The impacts of observing flawed and flawless demonstrations on clinical skill learning

Kurt Domuracki,[1] Arthur Wong,[2] Lori Olivieri[1] & Lawrence E M Grierson[2,3,4]

OBJECTIVES Clinical skills expertise can be advanced through accessible and cost-effective video-based observational practice activities. Previous findings suggest that the observation of performances of skills that include flaws can be beneficial to trainees. Observing the scope of variability within a skilled movement allows learners to develop strategies to manage the potential for and consequences associated with errors. This study tests this observational learning approach on the development of the skills of central line insertion (CLI).

METHODS Medical trainees with no CLI experience (n = 39) were randomised to three observational practice groups: a group which viewed and assessed videos of an expert performing a CLI without any errors (F); a group which viewed and assessed videos that contained a mix of flawless and errorful performances (E), and a group which viewed the same videos as the E group but were also given information concerning the correctness of their assessments (FA). All participants interacted with their observational videos each day for 4 days. Following this period, participants returned to the laboratory and performed a simulation-based insertion, which was assessed using a standard checklist and a global rating scale for the skill. These ratings served as the dependent measures for analysis.

RESULTS The checklist analysis revealed no differences between observational learning groups (grand mean $\pm$ standard error: $[20.3 \pm 0.7]/25$). However, the global rating analysis revealed a main effect of group ($d.f._{2,36} = 4.51$, $p = 0.018$), which describes better CLI performance in the FA group, compared with the F and E groups.

CONCLUSIONS Observational practice that includes errors improves the global performance aspects of clinical skill learning as long as learners are given confirmation that what they are observing is errorful. These findings provide a refined perspective on the optimal organisation of skill education programmes that combine physical and observational practice activities.

Medical Education 2015: 49: 186–192
doi: 10.1111/medu.12631

Discuss ideas arising from the article at "www.mededuc.com discuss".

[1]Department of Anaesthesia, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada
[2]MD Undergraduate Program, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada
[3]Department of Family Medicine, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada
[4]Program for Educational Research and Development, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada

Correspondence: Lawrence E M Grierson, Department of Family Medicine, McMaster University, McMaster Innovation Park, 175 Longwood Road South, Suite 201A, Hamilton, Ontario L8P 0A1, Canada. Tel: 00 1 905 525 9140;
E-mail: lawrencegrierson@gmail.com

INTRODUCTION

In response to reductions in resident duty hours and society's increased focus on patient safety, health professions training programmes have turned to simulation to meet their educational needs. This pedagogical approach allows for the practice and assessment of a wide spectrum of skills in environments that are not beset by the risks associated with training directly in the clinical setting.[1,2] However, the potential of this approach for clinical education is limited by the temporal and spatial constraints associated with operating a simulation centre.[3] To address these constraints, health professions education researchers have started to explore the utility of video-based observational learning as a way to extend the impact and efficiency of simulation-based training.[4,5]

The observational learning approach capitalises on the influence of observing demonstrations on the acceleration of skill learning.[6,7] It is widely understood that the acquisition of new precision technical skills involves the development of internal representations of action that describe the amplitude and timing of the neural inputs required to achieve desired movement outcomes. When a learner practises a precision skill, he or she is provided with important afferent and efferent stimuli, which provide the bases for building, updating and storing these representations for use at a later time and under a variety of different conditions.[8] Importantly, a portion of the same neural reorganisation that occurs during physical practice also occurs during skill observation.[9] Consequently, the act of watching the performance of a skill can contribute to the development of that skill.

Recent empirical evidence suggests that the impulse timing and magnitude information at the foundation of internal skill representations is organised in a strategic manner, which accounts for the inherent variability of the neuromuscular system,[10,11] and which protects against the consequences associated with certain types of error.[12,13] The implication for observational learning is that skill acquisition is enhanced when trainees have the opportunity to observe skill demonstrations that include flawed performances.[14] This is based on the theory that observing the scope of variability that surrounds a skilled movement provides learners with the information necessary to develop error detection and mitigation strategies. Essentially, errorless demonstrations show the learner how to perform the skill, and flawed demonstrations show him or her how not to perform the skill. There is evidence to suggest that this 'mixed' observation of both flawless and flawed demonstrations results in superior motor skill learning compared with either flawless or flawed skill observation.[15]

The current study used the medical simulation environment to explore the benefits of observational learning with errors on the development of the skill of central line insertion (CLI). Central line insertion is a critical, high-risk medical procedure associated with significant morbidity and mortality, which is performed by a wide spectrum of practitioners in both emergency and elective situations, and which involves the ultrasound-assisted insertion of a catheter into the internal jugular vein for the administration of resuscitative fluids, medications or nutrition.[16] Recently, consensus guidelines regarding the teaching of CLI and recommendations for the composition of CLI teaching activities have evolved.[17] Typically, CLI is taught at the bedside; however, this approach is slowly giving way to simulation experiences, which have led to lower rates of complications in CLIs performed by novice practitioners.[18,19]

To accomplish this, medical student trainees with no CLI experience were randomised into three observational learning groups that differed with respect to the quality of the performances they observed: one group viewed only an errorless video-based performance, and two groups viewed a mix of one errorless and three flawed video-based performances. We hypothesised that observational learning that includes errorful demonstrations will result in better learning as measured by a criterion test after a period of post-observation retention. Of secondary interest, however, is the issue of whether any effect of error observation requires that the errorful components are recognised as such. Thus, as part of the observational intervention, each participant was also required to fill out a standardised checklist and global rating scale (GRS) for each CLI performance he or she watched. After completing the rating scales, members of one of the mixed demonstration groups compared their own ratings against a correct reference checklist and GRS rating for the observed performance. The other mixed demonstration group received no feedback. The flawless demonstration group also completed rating scales.[14]

METHODS

Participants

Thirty-nine medical students (mean age: 27.7 years; 17 women, 22 men) from McMaster University with no previous CLI experience, simulated or otherwise, volunteered to participate in this study. All participants provided informed consent in accordance with the guidelines set out by the Hamilton Health Sciences/Faculty of Health Sciences Research Ethics Board (HHS/FHS REB) and the Declaration of Helsinki (1954).

Protocol

Before beginning the educational intervention, all participants viewed a standard instructional video detailing the correct procedure for CLI[16] and took part in a short (~ 1 hour) expert lead tutorial, which included CLI kit familiarisation and ultrasound-guided needle localisation practice on a part-task simulator.[20] Participants were not permitted to practise the CLI skill during the tutorial.

Following the introductory video and tutorial, participants ($n = 39$) were randomised into one of three observational practice groups: a flawless demonstration (F, $n = 15$) group that viewed and assessed, by way of a validated checklist and GRS which were modified to reflect institutional practice,[21,22] a single video of an expert performing CLI without any errors; an errorful demonstration (E, $n = 11$) group that viewed and assessed videos that contained flawless and errorful performances, and a feedback-augmented (FA, $n = 13$) group that viewed the same videos as the E group, but also received a set of reference scales that provided information concerning the correctness of their assessments. A simple method of randomisation was used in an attempt to allocate participants in a 1 : 1 : 1 ratio to each group. The numbers of participants in each of the three groups ($n = 15$, $n = 11$, $n = 13$) were deemed to be sufficiently well distributed to remove the need for any block restrictions during randomisation. Participants in groups E and FA were notified that their observational sets would include videos that may contain performance errors. The video demonstrations were produced such that two errorful videos contained one procedural error as determined by checklist items 17 and 20, and the third video contained two procedural errors determined by checklist items 23 and 24 (Table 1). The first two videos in each cohort also contained one global technique error, whereas the third contained two, defined as consensus expert scores of 1 on the *Respect for tissue*, *Time and motion*, *Instrument handling*, and *Flow of procedure* global rating items, respectively (Table 2). In total, there were three errorful and one flawless video demonstration, the presentation of which was distributed and counterbalanced across observational learning groups.

The participants watched and assessed one of their group-specific observational videos each day within a web-based learning environment over the course of 4 days (see Grierson *et al.*[4] for a review of the Internet-mediated observational network). Access to each video was permitted for a single day only and participants were required to complete the assigned checklist and GRS immediately upon viewing the video demonstration. Participants were limited to a single observation of each assigned video. All participants completed the observational assignments in this manner. The day following the completion of the observational learning period (~ 24 hours later), participants returned to the simulation laboratory and performed a supervised and video-recorded CLI on a manikin (Blue Phantom Gen II Ultrasound Central Line Training Model; CAE Healthcare, Kirkland, WA, USA). If the participant required more than four attempts or 3 minutes to cannulate the simulated internal jugular vein with the needle (checklist item 14; Table 1), the supervising experimenter interrupted the performance and advanced the participant to the next stage of the skill. This occurred in 12 instances (F group, four instances; E group, five instances; FA group, three instances).

The video-recorded performances were assessed by two experts using the checklist and GRS. These experts were unfamiliar with the study participants and were blinded to the observational learning conditions under which each participant engaged in the experiment. The expert raters were advised that if experimenter intervention occurred, they were to mark the checklist item as not completed and to consider the quality of the participant's attempts when filling in the GRS.

Analysis

The mean expert ratings attained from the checklist and GRS served as the dependent measures for separate one-way, three-group (F, E, FA) analyses of variance (ANOVAs). The final checklist score for each expert rater involved the sum of the checklist items

Table 1 Checklist of procedures for an ultrasound-guided internal jugular central line insertion

	Central line insertion checklist
1	Using ultrasound, the internal jugular vein is identified in a non-sterile fashion
2	The patient is placed in a slight Trendelenburg position
3	The demonstrator dons a cap, mask, sterile gown and sterile gloves
4	The appropriate area on the patient is cleaned with chlorhexidine
5	The appropriate area is draped in sterile fashion using a full body drape
6	All ports on the central line catheter are flushed with sterile saline
7	After flushing, each port on the central line is clamped or closed
8	The brown port is opened from the end of the catheter to accommodate the guidewire
9	The ultrasound probe is properly set up with a sterile sheath and sonographic gel
10	The internal jugular vein is localised using anatomical landmarks with the ultrasound probe
11	The skin is anaesthetised with 1% lidocaine in a small wheel
12	The deeper structures are also anaesthetised
13	Using ultrasound, the internal jugular vein is localised with the needle
14	Using ultrasound, a large needle is used to cannulate the vein while aspirating for blood
15	Once flashback of venous blood is seen in the syringe, the body of the syringe is detached from the needle tip; the needle remains in the patient
16	The guidewire is advanced through the needle tip into the vein no more than 12–15 cm
17	Once the guidewire is in, the needle is removed; the patient's skin is then nicked with a scalpel to accommodate entry of the dilator
18	The dilator is advanced over the guidewire to dilate the vein
19	The central line catheter is advanced over the guidewire
20	The demonstrator never lets go of the guidewire through any of the previous steps
21	Once the catheter is inserted, the guidewire is entirely removed
22	The catheter is advanced to approximately 14–16 cm on the right side of the patient
23	Blood flow is aspirated from all catheter ports; each port is then flushed and closed
24	A dressing is placed over the catheter
25	A sterile technique is maintained throughout the procedure

presented in Table 1. The performance of the correct item in the correct order was ascribed a score of 1, and the performance of an incorrect item or an item in an incorrect order was ascribed a score of 0. Under this scoring system a top score of 25 is obtainable. The final global rating score reflected the mean score of each scale component, for a rating out of 5.

The participants' checklist and global ratings of the videos observed over the experimental period were also subjected to independent comparisons in order to determine each group's accuracy in evaluating the observed content. Specifically, participants' mean accuracy (out of 25) on the checklist ratings and the mean error in their GRS submissions were subjected to separate one-way, three-group (F, E, FA) ANOVAs. The participants' total rating accuracy on the four checklist items that were manipulated as errorful within the observational videos, and their mean rating accuracy for the GRS items that were manipulated as errorful within the observational videos were analysed in separate two-group (E, FA) ANOVAs.

Effects found to be significant at an alpha value set at $p < 0.05$ were further decomposed using Tukey's honestly significant differences (HSD) *post hoc* methodology.

RESULTS

Analysis of expert checklist rating of participant performance

The analysis of checklist scores revealed no significant performance differences between observational learning groups (d.f.$_{2,36}$ = 1.14, p = 0.33; d = 0.46; grand mean ± standard error [SE]: [20.3 ± 0.7]/

Table 2 Global rating scale for an ultrasound-guided internal jugular central line insertion

	Central line insertion global rating scale				
	1	**2**	**3**	**4**	**5**
Respect for tissue	Frequently used unnecessary force on tissue or caused damage		Careful handling of tissue but occasionally caused inadvertent damage		Consistently handled tissues appropriately with minimal damage
Time and motion	Many unnecessary moves		Efficient time/motion but some unnecessary moves		Clear economy of movement and maximal efficiency
Instrument handling	Repeatedly made tentative or awkward moves with instruments by inappropriate use of instruments		Competent use of instruments but occasionally appeared stiff or awkward		Fluid moves with instruments and no awkwardness
Knowledge of instruments	Frequently asked for wrong instruments or used an inappropriate instrument		Knows names of most instruments and used appropriate instrument		Obviously familiar with the instruments and their names
Flow of procedure	Frequently stopped procedure and seemed unsure of next move		Demonstrated some forward planning with reasonable progression of procedure		Obviously planned course of procedure with effortless flow from one move to the next
Knowledge of procedure	Deficient knowledge		Knows all important steps of operation		Demonstrated familiarity with all aspects of operation
Overall performance	Very poor		Competent		Clearly superior

25). A set of sub-analyses that looked specifically at the performance of the checklist items that were manipulated as errorful within the observational videos revealed no significant differences between the three groups (Item 17: $d.f._{2,36} = 1.31$, $p = 0.28$, $d = 0.42$; Item 20: $d.f._{2,36} = 2.23$, $p = 0.12$, $d = 0.63$; Item 23: $d.f._{2,36} = 1.21$, $p = 0.31$, $d = 0.45$; Item 24: $d.f._{2,36} = 2.51$, $p = 0.10$, $d = 0.75$).

Analysis of expert global ratings of participant performance

Analysis of the GRS measures of participant performance revealed a main effect of group ($d.f._{2,36} = 4.51$, $p = 0.018$; $d = 1.00$). *Post hoc* decomposition of this effect indicated that the FA group performed significantly better than the F and E groups, which did not differ significantly from one another (Fig. 1).

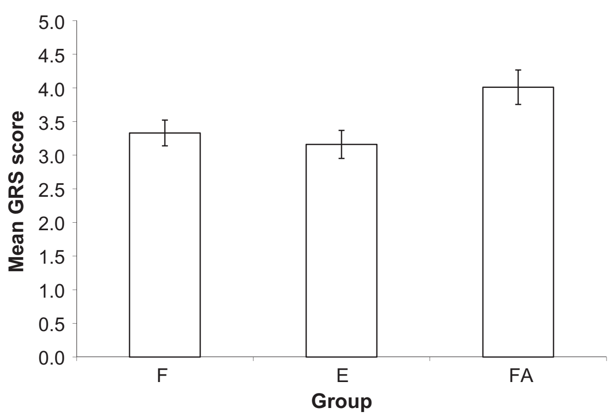

Figure 1 Mean ± standard error global rating scale (GRS) scores plotted as a function of observational learning group. Groups: F = flawless demonstrations; E = flawless and errorful demonstrations; FA = feedback-augmented flawless and errorful demonstrations

Analysis of participant accuracy in completing checklists for the observed videos

Analysis of participant mean accuracy in completing the checklists for the observed videos revealed no significant differences between groups ($d.f._{2,36} = 0.73$, $p = 0.49$, $d = 0.23$; grand mean ± SE: [22.5 ± 0.4]/25). Similarly, analyses of the E and FA group participants' total accuracy on those items that were manipulated to be errorful also revealed no significant differences ($d.f._{1,26} = 3.14$, $p = 0.08$, $d = 0.66$; grand mean ± SE: [2.76 ± 0.2]/4).

Analysis of participant accuracy in completing the GRS for the observed videos

Analysis of participant error in rating the performances observed using the associated GRSs revealed a significant group effect ($d.f._{2,36} = 6.90$, $p = 0.003$, $d = 1.23$), whereby the E group performed the assessment with less error (mean ± SE: [0.63 ± 0.18]/5) than the F group (mean ± SE: [1.22 ± 0.11]/5) and the FA group (mean ± SE: [1.25 ± 0.08]/5), which did not significantly differ from one another.

DISCUSSION

This study demonstrates that observational learning in a context that includes errors can advance the development of clinical skills in novices. The results presented here are consistent with recent evidence which supports the use of mixed observational content for both experimentally constrained spatio-temporal motor tasks[14,15] and clinical procedural skills.[4] However, our findings indicate that observational learning that includes error demonstrations has this benefit on learning only if the observer is given information that confirms that what he or she has observed is errorful. This is revealed through the comparison of the groups' relative abilities to assess the demonstrations they viewed with accuracy and their post-observation performances of the CLI skill. In particular, it is noteworthy that the E group provided the best global ratings of the observed demonstrations, but did not perform as well as the FA group at the skill. This is an indicator that the mere presence of a performance flaw in an observational demonstration, whether it is recognised by the learner or not, is insufficient to promote this type of learning, at least in novice learners. Rather, it appears that variations of performance are better integrated into internal memorial representations of action for this type of skill when the learner has confirmatory information about the specific qualities that define that performance as flawed.

Interestingly, the error observation effects were specifically localised to the aspects of performance that are captured by way of the GRS. This finding is directly in line with the underlying theory that observation of errorful performances provides learners with information regarding the scope of variability that surrounds their movements. It appears that exposure to this variability helps learners understand the potential for and consequences of certain errors in any particular performance context, specifically the types of error that result from neuromuscular inefficiency (i.e. target overshooting, instrument mishandling, uncoordinated bimanual movements, etc.). That the checklist ratings did not reveal any performance differences between observational groups suggests that the presence or absence of errors did not impact on learners' ability to acquire the procedural, or order of events, knowledge associated with the CLI skill. Indeed, our participants' performances, regardless of group assignment, were completed with an average checklist score of 81.2%, which is considered an adequate representation of procedural competence for this skill.[23] It would seem that novices can learn the procedural steps of this complex skill with minimal observational training, and that the inclusion of errors does not improve this learning. It is likely that the participants' iterative experience with the checklist drives this learning as much as, if not more than, the specific content of the observed videos.

Given the present results and the fact that modern health professional simulation spaces are almost uniformly outfitted with video-recording equipment, observational learning appears to be an effective and readily available method of refining simulation-based education. Nonetheless, we recognise that more research is needed to determine the optimal organisation and utilisation of such methods. In particular, it is important that we extend the line of inquiry described in this study to include considerations for the possible positive beneficial effects that watching multiple videos has on observer skill acquisition. Variation, whether error-laden or not, is a well-known feature of effective learning paradigms,[8,11,15] and it may be that the improved acquisition demonstrated by the FA group relative to the F group derives from the fact that participants in the former group observed a number of different demonstrations during the intervention period. However, the differential findings elicited between the FA and E groups suggests that through the use of video technology, an observational practice system – which may include the intentional use of errors – may liberate a portion of motor skill learning from the clinical or simulated learning environment in a way that can reach a large number of learners. Certainly, a hybrid programme of systematic physical and observational practice represents an evidence-aligned model for the development of clinical skill competence.

Contributors: KD designed and performed the experiments, analysed the data and wrote the manuscript. AW assisted in participant recruitment, data collection and manuscript generation. LO supervised data collection. LEMG planned and supervised all aspects of the project. All authors contributed to the critical revision of the paper, approved the final manuscript for publication, and have agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.
Acknowledgements: the authors acknowledge Dr David Musson and the staff and faculty at the Centre for Simulation-based Learning at McMaster University for their valuable assistance and guidance throughout the data collection process, and Dr Bill Kapralos and his research team at the Ontario University Institute of Technology for their help in operating the observational network.
Funding: none.
Conflicts of interest: none.
Ethical approval: this study was approved by the Hamilton Health Sciences and McMaster University Faculty of Health Sciences Integrated Research Ethics Board.

REFERENCES

1 Reznick R, MacRae H. Teaching surgical skills – changes in the wind. *N Engl J Med* 2006;**355** (25):2664–9.
2 Waurick R, Weber T, Bröking K, van Aken H. The European Working Time Directive: effect on education and clinical care. *Curr Opin Anaesthesiol* 2007;**20** (6):576–9.
3 Kurrek MM, Devitt JH. The cost for construction and operation of a simulation centre. *Can J Anaesth* 1997;**44**:1191–5.
4 Grierson LEM, Barry M, Kapralos B, Carnahan H, Dubrowski A. The role of collaborative interactivity in the observational practice of clinical skills. *Med Educ* 2012;**46**:409–16.
5 Martineau B, Mamede S, St-Onge C, Rikers RMJP, Schmidt HG. To observe or not to observe peers when learning physical examination skills; that is the question. *BMC Med Educ* 2013;**13**:55.
6 Ashford D, Bennett SJ, Davids K. Observational modelling effects for movement dynamics and movement outcome measures across differing task constraints: a meta-analysis. *J Motor Behav* 2006;**38**:185–205.
7 Mattar AG, Gribble PL. Motor learning by observing. *Neuron* 2005;**46** (1):153–60.
8 Schmidt RA. A schema theory of discrete motor skill learning. *Psychol Rev* 1975;**82**:225.
9 Maslovat D, Hodges NJ, Krigolson OE, Handy TC. Observational practice benefits are limited to perceptual improvements in the acquisition of a novel coordination skill. *Exp Brain Res* 2010;**204**: 119–30.
10 Meyer DE, Abrams RA, Kornblum S, Wright CE, Smith JE. Optimality in human motor performance: ideal control of rapid aimed movements. *Psychol Rev* 1988;**95** (3):340–70.
11 Schmidt RA, Zelaznik H, Hawkins B, Frank JS, Quinn JT. Motor-output variability: a theory for the accuracy of rapid motor acts. *Psychol Rev* 1979;**86**:415–51.
12 Elliott D, Hansen S, Grierson LEM, Lyons J, Bennett SJ, Hayes SJ. Goal-directed aiming: two components but multiple processes. *Psychol Bull* 2010;**136**:1023–44.
13 Elliott D, Grierson LEM, Hayes SJ, Lyons J. Action representations in perception, motor control and learning: implications for medical education. *Med Educ* 2011;**45**:119–31.
14 Andrieux M, Proteau L. Observation learning of a motor task: who and when? *Exp Brain Res* 2013;**229** (1):125–37.
15 Rohbanfard H, Proteau L. Learning through observation: a combination of expert and novice models favours learning. *Exp Brain Res* 2011;**215** (3–4):183–97.
16 Graham AS, Ozment C, Tegtmeyer K, Lai S, Braner DA. Videos in clinical medicine. Central venous catheterisation. *N Engl J Med* 2007;**356** (21):e21.
17 Moureau N, Lamperti M, Kelly LJ, Dawson R, Elbarbary M, van Boxtel AJH, Pittiruti M. Evidence-based consensus on the insertion of central venous access devices: definition of minimal requirements for training. *Br J Anaesth* 2013;**110** (3):347–56.
18 Dong Y, Suri HS, Cook DA, Kashani KB, Mullon JJ, Enders FT, Rubin O, Ziv A, Dunn WF. Simulation-based objective assessment discerns clinical proficiency in central line placement: a construct validation. *Chest* 2010;**137** (5):1050–6.
19 Ma IW, Brindle ME, Ronksley PE, Lorenzetti DL, Sauve RS, Ghali WA. Use of simulation-based education to improve outcomes of central venous catheterisation: a systematic review and meta-analysis. *Acad Med* 2011;**86** (9):1137–47.
20 Pollard BA. New model for learning ultrasound-guided needle to target localisation. *Reg Anesth Pain Med* 2008;**33** (4):360–2.
21 Bould MD, Crabtree NA, Naik VN. Assessment of procedural skills in anaesthesia. *Br J Anaesth* 2009;**103** (4):472–83.
22 Ma IW, Zalunardo N, Pachev G, Beran T, Brown M, Hatala R, McLaughlin K. Comparing the use of global rating scale with checklists for the assessment of central venous catheterisation skills using simulation. *Adv Health Sci Educ Theory Pract* 2012;**17** (4):457–70.
23 Barsuk JH, McGaghie WC, Cohen ER, Balachandran JS, Wayne DB. Use of simulation-based mastery learning to improve the quality of central venous catheter placement in a medical intensive care unit. *J Hosp Med* 2009;**4** (7):397–403.

Received 7 April 2014; editorial comments to author 12 June 2014, 8 September 2014; accepted for publication 17 September 2014

