocated patients who moved between such settings, eg primary to secondary care (Hauer et al., 2012). Some students also reported that they felt they had a positive impact on patient care (Poncelet et al., 2011). However there was no data as to whether these particular effects impacted on measurable patient outcomes.

Choice of career

Longitudinal rural placements appeared to be improving attitudes towards rural practice in Australia, Canada and the USA with potential and actual enhancement of rural recruitment and retention (Couper et al., 2011; Florence et al., 2007; Halaas et al., 2008; Stagg & al, 2009; Verby 1998). However some primary care specific programs did not seem to be affecting choice of career (Herold et al., 1993; Hirsh et al., 2012).

Patients' feedback and impact

Wee et al. (2011) interviewed 355 patients via a structured questionnaire about their experiences of the longitudinal interprofessional student-run home visit program in Singapore. This program focused specifically on screening an underserved population for chronic disease and management of existing conditions, with students making visits to patients fortnightly and then monthly for six months. Patients reported that their health had improved over this time and that students had been able to assist them in addressing their health issues. Treatment rates for hypertension also increased significantly during the program and these were maintained over the following year.

Berger & Schaffer (1986) surveyed patients following their involvement in a weekly student run clinic in a deprived and underserved area of East Virginia. While the number of patients was not reported, all those who responded stated that the care they received at the clinic was better than any other previous care they had experienced.

How do longitudinal placements promote learning?

The placements appeared to promote learning through engaged participatory learning and the great variety and extent of clinical experiences (Couper et al., 2011). Teaching tended to be learner-centred (Zink et al., 2008). Students reportedly received a better quality of feedback and were more often observed performing clinical skills (Poncelet et al., 2011) than on traditional placements. They fostered 'a sense of intellectual and multidisciplinary inquiry' (Ogur & Hirsh, 2009, p. 845). Students were expected to be self-directed learners (Denz-Penhey et al., 2004) and were better able to seek opportunities for education (Denz-Penhey & Murdoch, 2008; Zink et al., 2008), while engaging in self-reflection (Hirsh et al., 2012). This increased their confidence (Denz-Penhey & Murdoch, 2008; Prislis et al., 1998) and trust in their own judgment (Denz-Penhey & Murdoch, 2008), while enhancing their sense of personal efficacy (Denz-Penhey & Murdoch, 2009) and their ability to deal with ambiguity (Hirsh et al., 2012). Students began to assume a more doctor-like role with their preceptors framing this role as that of a legitimate team member (Hauer et al., 2012); this role was meaningful to students (Konkin & Suddards, 2011). Hauer et al. (2012) compared students on longitudinal placements with those on block rotations across three institutions (University of California, Harvard Medical School, University of South Dakota) and while both sets of students did assume this doctor-like role to some extent, the block students were only able to do this in ambulatory settings and even then reported this as rare. In comparison, the role was more commonly adopted by longitudinal students.

Factors inhibiting learning

The longitudinal type of experience emphasising self-direction did not suit all students – some needed more structure (Couper et al., 2011). Students accustomed to being directed could feel as if they were 'floundering' in the longitudinal environment (Worley et al., 2000, p. 563). Even in the same program, the location and nature of the clinics where students were attached could

affect the quality of the experience. Students in subspecialty clinics lacked variety and some were unable to follow up patients as often as they would have liked (Walmsley et al., 2009). Having only a single supervisor could be a problem if relationship issues arose (Worley et al., 2006).

While students had advantages in being part of the clinical team, they could also feel overworked - with little time for specific learning (Couper et al., 2011; Wilson & Cleland, 2008). Enthusiastic tutors expected students to see interesting patients at all times of the day and out-of-hours and some students found it hard to set boundaries around their working day and to manage their time effectively (Denz-Penhey et al., 2004). Students in more remote locations could feel isolated (Couper et al., 2011). Students also reported some uncertainty about their role with patients: there could be a blurring of the line between being a friend and a care provider, and feelings of pressure because patients viewed them as being fully responsible for clinical care (Ogur & Hirsh, 2009).

As ever, dealing with problematic students could be difficult for supervisors, and this was compounded by the distance from central administration in more remote locations (Couper et al., 2011). However, staff sometimes blamed students if they complained about poor organisation or excessive workload rather than trying to discuss the underlying issues (Denz-Penhey & Murdoch, 2009).

New models of education resulted in some clinicians in traditional programs making negative remarks to LIC students about their ability to pass examinations due to their primarily community experience (Oswald, 2001), highlighting the need to brief staff adequately and inform them of the evidence about assessments.

Student anxieties and concerns

Those programs in which only a proportion of students were undertaking the longitudinal placements precipitated some anxiety in students about learning opportunities and equitable assessment in the first few years (Denz-Penhey et al., 2004). Students could take a while to adapt to a non-discipline specific (integrated) attachment and to feel competent in individual disciplines; however, they were also able to be more flexible in filling in the gaps in their experiences (Mihalynuk et al., 2008). There could be gaps in exposure to certain conditions in some specialties in rural sites, which did not have certain services locally, for example, paediatrics and psychiatry, and students reported missing specialist contact in certain areas (Couper et al., 2011; Denz-Penhey & Murdoch, 2008). Students could also become frustrated if their high expectations of teaching were not met, and expressed dissatisfaction with locations if resources (such as up-to-date books and an adequate number of rooms) were lacking (Denz-Penhey & Murdoch, 2009).

Prislin et al.'s (1998) study of ambulatory primary care longitudinal placements across five USA medical schools reported on 429 students' perceptions from one academic year. While satisfied in many areas, students generally felt less impressed by their acquisition of medical knowledge in the longitudinal attachments and with their communication with members of health care teams. This however contrasts with their feedback in terms of enhanced learning with respect to their on-going management of chronic disease, and recognising and engaging with patients' hidden agendas and psychosocial problems. Wilson & Cleland (2008) reported on a rural year-long attachment in Scotland. Students were highly satisfied with the placement but expressed concerns about book learning; they felt they continued to be working on the wards while their

colleagues at the medical school and teaching hospital were able to focus on learning in the library.

Initiatives in locations new to having students and with inexperienced tutors could lead to problems. Students who undertook the first year of a rural program in New Zealand reported that only 42% of their supervisors helped with identification of learning needs and there were also concerns about the supervisors' lack of knowledge about curriculum and examinations (Poole et al. 2010). This highlights the need for adequate preparation of staff and students.

Discussion

This review aimed to explore, analyse and synthesise evidence relating to the effectiveness of longitudinal placements including LICs, as a means of achieving learning outcomes in medical student pre-qualification training programmes. The review examined the nature and scope of programs and factors associated with successful outcomes.

Taking all the evidence and evaluations into account, longitudinal placements appear to have an effect in helping students meet defined learning outcomes such as the integration of basic and clinical science, the development of a patient-centred approach to clinical care, presentations of illness in primary care settings and the development of insights into community-based medicine. Students develop an understanding of the importance of continuity of care through their own continuity experiences. In rural settings they are able to experience the lifestyle and work of rural family doctors and their families, with this experience appearing to have an effect on later career choice. Our findings relating to LICs in particular are similar to those of Walters et al. (2012), which also highlight that these programs help students develop a more patient-centred approach through the provision of activities that promote higher order clinical skills. However, much of the evaluation data is related to student reaction, most of the studies and LIC programs are small and also are concentrated in a small group of medical schools, and many studies involve students who have self-selected. Despite these limitations, we feel that there is enough evidence to support further development of longitudinal placements in conjunction with more robust realist evaluation.

In those medical schools where traditional block and integrated longitudinal placements exist in parallel, there is generally no difference in assessment results, except where assessments are designed to reflect hospital specific skills and/or knowledge. Where such interventions are concurrent, it is important that students are informed of the equity in learning experience and assessment outcomes.

The underlying mechanism promoting learning is overwhelmingly that of continuity in its varying forms of patient, mentor and location longitudinal exposure. As Hirsh et al. (2007) noted clinical continuity fits with the theory of adult learning and its progressive development of knowledge and skills through experience. Hirsh et al. (2007) go on to write that: '...educational continuity subsumes two interrelated integrating forces: horizontal integration (enhancing the development of general competency by linking learning experiences between and across clinical specialties) and vertical integration ... linking advances in the biomedical and clinical sciences to clinical problem solving) (p. 858). Such continuity builds up trust between tutors/preceptors and students, and longer term relationships between patients and students. In realist terms, this mechanism is linked to the theory of communities of practice: students become members of a community of practice, developing legitimate peripheral participation (Lave & Wenger, 1991) with attendant responsibility depending on their competency and seniority. Their enhanced understanding of how illness changes and its

effects on patients and families promotes a patient-centred approach, rapport building, compassion and caring. Some interventions also enrich students' experiences of teamwork.

The immersion aspects of longitudinal placements also reinforce the community of practice model. This immersion varies from an intermittent (i.e. not continuous and consecutive days) but sustained exposure over time to full-time placement in a practice. The intermittent models, most commonly occurring during the first half of a medical program, are often required activities and give students an understanding of the psychosocial aspects of patients' lives. Other outcomes including the development of compassion and rapport building are, not surprisingly, similar to those arising from early patient contact, though the BEME review on early patient contact did not focus on longitudinal experience (Dornan et al., 2005). The immersion models placed in later years of medical courses are usually voluntary. Students feel valued and, both they and their tutors recognise their worth. In the majority of cases, preceptors also gain from hosting students, but have to be mindful of their own workload and have skills and support in dealing with problem students.

Within the community of practice model student learning on longitudinal placements also resonates with adult learning theory (Knowles, 1978), with its emphasis on experiential learning, problem solving, self-direction and relevance of defined learning outcomes to students' current and future practice.

While continuity is the underlying mechanism, important aspects of the context include: motivated and orientated mentors/preceptors who have the understanding of their peers in terms of workload; dedicated administration staff in geographically dispersed locations; adequate selection and preparation of students who are able to direct some of their own learning and take advantage of their clinical experiences; attention to equity issues when students are undertaking different experiences across the same program; involvement of patients in planning if they are to receive hands-on care from students, particularly those in their junior years.

Conclusion

If students are able to have a continuity of experience, that is, some form of longitudinal placement, they are able to develop their capabilities in patient-centred care, while gaining the trust of their tutors. For the placements to work, tutors and students need adequate support, especially if both are a distance away from the organisational medical school hub. The different models have different outcomes and effects. The intermittent models are especially suitable to junior students who are not ready for the more complex clinical tasks, while the immersion and integrated models are effective for senior students who are able to take advantage of the experience of becoming members of the patient care team.

Placements are less likely to work if academic and administrative staff are ill-prepared or ill-supported, if students feel disadvantaged in terms of resources and specialty experiences, and/or if students are unable to take responsibility for their own learning.

Practice points

- Longitudinal placements have been developed in contrast to more traditional shorter clinical placements and come in a variety of formats includes longitudinal integrated clerkships (LICs).

- The major factors, or mechanisms, associated with their effectiveness, appear to be continuity of patient care, quality of participation, continuity of preceptor/mentor, learning environment and, where there are different types of placements within the same medical school, students feeling that there is equity in learning and assessment.
- Longitudinal placements enhance students' understanding of patient-centredness, and the importance of the life perspective, family dynamics and social contexts of patients' presentations.
- The immersion aspects of longitudinal placements reinforce the community of practice model and students, as adult learners, learn through experience and problem solving.

Recommendations for medical education

- All medical students should have the opportunity to undertake a longitudinal placement with continuity of tutor/preceptor: however, there may be resource and logistical implications for such placements, particularly in medical schools with high student numbers.
- Given the trust and rapport that builds up between student and tutor, and student and patients over time, placements should be greater than the 8 weeks duration of a common standard block or discipline-specific rotation.
- If there are parallel placements at the medical school (e.g. longitudinal and block) the learning outcomes should be similar for all placements to ensure equity and they should form the basis of the assessment.
- It is important to prepare administrators, students and clinicians for the placements – all must have a good understanding of what this commitment entails.
- Continuity of supervision and patient access needs to be a defining feature of the placements.
- Some students on longitudinal placements located a distance from the medical school need support and structure to their learning.
- Students having only one supervisor for long periods of time may be disadvantaged if the relationship is not optimum. Both students and supervisors need access to program staff in case of such problems.
- New supervisors/preceptors require mentoring.
- The workload of students and staff needs to be monitored to avoid it becoming excessive.

Limitations of the review

Our grading of papers in relation to their 'quality' and significance was to some extent subjective, though we had good agreement and all papers included in the final 53 were coded and graded by two members of the TRG. There was considerable overlap in data published by the same institutions and programs but again the consistency of outcome data and our reflection on program evaluation suggest that our conclusions are valid for the data provided.

Recommendations for further research

We have provided a definition of longitudinal placements and a realist evaluation of data published in relation to these. We would suggest further exploration of the effect of continuity and the minimum time and frequency for placements to achieve this effect. As with other innovations in health professional education delivery, there also needs to be more exploration of the impact of the placements on patients and their health outcomes. While there are now a number of medical

schools with longer term data, the impact of LICs does require on-going research over time, often a difficult proposition given the lack of resources and continuity of researchers within health professional education and universities.

Further realist evaluation should focus on why LICs are effective for some students and not others. How may LICs be tailored in terms of length, integration and location to be suitable for more students? Attention also needs to be given to the faculty development needs of mentors/tutors.

Appendix 1: Search syntax

PUBMED

“General Practitioners”[Mesh] OR “General Practice”[Mesh] OR “General Practitioner” OR “General Practitioners” OR “General practice” OR “Family Physician” OR “Family Physicians” OR “Physicians, Family” OR “Primary Care” OR “Rural Health Services”[Mesh] OR Rural OR Community OR Hospital

AND “Education”[Mesh] OR Education OR Educational OR Teaching AND “Students, Medical”[Mesh] OR “Medical Student” OR “Medical students” OR “Medical student's” AND “Longitudinal Studies”[Mesh] OR Longitudinal OR Extended OR Integrated AND “Clinical Clerkship”[Mesh] OR Clerkship OR Clerkships OR Rotations OR Attachments OR Placements OR Placement OR

Terms OR Modules OR Programs OR Program

MEDLINE

General Practitioner OR General Practitioners OR General practice OR Family Physician OR Family Physicians OR Physicians, Family OR Primary Care OR Rural Health Services OR Rural OR Community OR Hospital

AND Education OR Educational OR Teaching

AND Students, Medical OR Medical Student OR Medical students OR Medical student's AND Longitudinal Studies OR Longitudinal OR Extended OR Integrated AND Clinical Clerkship OR Clerkship OR Clerkships OR Rotations OR Attachments OR

Placements OR Placement OR Terms OR Modules OR Programs OR Program

CINAHL, EMBASE and WoK

“General Practitioner” OR “General Practitioners” OR “General practice” OR “Family Physician” OR “Family Physicians” OR “Physicians, Family” OR “Primary Care” OR “Rural Health Services” OR Rural OR Community OR Hospital

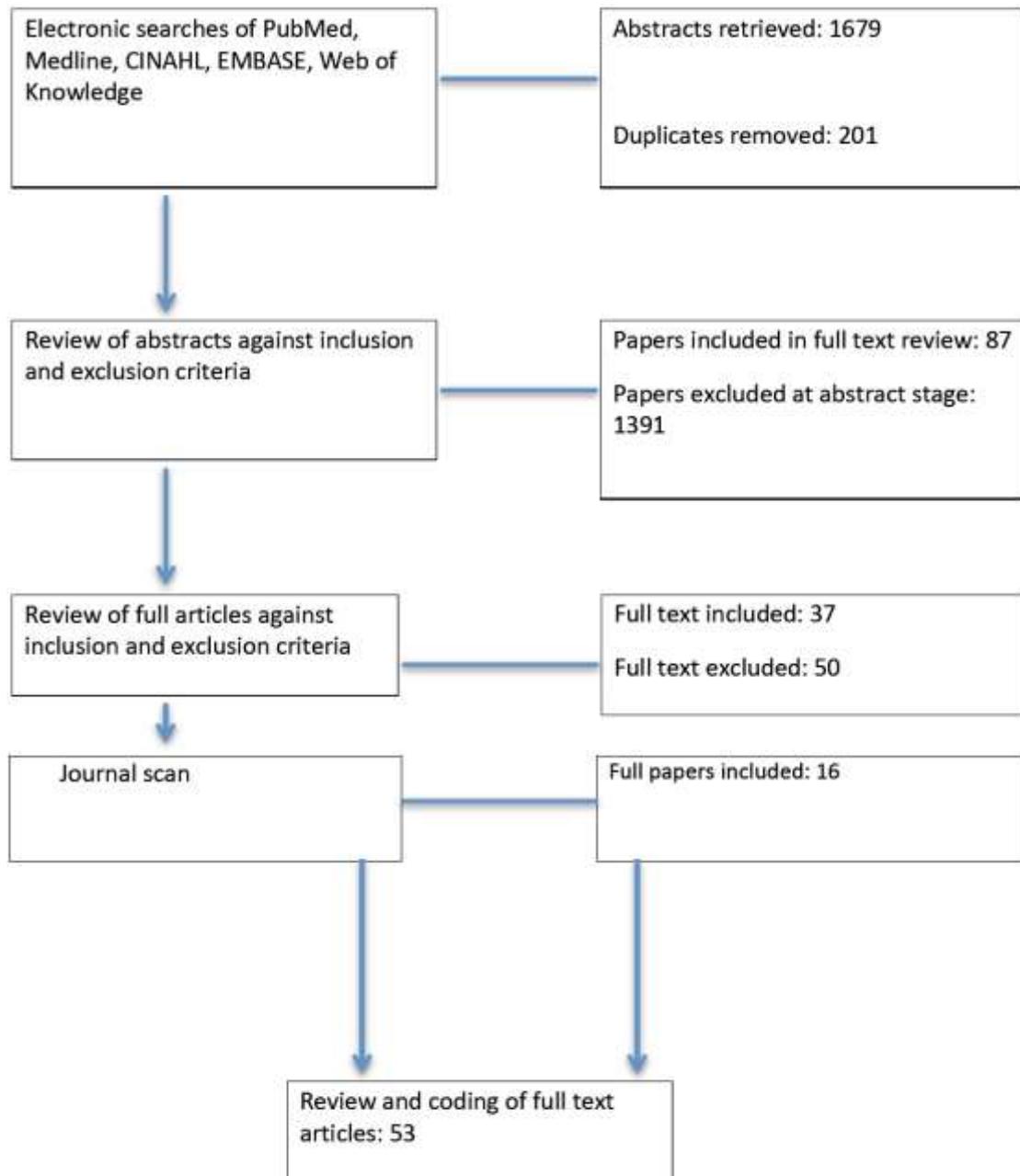
AND Education OR Educational OR Teaching AND

“Students, Medical” OR “Medical Student” OR “Medical students”

AND “Longitudinal Studies” OR Longitudinal OR Extended OR Integrated AND “Clinical Clerkship” OR Clerkship OR Clerkships OR Rotations OR Attachments OR

Placements OR Placement OR Terms OR Modules OR Programs OR Program

Appendix 2: Flow diagram of literature search and paper selection



Appendix 3: Realist approach (Pawson et al., 2005) as applied to this review

Step 1: Clarify scope

- *Identify review questions:* We identified our review questions as listed in Methods.
- *Define the nature and content of the intervention:* We defined longitudinal clinical placements for medical students as lasting more than 13 weeks, and where each student interacts with the same patient population and / or with the same clinical mentor/preceptor to ensure continuity of experience
- *The circumstances or context for the use of the intervention:* The placements were longer than usual placements and/or repeated experiences in the same location with continuity or mentorship and/or patient interactions.
- *Define the policy intentions or objectives:* These were broadly defined, though not always explicitly, as prequalification curriculum reform with the aims of improving students' learning through continuity experiences, with integrated specialty rotations in longer term placements than in traditional programs, and enhancing recruitment of doctors to rural areas through prolonged exposure to community placements.
- *Refine the purpose of the review:* this was not necessary though we did adjust our definition of longitudinal during the review process.
- *Theory integrity – does the intervention work as predicted?* The main theories articulated or implied were: the importance of a community of practice and the resultant development to enable students to move from peripheral participation to an authentic role in patient care because of continuity of interactions. The review results are highly suggestive of this being the case with the intervention having a predicted outcome in the majority of cases.
- *Theory adjudication – which theories fit best?* Community of practice; experiential learning through immersion.
- *Comparison – how does the intervention work in different settings, for different groups?* The different settings were: community (urban and rural) and hospital (wards and ambulatory clinics). The different groups were: medical students at different stages of training and varying clinicians as mentors.
- *Reality testing – how does the policy intent of the intervention translate into practice?* This involved comparing the aims/defined learning outcomes of the placements at the start with the outcomes through the evaluation data presented in the papers.
- *Articulate key theories to be explored:* During the review process we looked at the literature on continuity and communities of practice.
- *Evaluative framework:* Our evaluative framework was outcomes and process, with emphasis on exploring and synthesizing the data to answer our review questions.

Step 2: Search for evidence

- *Define search strategy:* Our search strategy is reported in appendix 1.
- *Final search for additional studies when review near completion:* we checked the more recent 2012 publications

Step 3: Appraise primary studies and extract data

- *Use judgment to supplement formal critical appraisal checklists, and consider 'fitness for purpose':* as described in the methods
- *Relevance – does the research address the theory under test?* We believe so as described above.
- *Rigour – does the research support the conclusions drawn from it by the researchers or the reviewers?* All the papers were coded by at least two reviewers and discussed to reach consensus.
- *Develop 'bespoke' set of data extraction forms and notation devices. Extract different data from different studies to populate evaluative framework with evidence:* as described in methods – BEME coding sheet and analysis focusing on the review questions.

Step 4: Synthesize evidence and draw conclusions

- *Synthesize data to achieve refinement of program theory – that is, to determine what works for whom, how and under what circumstances:* as addressed in our findings and discussion
- *Allow purpose of review to drive the synthesis process:* we made sure we focused on the review questions
- *Use 'contradictory' evidence to generate insights about the influence of context:* as discussed in findings and discussion
- *Present conclusions as a series of contextualized decision points of the general format 'If A, then B' or 'In the case of C, D is unlikely to work':* see conclusions section

Step 5: Disseminate, implement and evaluate

- Dissemination will be by publication of this review. In the longer term we aim to see what impact it has on further work in this area.

Appendix 4: Case studies

Case study 1: Rural clinical school, UWA

The rural clinical school (RCS) of the University of Western Australia (UWA) is one of many RCSs funded by the Australian Commonwealth Department of Health and Ageing. The purpose of these schools is to provide a rural education and training network to increase access to and sustainability of rural health services in underserved areas of rural and remote Australia. The UWA RCS was set up in 2002 to provide a full year of integrated clinical experience for year 5 medical students (undertaking a 6-year undergraduate program) with the first cohort of 21 students arriving in 2003. Between 3 and 9 students were allocated to one of five locations, where they covered the year 5 curriculum and participated in the same assessments as the rest of their city-based year group. There was some variation in curriculum delivery depending on the clinical site. In 2004 the number of students increased to 24, and they were able to spend up to one month of the year at another site to gain experience of a large rural hospital and/or a remote Aboriginal community. The number of sites and students increased over time: ten and 62 respectively by 2007. While one of the sites was large enough for students to undertake discipline-specific rotations (so not wholly integrated), at the others students' experience depended on patient availability and opportunistic presentations.

Publications from the RCS include: an evaluation of the 2003 cohort (Denz-Penhey et al., 2004); a comparison of the RCS experience between UWA and the Spencer Gulf Rural Health School (SGRHS) of the universities of Adelaide and South Australia relating to the cohorts of 2003 and 2004 (Denz-Penhey et al., 2005); a more in-depth longer term evaluation of the RCS's 2003 and 2004 students through interviews one year after completion of their longitudinal attachment (Denz-Penhey & Murdoch, 2008); a qualitative evaluation of the 2007 cohort involving students, academic and administrative staff (Denz-Penhey & Murdoch, 2009); and a comparison of students' perceptions of their experience between the larger and smaller rural sites (Denz-Penhey & Murdoch, 2010).

The evaluations are internal and mainly focus on student and staff perceptions through structured questionnaires and semi-structured interviews. Lessons learnt through the evaluation of the 2003 cohort (Denz-Penhey et al., 2004) affected the further development and delivery of the program. Of note was that the fewer the students at a site, the greater their anxiety as they felt less peer support in this new initiative. Students were less anxious if they could concentrate on one to three disciplines at a time rather than the full range of patient problems. They valued continuity of learning but appeared to miss the boundaries typical of hospital-based discipline-specific rotations. Students also felt they should be assessed on the patients they were seeing rather than be examined on what the city-based students were experiencing. The majority of students appreciated the teaching and learning opportunities, and not having to compete for patients with other students, plus the experience of living in and socializing with the rural community.

The comparison of the RCSs in the two states (WA and SA) also involved a trial of three different types of rotations (Denz-Penhey et al., 2005): long term in one centre; mainly long term in one centre but with short rotations of 3-6 weeks away from base; 6 week rotations without a home base (the last of these therefore not being longitudinal nor integrated). Most students preferred the home base approach for academic, clinical and social reasons. Academically they liked the continuity of having the same medical preceptors for the whole year; clinically this continuity allowed them to feel useful as they became known and therefore trusted by these preceptors so that they were able to contribute to practice and take a more active role in patient care; socially they were able to take part in a wide range of community activities. In conclusion the longer rotations were more optimal in terms of student satisfaction and reaction but this paper does not comment on assessment results.

The longer term evaluation a year after the rural immersion (Denz-Penhey & Murdoch, 2008) indicated that students felt more confident clinically in their final year because of the RCS's strong focus on practical experience and patient contact (though their confidence was not compared with that of the city-based students). Students also gained good experience of teamwork and developed relationships with other health professionals, contrasting this with the city experience where they rarely saw the same doctors or other health professionals on a regular basis. There was a lack of experience in some specialties at some of the sites compared to the tertiary hospitals.

The later evaluation of the RCS four years on (Denz-Penhey & Murdoch, 2010) reports some student dissatisfaction with resources such as books and rooms for interviewing patients by themselves, though overall they still perceived they were having more clinical opportunities than their peers in the city hospitals. There appeared to be a mismatch between student expectation of their teachers and what teaching was delivered at some sites: two of the less well functioning sites were in their first year of operation. The paper mentions that all 2007 RCS cohort passed the year and their marks were indistinguishable from the city-based students.

When comparing the larger (population > 20 000) and smaller (< 20 000) rural sites again there was no difference between marks though students were more satisfied with the smaller sites (Denz-Penhey & Murdoch, 2009). This last evaluation used DREEM (the Dundee Ready Education Environment Measure) to explore students' perceptions of their education.

Case study 2: Harvard Medical School

Harvard Medical School (HMS) established a primary care clerkship (PCC) in 1997 as part of curricular revision (Peters et al., 2001). The PCC, which is ongoing, has two goals: to give students the opportunity to care for patients over time (continuity); and to teach them aspects of modern primary care including health promotion and disease prevention, dealing with clinical uncertainty, working in a health care team; and shared decision making with patients. Running over nine months in year 3 of the four-year MD program, the PCC involves each student being attached to one primary care physician. Student spend three afternoons each month at the primary care practice, seeing three to five patients per session with the expectation of their following at least one patient over the nine month placement. On the fourth afternoon they have classroom based lectures and tutorials, while the rest of their time they undertake traditional specialty specific rotations. Peters et al. (2001) summarized the evaluation carried out through paper surveys, small group feedback and interviews and concluded that the PCC was meeting some but not all of its objectives. The majority of students were highly satisfied with their preceptors. In terms of continuity of patient care, data from 103 of the 132 student reports showed that students saw their allocated patient on average 3-4 times. There are no data given on the impact of this continuity nor about student learning as such.

HMS more recently introduced a parallel and alternate third year longitudinal clinical attachment. This change was specifically aimed at reducing the fragmentation of year 3 and it incorporated the PCC as described above but now extended to the full twelve months. Bell et al. (2008) state that this initiative was introduced in the academic year 2005-6 following design and development the previous year; however Ogur et al. (2007) state that the pilot year was 2004-2005. The discrepancy may be due to a time difference in implementation at the three pilot sites: Cambridge Hospital, Beth Israel Deaconess Medical Center (BIDMC) and the Brigham and Women's Hospital. According to Ogur et al. (2007) eight student volunteers undertook the HMS-Cambridge Integrated Clerkship (HMS-CIC) from a year group of 189, while Bell et al. (2008) report on the eight students undertaking the pilot at BIDMC.

The HMS-CIC involved students being paired with various discipline-specific preceptors (medicine, neurology, obstetrics & gynaecology, paediatrics and psychiatry) and spending 5 to 10 hours per week or alternate weeks at the preceptors' ambulatory clinic. Moreover students were allocated 'panels' of patients to follow throughout the year, accompanying them to consultations, through hospital admissions, births, operations and other healthcare visits. The clerkship is described in detail. Evaluation involved both quantitative and qualitative data and the eight pilot students were compared to 11 HMS students undertaking the traditional course, and for summative assessments with the whole year 3 cohort. Findings include that the HMS-CIC performed as well if not better than their peers; in particular their communication skills had improved significantly more by the end of the year. In terms of perceptions the HMS-CIC students found the year more rewarding and less marginalising than the 'controls' and felt that the year had better prepared them to be caring doctors, to deal with ethical dilemmas and to involve patients in decision making. The HMS-CIC students also ranked patients' psychosocial concerns more highly at the end of the year while the traditional students' ranking had decreased.

Bell et al. (2008) also outline the six principles on which the longitudinal clerkships were based: student-centred; provision of patient-centred care; fostering of humanism; provision of longitudinal primary care experience; integration of clinical and basic sciences; emphasis on a cross-disciplinary approach to patient care.

The BIDMC program differed from the HMS-CIC as described in various ways. At the medical centre they rotated through specialties but, compared to students on the 'standard' HMS course, they also followed one obstetrics-gynaecology patient for the year and received continuity of mentorship from faculty including consistent feedback throughout the year and monthly reflection sessions. Their evaluation data compares the eight pilot BIDMC students with eighteen volunteer students following the traditional course. Mid-year focus group data from the pilot students was limited to showing that students felt valuable to the team and less anonymous to hospital faculty. End of year comparisons across both cohorts indicated that pilot students found their experience more humanising, less hectic and less frustrating than traditional students. Pilot students reported they were more prepared to involve patients in decision-making and to deal with uncertainty, and had been more likely to follow their patients after discharge than the traditional students. There were, however, no statistically significant different total examination scores.

Ogur et al. (2009) published further work on four years of the HMS-CIC using an innovative evaluation approach. Out of the 38 students from these 4 years, 14 submitted 16 narratives in total of their experiences, having been asked to consider the way in which the longitudinal care of patients had influenced their learning. The authors themed the work into six clusters: creating a dynamic integrated learning environment; providing a broader understanding of all aspects of illness; permitting a deeper connection with patients; transforming of the student's role in ways that were challenging and empowering; improving patient care and inspiring commitment, advocacy and idealism. These themes are discussed in more detail. The authors draw attention to the limitations of such work from a small proportion of volunteer students without comparison to students on the traditional clerkships.

The most recently published evaluation of the HMS-CIC (Hirsh et al., 2012) compares 27 CIC students with 45 traditional students from three-year cohorts (2004-2007) during the piloting of the clerkship. As with the similar studies, the CIC performed as well or slightly better than their peers on the standard student assessments (NBME – National Board subject exams covering surgery, obstetric & gynaecology, paediatrics, psychiatry). They also scored higher on the year 4 OSCE particularly in history-taking skills. CIC students were more likely to feel they had established meaningful relationships with patients and made differences to patient care, through a patient-centred approach. They reported receiving more feedback and mentoring and were more satisfied with the learning environment, gaining more confidence and feeling less marginalised. Of note is that the CIC students appear to have had less exposure to the hidden curriculum (the authors presumably thus indicating that the hidden curriculum is always a negative influence). CIC students however felt less well prepared to practice in a hospital setting but better prepared for an ambulatory setting

Case study 3: The Parallel Rural Community Curriculum (PRCC)

The PRCC is an initiative of the School of Medicine, Flinders University, Adelaide (South Australia). It began as a pilot project in 1997 to attempt to address Australian rural workforce mal-distribution (with the hypothesis that training students in a rural location would enhance the likelihood of their becoming rural doctors in future) and to challenge the high proportion of clinical placements being undertaken in urban tertiary teaching hospitals. Medical students

undertaking the PRCC spend an entire year in rural general practice with GPs as their primary mentors and facilitators of learning. A number of papers have been published about the PRCC with eight including data about effectiveness: Couper & Worley, 2010; Couper et al., 2011; Stagg et al., 2009; Walters et al., 2011; Worley et al., 1999; Worley et al., 2000; Worley et al., 2004; Worley et al., 2006.

Preliminary evaluation of the first (pilot) year of the PRCC in the Riverland region of SA (Worley et al. 1999 – focus on assessment data; Worley et al., 2000 – description of PRCC and assessment data) involved eight students: two at each of four general practices. The 1997 cohort was from year 5 of the then 6-year undergraduate program (the program changed to a 4-year graduate entry program in 1996). The students participated in the same tutorials and lectures as the rest of their year group via videoconferencing and recorded teaching sessions. Their clinical placement was integrated rather than involving rotations through the five individual specialties: medicine, surgery, psychiatry, obstetrics & gynaecology, psychiatry and paediatrics. While acknowledging the small numbers of PRCC students, the evaluators compared the academic performance of those 8 to the rest of the year group in Adelaide (64 students). Each of the five year 5 specialties had separate end-of-rotation assessments (clinical performance in workplace; OSCE; written examination), which were then combined for an overall subject score. In addition there was an integrated end-of-year OSCE. As the PRCC students did not undertake specialty rotations, their specialty assessments were spread throughout the year. All students regardless of location took the same examinations including the endpoint OSCE.

The PRCC students performed significantly better than their peers when compared to the whole cohort (Worley et al., 2000) and in the aggregated specialty assessments when compared to eight matched controls (Worley et al., 1999).

After the change to a 4-year graduate entry program, the PRCC was undertaken in year 3. A further comparison of student competence from four consecutive years (1998-2002) was reported in 2004 (Worley et al., 2004). From 1998 as well as being based in Adelaide at the tertiary hospital or on the PRCC, 16 students undertook the whole of year 3 in Darwin (the capital of the Northern Territory) in the regional secondary referral hospital. The four years worth of assessment data relates to 371 students in total (262 Adelaide; 40 PRCC, Darwin 48). Mean assessment scores at the end of year 2 during which all the students were based in Adelaide were similar. Mean assessment scores for year 3 however differed significantly even though all students undertook the same examinations. The PRCC group and the Darwin group performed significantly better than the Adelaide students. The authors discuss possible limitations in that the students were not randomised to the three groups but chose their rotations. However the selection into medical school was the same process for all students, and their marks were similar in the first two years of the course.

As well as this quantitative evaluation (Kirkpatrick level 2b), Worley et al. (2006) carried out a case study using an interpretivist perspective (though the interviews are described as structured) of 6 PRCC and 16 students from Adelaide in year 3 in 1998. The main differences that came from the analysis of the transcripts were: PRCC students felt involved in the care of patients (and referred to 'my' patients') while Adelaide students varied in the amount of participation with some feeling they provided mainly social support for patients and some feeling they provided no useful role (and referred to 'a' patient). Both sets of students highlighted the importance of patient contact for learning. The Adelaide students reported competition to gain access to patients and hierarchical relationships in the clinical setting. The PRCC students described a more collegial environment and fewer students competing for either

patients or supervisor time. The PRCC students also felt valued rather than being seen as inconvenient or in the way as the Adelaide students did. Not all the PRCC students however found the integrated nature of the placement helpful for learning and were anxious they would have gaps in knowledge. But they did acknowledge the continuity of care that was possible in the rural location – continuity with patients and continuity of mentoring from their GP and other staff.

An external international evaluation of the PRCC was undertaken in 2006 and the data explored further in 2009 with the aim of exploring how the location of clinical training might be expanded, how graduating doctors might be motivated to work in rural practice, how the private health sector might be enticed to engage in student teaching and what would reduce the loss of humanism by students during their training (Couper & Worley, 2010). Data were extracted from interviews and focus groups with staff, academics, GPs, other specialists, managers and students. In terms of learning, students highlighted the importance of continuity in respect of both patients and tutors, and their ability to live in the same environment as the patients. The fact that they contributed directly to care rather than simply being observers allowed students to take responsibility. The findings of this external evaluation were then compared to an evaluation of a similar program introduced at the Northern Ontario School of Medicine (NOSM) in 2007 (Couper et al., 2011). There were similar conclusions from both schools: the programs help produce confident and skilful students through continuity of relationships, teamwork and workplace participation. There were also similar concerns about potential student isolation, ensuring sufficient experience in all specialties, the need for adequate support and optimal communication.

Further evaluation of the PRCC involved semi-structured interviews not only with students (7) but also GP preceptors (21) and practice managers (4) at three points during one academic year (Walters et al., 2011). This paper contains a lot of data categorised into themes with an emphasis on the maturation of the doctor-student relationship during the longitudinal clinical attachment. The GPs observed that supervising students took less time as the year progressed as students took on more responsibility and more complex consultation tasks. The teacher-student relationship matured as students increased their participation in authentic clinical work, and this was a result of the length of that relationship in this model.

Stagg et al. (2009) focused their evaluation on whether the PRCC had achieved the original aim of enticing graduates to practise in rural areas. Their survey of 86 graduates had a response rate of 53%. The data suggested that the PRCC graduates as doctors could be divided into four groups (rural background, rural career; rural background, urban career; urban background, rural career; urban background, urban career) but that overall 50% of respondents were on a rural career training pathway - higher percentage than the Flinders graduates in total.

Case study 4: The Rural Physician Associate Program (RPAP)

The RPAP of Minnesota began in 1971 funded by state money and with the aim of increasing the number of primary care physicians practicing in rural locations. The RPAP is defined as a community-based, clinical continuity of care experience and takes place over nine months in year 3. Students are mentored by primary care doctors in one of 110 rural communities. Continuity of care and responsibility are important components of the program, with students following up patients they assess. Verby (1988) described the program in detail. The number of students placed in any one year varies from 19 to 46, with an average of 33 between 1971 and 2007. Six papers provide evaluation data relating to the RPAP: Halaas et al., 2008; Power et al., 2006; Verby, 1988; Zink et al., 2008; Zink et al., 2010a; Zink et al., 2010b.

The impact of the program was such that by 1988 all 87 counties in Minnesota had an acceptable ratio of one general physician per 2500 people or better for the first time in the state's history (Verby, 1988) but there are no details about whether the number of medical students graduating had increased over this time period. The first basic evaluation data are from 1988 and are based on students' written summaries of their experiences and their suggested improvements. There are also comparisons between RPAP and non-RPAP students' performance on the then national board examinations (NBME) parts I and II for seven of the 16 years of the program. There were no statistical differences in marks (though no figures are given in the paper).

In 2003 a comparison of RPAP (33) with one cohort of non-RPAP students (35 of a total year group of 180 who did not do the RPAP) undertaking a 15-station OSCE at the beginning of year 4 was reported (Power et al., 2006). The OSCE assessed the primary care clerkship: an 8-week course for non-RPAP students but integrated throughout nine months for RPAP students. On each station there was no significant difference between the two student groups except for the case of a middle-aged woman with chest pain for which the RPAP students gained significantly higher marks. A further comparison between RPAP (201) and non-RPAP (1129) students also showed no major differences for the USMLE step one and step two examinations or an OSCE except for the obstetrics & gynaecology components (Zink et al., 2010a). The RPAP students scored significantly lower in this specialty even though other data suggested that RPAP students performed more deliveries. The authors suggest that the reason for this result was that RPAP students focus more on clinical work than reading and have less structured lecture time. Zink et al. (2010b) evaluated the OSCE performance of 16 students (8 RPAP and 8 non-RPAP) via videotaping the stations and qualitatively analysing the interactions. The students for this analysis were chosen from the high and low ends of the marking spectrum. The RPAP students were found to be more consistent in rapport building and were better at discussing prevention and screening, while following an effective consultation pattern. Non-RPAP students were better interacting with an adolescent patient using a mnemonic taught during a lecture.

Halaas et al. (2008) focus on the career outcomes of the RPAP students. Data from 1175 medical students who had undertaken the RPAP of whom 901 were practising physicians suggested that the longitudinal rural community experience had increased the number of students choosing to practice as primary care physicians in a rural setting. The authors postulated that a major factor in this choice and a key component of the RPAP is the length of the rotation and its longer-term nature compared to the standard program. This hypothesis was based results from three consecutive cohorts of RPAP students' essays about their experience (Zink et al., 2008). These were thematically analysed and findings included: students' learning from patients; the teaching being learner-centred; trust developing between student and physicians as well as patients; and learning being hands-on. Students cared for large numbers of patients compared to their urban peers and were said to gain confidence, autonomy and competence from their experience.

Appendix 5

Authors	Source/ database	Journal	Year	Loctn	No of students in evaluation + number in comparison if applic	Students	No of others/ non students	Year of prog	Location	Time	Year began	Groups	Timing of eval	K level	Strength of findings	Grade of paper
Bell et al.	J Scan	Acad Med	2008	N Am	8 + 18	volunteers	0	3	Hospital/ ambulatory	Full yr	2005	Single + comparison	pre/post	1, 2a, 2b	3	good
Berger & Schaffer	J Scan	J Med Ed	1986	N Am	10 + 24	volunteers	0	2	Primary care	0.5 days a week for 9/12	1984	2 yr groups	post	1	3	acceptable
Carney et al.	Pub Med	Acad Med	2005	N Am	155	required	0	2	Academic medicine clinic	4 hrs /wk for 2 yrs	2002	2 yrs, 3 sites	post	2b	4	excellent
Couper et al.	Pub Med	MJA	2011	Aus/ N Am	45 Aus; 7 N Am	volunteers	87 Aus 39 N Am	3	Rural	Full yr	1997/2008	2 sites	post	1	2	acceptable
Couper & Worley	Pub Med	RRH	2010	Aus	45	volunteers	87	3	Rural	Full yr	1997	Single year	post	1	3	excellent
Denz-Penhey & Murdoch	J Scan	Med Teach	2008	Aus	8	volunteers	0	5	Rural	Full yr	2003	2 yr groups	post	1	3	acceptable
Denz-Penhey & Murdoch	PubMed	Med Teach	2009	Aus	60 (of 62)	volunteers	27 acad 15 admin	5	Rural	Full yr	2003	single	mid/post	1	3	good
Denz-Penhey & Murdoch	PubMed	RRH	2010	Aus	245	volunteers	0	5	Rural	Full yr	2003	Multi yrs + comp	post	1. 2b	4	good
Denz-Penhey et al.	J Scan	RRH	2004	Aus	21	volunteers	Staff	5	Rural	Full yr	2003	Staff & students - single yr	mid/post	1	2	acceptable
Denz-Penhey et al.	WoK	RRH	2005	Aus	29 WA 25 SA	Volunteers	0	5	Rural	Full yr	2003	2 yrs 2 sites (WA & SA)	mid/post	1	3	acceptable
Deterding et al.	Pub Med	Arch Ped Ad	1999	N Am	109 of whom 56 did self- assessment	required	0	1 st 3 years	Community	0.5 days per wk , 3.52 per month	1996	1 yr group	post	2b	3	acceptable
Florence et al.	Pub Med	J Rur Health	2007	N Am	58 of 84 + 72 of 168 comparison	volunteers	0	1 st 3 years	Community	1/7 per week	1990-2002	Multi yrs + comparison	post	1, 2a career	2	acceptable
Fraterelli & Kamamoto	Pub Med	Am J Ob Gyn	2004	N Am	40 + 227	volunteers	0	3	Community	0.5 days per wk for 5/12	1995	Multi yr + comp	post	2b	3	good
Geyman et al.	Embase	J Fam Pract	1984	N Am	44	elective	0	2	Family medicine	0.5 days per wk 9/12	1974	single	post	1	3	acceptable
Hadac et al.	Pub Med	J Med Ed	1979	N Am	30	elective	0	2	Family medicine	0.5 days per wk 9/12	1974	single	post	1	2	acceptable
Halaas et al.	Pub Med	Ed Rur Pract	2008	N Am	1175	Apply + interview	0	3	Rural	9/12	1971	Multi yrs	post	Career	3	acceptable
Hamilton et al.	WoK	J Nat MA	1998	N Am	29	required	0	3 & 4	Primary care	0.5 days per wk for 40 wks	1995	single	post	1	2	poor
Hauer et al.	J Scan	Med Ed	2012	N Am	29 + 25	Rank preference most receive 1 st choice	0	3	Ambulatory clinics	Full yr		3 sites + comparison	mid/post	1	3	good
Henley et al.	Pub Med	Acad Med	2000	N Am	97	required	0	2-4	Primary care	0.5 to 1 day per wk for 3 yrs	1986	Yr 3 & yr 4 students of 1 yr	post	1	2	acceptable
Herold et al.	Pub Med	Acad Med	1993	N Am	93 + 108	volunteers	0	All 4	Community	0.5 days per wk	1987	Multi yrs + comp	post	career	3	poor
Hirsh et al.	J Scan	Acad Med	2012	N Am	27 + 45	volunteers	0	3	Different venues of care	Full year	2004	Multi yr + comp	Pre/post	2a, 2b	3	good
Konkin & Suddards	Pub Med	AHSE	2012	N Am	25 of 33	volunteers	0	3	Community	9/12	2007	Multi yrs	post	1, 2a	3	acceptable

Lemon et al.	Pub Med	Acad Med	1995	N Am	180 students per year but not clear how many evaluated	Volunteers for 1 st 3 yrs, then mandatory	Evaluation includes staff but numbers not given	Yrs 1-3	Primary care	Variable over the 3 yrs	Mandatory from 1994	Multi yrs	post	1	2	poor
Lewin et al.	Pub Med	Acad Med	1999	N Am	24 + 81	volunteers	0	1-3	Primary care	Yr 3 full year	1994	Single + comp	post	1, 2b	3	good
Lubetkin et al.	Pub Med	Acad Med	1999	N Am	114	required	74/114 precept	1	Community	0.5 days per week	1996	single	mid/post	1, 2b	3	acceptable
McLaughlin et al.	J Scan	Acad Med	2011	N Am	27 + 108	voluntary	0	3	Family medicine	9-12/12		3 sites 1 year + comp	post	2b	3	acceptable
Mihalynuk et al.	Pub Med	Med Ed	2008	N Am	6 + 6	voluntary	0		Family medicine	Full yr		Single + comp	mid/post	1	4	acceptable
O'Brien et al.	J Scan	Med Ed	2012	N Am	Dakota: 9 + 25 HMS: 10 + 50 UCSF: 16 + 39	Volunteers through ranking	0	3	3 different models	Full yr or block	1994 Dakota 2004 HMS 2007 UCSF	3 sites	mid/post	observ ed	3	acceptable
Ogur & Hirsh	J Scan	Acad Med	2009	N Am	14 of 38	volunteers	0	3	Urban ambulatory clinic	Full yr	2004	single	post	1, 2a,	3	acceptable
Ogur et al., 2007	J Scan	Acad Med	2007	N Am	8 + 11	volunteers	0	3	Urban ambulatory clinic	Full yr	2004	Single + comp	mid/post	1, 2a, 2b	3	excellent
Oswald et al.	Pub Med	Med Ed	2001	UK	13 of 14	volunteers	0	3 & 4	General practice	15/12	1993	Multi yrs + comp	mid/post	1, 2b	4	good
Peters et al.	Pub Med	Acad Med	2001	N Am	103 of 132	required	0	3 & 4	Primary care	9/12	1997	single	post	1	3	acceptable
Poncelet et al.	Pub Med	Med Ed	2011	Aus	23 + 206	Volunteers lottery	0	3	Family medicine	Full yr	2007	2 yrs + comp	mid/post	1, 2b	4	good
Poole et al.	CINAHL	RRH	2010	Aus	19	volunteers	0	5	Rural	27/52	2008	Single + comp	post	1, 2b	3	acceptable
Power et al.	J Scan	T Learn Med	2006	N Am	33 + 35	Selected volunteers	0	3	Rural	9/12	1971	Single + comp	post	1, 2b	3	good
Prislin et al.	Pub Med	Acad Med	1998	N Am	429	Not clear	0	Variable	Primary care	12/12; 21/12; 44/12		5 sites	post	1	3	acceptable
Ringdahl et al.	Pub Med	Fam Med	2009	N Am	40 + 128	Volunteers + selection	0	4	Family medicine	0.5 days per wk	1992	Multi yrs + comp	post	1, 2b	4	good
Schauer & Schieve	Pub Med	Acad Med	2006	N Am	29 + 296	volunteers	0	3	Rural	7/12	1998	Multi yrs + comp	post	2b	4	good
Stagg et al.	Pub Med	RRH	2009	Aus	46 of 86	volunteers	0	3	Rural	Full yr	1997	Multi yrs	post	career	4	acceptable
Teherani et al.	J Scan	Acad Med	2009	N Am	0	Volunteers + lottery	27 precep	3	Ambulatory integrated clinics	Full yr	2007	single	post	Precept or reaction	4	good
Verby	J Scan	J Med Ed	1988	N Am	462	volunteers	0	3	Rural	Full yr	1971	Multi + comp	post	1, 2b	4	good
Von Below et al.	Pub Med	BMC Med Ed	2008	Eur	60 of 86	required	15 facil	1 & 2	Primary & secondary care	2 yrs	2001	Single yr	post	1	4	acceptable
Walmsley et al.	Pub Med	Med Ed	2009	N Am	12 of 154	required	0	3	Primary care	Full yr	2006?	Single yr	post	1	4	good
Walters et al.	Pub Med	Med Ed	2011	Aus	7	volunteers	21 GPs 4 pr man	4	Rural	Full yr	1997	Single yr	During/post	1	3	acceptable
Wee et al.	Pub Med	Acad Med	2011	Asia	222/240	volunteers	355 patients	variable	Community	6/12 - x2 per wk and then monthly	2009	Single yr	post	2b, 4	3	acceptable
Wilson & Cleland	Pub Med	RRH	2008	UK	14	volunteers	0	4	Rural	Full yr	2006	Single + comp	Pre/post	1, 2b	3	acceptable
Worley et al.	J Scan	BMJ	2004	Aus	40 + 331	volunteers	0	3	Rural	Full yr	1997	Multi yrs + comp	post	2b	4	good
Worley & Lines	J Scan	Med Teach	1999	Aus	8	volunteers	0	3	Rural	Full yr	1997	Single + comp	post	2b	3	acceptable
Worley et al.	J Scan	Med Ed	2006	Aus	6 + 16	volunteers	0	3	Rural	Full yr	1997	Single + comp	During/post	1	3	good

Worley et al.	Embase	Med Ed	2000	Aus	8	volunteers	0	3	Rural	Full yr	1997	Single + comp	During/post	1, 2b	3	good
Zink et al.	Pub Med	JRH	2008	N Am	95	volunteers	0	3	Rural/family med	9/12	1971	Multi yrs	post	1	3	acceptable
Zink et al.	Pub Med	Fam Med	2010a	N Am	201 + 1129	volunteers	0	3	Rural/family med	9/12	1971	Multi + comp	post	2b	4	excellent
Zink et al.	Pub Med	Fam Med	2010b	N Am	8 high performers 8 low performers	volunteers	0	3	Rural/family med	9/12	1971	2 groups from single yr	post	2b	4	good

References

- Barr H, Koppel I, Reeves S, Hammick M, Freeth D. 2005. Effective interprofessional education. Argument, assumption and evidence. Oxford: Blackwell Publishing.
- Dornan T, Littlewood S, Margolis S, Scherpbier A, Spencer J, Ypinazar. 2005. BEME Guide 6. How can experience in clinical and community settings contribute to early medical education? Available at: http://www.bemecollaboration.org/downloads/764/early_experience_review.pdf
- Hart JT. 1985. The world turned upside down: Proposals for community based undergraduate medical education. *J R Coll Gen Pract* 3: 63-68.
- Hirsh DA, Ogur B, Thibault GE, Cox, M. 2007. New models of clinical clerkships: "Continuity" as an organizing principle for clinical education reform. *N Engl J Med* 356: 858-866.
- Hauer KE, O'Brien B, Poncelet AN. 2009. Longitudinal, integrated clerkship education: Better for learners and patients. *Acad Med* 84: 821.
- Kirkpatrick D. 1967. Evaluation of Training. In Craig R, Bittel L, editors. *Training and Development Handbook*. New York: McGraw- Hill. p. 131-167.
- Knowles M. 1978. *The adult learner: A neglected species*. Oxford, England: Gulf Publishing.
- Krupat E, Hiam CM, Fleming MZ, Freeman P. 1999. Patient-centeredness its correlates among first year medical students. *Int J Psychiatry Med* 29:347-56.
- Lave J, Wenger E. 1991. *Situated learning: legitimate peripheral participation*. Englewood Cliffs, NJ: Prentice Hall.
- Norris TE, Schaad DC, deWitt D, Ogur B, Hunt D. 2009. Longitudinal integrated clerkships for medical students: An innovation adopted by medical schools in Australia, Canada, South Africa and the United States. *Acad Med* 84: 902-907.
- Ogrinc G, Mutha S, Irby D. 2002. Evidence for longitudinal ambulatory care rotations: a review of the literature. *Acad Med* 77: 688-693.
- Oswald N. 1989. Why not base clinical education in general practice? *Lancet* 334: 148-149.
- Oswald N, Alderson T, Jones S. 2001. Evaluating primary care as a base for medical education: The report of the Cambridge Community-based Clinical Course. *Med Educ* 35: 782-788.
- Oswald N, Jones S, Date J, Hinds D. 1995. Long-term community-based attachments: The Cambridge course. *Med Educ* 29: 72-76.
- Pawson R, Greenhalgh T, Harvey G, Walshe K. 2005. Realist review – a new method of systematic review designed for complex policy interventions. *Journal of Health Services Research & Policy* 10 Suppl 1: 21-34
- Pawson R, Tilley N. 1997. *Realist Evaluation*. London: Sage.
- Shortell S. 1976. Continuity of medical care: Conceptualisation and Measurement. *Med Care* 14: 377-391.
- Thistlethwaite JT, Kidd MR, Hudson JN. 2007. General practice: A leading provider of medical student education in the 21st century? *Med J Aust* 187: 124-128.
- Walters L, Greenhill J, Richards J, Ward H, Campbell N, Ash J, Schuwirth LT. (2012). Outcomes of longitudinal integrated clinical placements for students, clinicians and society. *Medical Educ* 46: 1028-1041.
- Wong G, Greenhalgh T, Westhorp G, Pawson R. (2012). Realist methods in medical education research: what are they and what can they contribute? *Med Educ* 46: 89-96.
- Worley P, Silagy C, Prideaux D, Newble D, Jones A. 2000. The parallel rural community curriculum: An integrated clinical curriculum based on rural general practice. *Med Educ* 34: 558-565.
- Yardley S, Dornan T. 2012. Kirkpatrick's levels and education 'evidence'. *Med Educ* 29: 97-106.

REFERENCES INCLUDED IN THE REVIEW PROCESS

1. Bell SK, Krupat E, Fazio FB, Roberts DH, Schwartzstein RM. 2008. Longitudinal pedagogy: a successful response to the fragmentation of the third-year medical student clerkship experience. *Acad Med* 83(5): 467-475.
2. Berger A, Schaffer S. 1986. An interdisciplinary continuity of care experience for preclinical medical students. *J Med Educ* 61: 771-773.
3. Carney PA, Ogrinc G, Harwood BG, Schiffman JS, Cochran N. 2005. The influence of teaching setting on medical students' clinical skills development: is the academic medical center the "gold standard"? *Acad Med* 80(12): 1153-1158.
4. Couper I, Worley PS, Strasser R. 2011. Rural longitudinal integrated clerkships: lessons from two programs on different continents. *Rural Remote Health* 11(2): 1665.
5. Couper ID, Worley PS. 2010. Meeting the challenges of training more medical students: lessons from Flinders University's distributed medical education program. *Med J Aust* 193(1): 34-36.
6. Denz-Penhey H, Murdoch JC. 2008. Rural learning is more than marks: sensitised to knowledge. *Med Teach* 30(8): 781-786.
7. Denz-Penhey H, Murdoch JC. 2009. 'It's really, really good, but it could be a lot better': qualitative evaluation of a rural clinical school, four years on. *Med Teach* 31(10): e443-448.
8. Denz-Penhey H, Murdoch JC. 2010. Is small beautiful? Student performance and perceptions of their experience at larger and smaller sites in rural and remote longitudinal integrated clerkships in the Rural Clinical School of Western Australia. *Rural Remote Health* 10(3): 1470.
9. Denz-Penhey H, Murdoch JC, Lockyer-Stevens VJ. 2004. 'What makes it really good, makes it really bad.' An exploration of early student experience in the first cohort of the Rural Clinical School in the University of Western Australia. *Rural Remote Health* 4(3): 300.
10. Denz-Penhey H, Shannon S, Murdoch CJ, Newbury JW. 2005. Do benefits accrue from longer rotations for students in Rural Clinical Schools? *Rural Remote Health* 5(2): 414.
11. Deterding R, Kamin C, Barley G, Adams L, Dwinnell B, Merenstein G. 1999. Effect of a longitudinal course on student performance in clerkships. *Arch Pediatr Adolesc Med* 153(7): 755-760.
12. Florence JA, Goodrow B, Wachs J, Grover S, Olive KE. 2007. Rural health professions education at East Tennessee State University: survey of graduates from the first decade of the community partnership program. *J Rural Health* 23(1): 77-83.
13. Frattarelli LC, Kamemoto LE. 2004. Obstetrics and gynecology medical student outcomes: longitudinal multispecialty clerkship versus traditional block rotations. *Am J Obstet Gynecol* 191(5): 1800-1804.
14. Geyman C, Smith L, Hadac R et al. 1984. Benefits of predoctoral experiences in longitudinal patient care. *J Fam Pract* 18: 911-914.
15. Hadac RR, Smith CK, Gordon MJ. 1979. Can continuity of medical care be taught? *J Med Educ* 54(7): 525-533.
16. Halaas GW, Zink T, Finstad D, Bolin K, Center B. 2008. Recruitment and retention of rural physicians: outcomes from the rural physician associate program of Minnesota. *J Rural Health* 24(4): 345-352.
17. Hamilton LA, Towns AB, Edelstein RA. 1998. Development of a multidisciplinary primary care program at the Drew University of Medicine and Science. *J Natl Med Assoc* 90(4): 205-207.
18. Hauer KE, Hirsh D, Ma I, Hansen L, Ogur B, Poncelet AN, Alexander EK, O'Brien BC. 2012. The role of role: learning in longitudinal integrated and traditional block clerkships. *Med Educ* 46(7): 698-710.
19. Henley E, Glasser M, May J. 2000. Medical student evaluation of family nurse practitioners as teachers. *Fam Med* 32(7): 491-494.

20. Herold AH, Woodard LJ, Pamies RJ, Roetzheim RG, Van Durme DJ, Micceri T. 1993. Influence of longitudinal primary care training on medical students' specialty choices. *Acad Med* 68(4): 281-284.
21. Hirsh D, Gaufberg E, Ogur B, Cohen P, Krupat E, Cox M et al. 2012. Educational outcomes of the Harvard Medical School-Cambridge Integrated Clerkship: A way forward for medical education. *Acad Med* 87(5): 643-650.
22. Konkin J, Suddards C. 2012. Creating stories to live by: caring and professional identity formation in a longitudinal integrated clerkship. *Adv Health Sci Educ Theory Pract* 17(4): 585-596.
23. Lemon M, Yonke A, Roe B, Foley R. 1995. Communication as an essential part of program and institutional development. *Acad Med* 70(10): 884-886.
24. Lewin LO, Papp KK, Hodder SL, Workings MG, Glover P, Headrick LA. 1999. Performance of third-year primary-care track students in an integrated curriculum at Cae Western Reserve University. *Acad Med* 74(1 Suppl): S82-S89.
25. Lubetkin EI, Krackov SK, Storey-Johnson C. 1999. The use of questionnaires to assess achievement of course goals in medical students' longitudinal community-based clinical experiences. *Acad Med* 74(12): 1316-1319.
26. McLaughlin K, Bates J, Konkin J, Woloschuk W, Suddards CA, Regehr G. 2011. A comparison of performance evaluations of students on longitudinal integrated clerkships and rotation-based clerkships. *Acad Med* 86(10 Suppl): S25-29.
27. Mihalyuk T, Bates J, Page G, Fraser J. 2008. Student learning experiences in a longitudinal clerkship programme. *Med Educ* 42: 729-732.
28. O'Brien BC, Poncelet AN, Hansen L, Hirsh DA, Ogur B, Alexander EK, Krupat E, Hauer KE. 2012. Students' workplace learning in two clerkship models: a multi-site observational study. *Med Educ* 46(6): 613-624.
29. Ogur B, Hirsh D. 2009. Learning through longitudinal patient care-narratives from the Harvard Medical School-Cambridge Integrated Clerkship. *Acad Med* 84(7): 844-850.
30. Ogur B, Hirsh D, Krupat E, Bor D. 2007. The Harvard Medical School-Cambridge integrated clerkship: an innovative model of clinical education. *Acad Med* 82(4): 397-404.
31. Oswald N, Alderson T, Jones S. 2001. Evaluating primary care as a base for medical education: the report of the Cambridge Community-based Clinical Course. *Med Educ* 35(8): 782-788.
32. Peters AS, Feins A, Rubin R, Seward S, Schnaidt K, Fletcher RH. 2001. The longitudinal primary care clerkship at Harvard Medical School. *Acad Med* 76(5): 484-488.
33. Poncelet A, Bokser S, Calton B, Hauer KE, Kirsch H, Jones T, Lai CJ, Mazotti L, Shore W, Teherani A, Tong L, Wamsley M, Robertson P. 2011. Development of a longitudinal integrated clerkship at an academic medical center. *Med Educ Online* 16.
34. Poole P, Bagg W, O'Connor B, Dare A, McKimm J, Meredith K, Shulruf B, Reid P. 2010. The Northland Regional-Rural program (Pukawakawa): broadening medical undergraduate learning in New Zealand. *Rural Remote Health* 10(1): 1254.
35. Power DV, Harris IB, Swentko W, Halaas GW, Benson BJ. 2006. Comparing rural-trained medical students with their peers: performance in a primary care OSCE. *Teach Learn Med* 18(3): 196-202.
36. Prislun MD, Feighny KM, Stearns JA, Hood J, Arnold L, Erney S, Johnson L. 1998. What students say about learning and teaching in longitudinal ambulatory primary care clerkships: a multi-institutional study. *Acad Med* 73(6): 680-687.
37. Ringdahl E, Kruse RL, Lindbloom EJ, Zweig SC. 2009. The University of Missouri integrated residency: evaluating a 4-year curriculum. *Fam Med* 41(7): 476-480.
38. Schauer RW, Schieve D. 2006. Performance of medical students in a nontraditional rural clinical program, 1998-99 through 2003-04. *Acad Med* 81(7): 603-607.
39. Stagg P, Greenhill J, Worley PS. 2009. A new model to understand the career choice and practice location decisions of medical graduates. *Rural Remote Health* 9(4): 1245.

40. Teherani A, O'Brien BC, Masters DE, Poncelet AN, Robertson PA, Hauer KE. 2009. Burden, responsibility, and reward: preceptor experiences with the continuity of teaching in a longitudinal integrated clerkship. *Acad Med* 84(10 Suppl): S50-53.
41. Verby JE. 1988. The Minnesota Rural Physician Associate Program for medical students. *J Med Educ* 63(6): 427-437.
42. von Below B, Hellquist G, Rödger S, Gunnarsson R, Björkelund C, Wahlqvist M. 2008. Medical students' and facilitators' experiences of an Early Professional Contact course: active and motivated students, strained facilitators. *BMC Med Educ* 8: 56.
43. Walmsley MA, Dubowitz N, Kohli P, Cooke M, O'Brien BC. 2009. Continuity in a longitudinal out-patient attachment for year 3 medical students. *Med Educ* 43: 895-906.
44. Walters L, Prideaux D, Worley P, Greenhill J. 2011. Demonstrating the value of longitudinal integrated placements to general practice preceptors. *Med Educ* 45(5): 455-463.
45. Wee LE, Koh GC, Lim VK. 2011. Caring for underserved patients through neighborhood health screening: outcomes of a longitudinal, interprofessional, student-run home visit program in Singapore. *Acad Med* 86(7): 829-839.
46. Wilson M, Cleland J. 2008. Evidence for the acceptability and academic success of an innovative remote and rural extended placement. *Rural and Remote Health* 8: 960 online.
47. Worley P, Esterman A, Prideaux D. 2004. Cohort study of examination performance of undergraduate medical students learning in community settings. *BMJ* 328(7433): 207-209.
48. Worley P, Lines D. 1999. Can specialist disciplines be learned by undergraduates in a rural general practice setting? Preliminary results of an Australian pilot study. *Med Teach* 21: 482-484.
49. Worley P, Prideaux D, Strasser R, Magarey A, March R. 2006. Empirical evidence for symbiotic medical education: a comparative analysis of community and tertiary-based programmes. *Med Educ* 40(2): 109-116.
50. Worley P, Silagy C, Prideaux D, Newble D, Jones A. 2000. The parallel rural community curriculum: an integrated clinical curriculum based in rural general practice. *Med Educ* 34(7): 558-565.
51. Zink T, Halaas GW, Finstad D, Brooks KD. 2008. The rural physician associate program: the value of immersion learning for third-year medical students. *J Rural Health* 24(4): 353-359.
52. Zink T, Power DV, Finstad D, Brooks KD. 2010a. Is there equivalency between students in a longitudinal, rural clerkship and a traditional urban-based program? *Fam Med* 42(10): 702-706.
53. Zink T, Power DV, Olson K, Harris IB, Brooks KD. 2010b. Qualitative differences between traditional and rural-longitudinal medical student OSCE performance. *Fam Med* 42(10): 707-711.