Clinical Simulation in Nursing (2019) 36, 37-46

Clinical Simulation in Nursing

www.elsevier.com/locate/ecsn

Featured Article

Quality of Virtual Patients on WeChat for Nursing History-Taking Training

Jie Luo, RN, MScN[a,b,†], Jing Zheng, PhD, Lecturer[c,†], Huan He, RN, BScN[d], Ting Liu, BScN, Master Candidate[b], Juanjuan Zhao, MScN, Lecturer[b], Kun Li, PhD, Associate Professor[b,*]

[a]*Department of Endocrinology, The First Affiliated Hospital of Zhejiang University, Hangzhou, China*
[b]*School of Nursing, Sun Yat-sen University, Guangzhou, China*
[c]*School of Nursing, Guangdong Pharmaceutical University, Guangzhou, China*
[d]*Cardiovascular Medicine Department, The Fifth Affiliated Hospital of Guangzhou Medical University, Guangzhou, China*

KEYWORDS
medical history taking;
nursing education;
virtual patient;
simulation;
WeChat

Abstract

Background: Virtual patients (VPs) have long been used in health care education. However, their high cost restricts their widespread use, and personal use of VPs has not yet been achieved. In addition, the quality of VPs, such as validity and accuracy, has rarely been reported, although the effects of VPs have been widely discussed. "Sun Yat-sen University Nursing Education Platform" (SYSU-NEP) is a software program designed for use on WeChat that provides VPs for nursing history-taking sessions. SYSU-NEP is characterized by ease of use, accessibility, and capacity to facilitate autonomous learning. This study aimed to examine the quality of the 13 VPs in SYSU-NEP based on content completeness and accuracy.

Method: This was a retrospective data analysis study. The data were derived from 570 history-taking sessions involving conversations generated by 57 baccalaureate nursing students and 13 VPs. Among the conversations, the students' queries without a response and the queries with an incorrect response were analyzed. The content completeness of each VP was defined as the percentage of queries with a response. The accuracy was defined as the percentage of queries with a correct response.

Results: There were 13,312 student queries. The content completeness of each VP ranged from 73.7% to 97.3% (median: 83%; interquartile range: 78.1%—88.3%). The content accuracy of each VP ranged from 85.7% to 98.8% (median: 97.7%; interquartile range: 95.7%—98%). The Present Conditions part had the highest rate of queries without a response and with an incorrect response.

Conclusions: The content completeness and accuracy of the VPs were satisfactory but had room for improvement. The queries without a response and with an incorrect response were also identified, which is helpful for the continuous improvement of the VPs.

Cite this article:
Luo, J., Zheng, J., He, H., Liu, T., Zhao, J., & Li, K. (2019, November). Quality of virtual patients on wechat for nursing history-taking training. *Clinical Simulation in Nursing, 36*(C), 37-46. https://doi.org/10.1016/j.ecns.2019.07.001.

[†]Contributed equally to this work.

Funding: The study was supported by the China Medical Board [grant 15-242], National Natural Science Foundation of China [grant 71974213], and Sun Yat-sen University [grants kcjs201814 and 53000-16300008].

Declarations of interest: none.

* Corresponding author: likun22@mail.sysu.edu.cn (K. Li).

1876-1399/$ - see front matter
https://doi.org/10.1016/j.ecns.2019.07.001

Virtual patients (VPs) are defined by the American Association for Medical Colleges as a "specific type of computer program that simulates real-life clinical scenarios," which can be used by learners to "emulate the roles of health care providers to obtain a history, conduct a physical examination, and make diagnostic and therapeutic decisions" (Association of American Medical Colleges, 2007). Since they were first described in 1971, VPs have been used in health care education for nearly half a century (Harless, Drennon, Marxer, Root, & Miller, 1971). The formats of VPs have improved a great deal. VPs are often delivered as programs that involve virtual worlds, text-based VPs, high-fidelity simulations, or conversational agents (Hege, Kononowicz, Tolks, Edelbring, & Kuehlmeyer, 2016). VPs are mainly used for clinical reasoning training, history-taking practice, examination skills, and so on (Kononowicz, Zary, Edelbring, Corral, & Hege, 2015).

Key Points

- The quality of virtual patients has rarely been reported, although their effects have been widely discussed.
- An ideal conversational virtual patient should have guaranteed content completeness and accuracy.
- The content completeness and accuracy of the 13 virtual patients on WeChat developed in this study were shown to be satisfactory but had room for improvement.

Compared with traditional clinical teaching, the advantages of VPs mainly involve avoidance of risk, repeatability, and objectivity (Zary, Johnson, Boberg, & Fors, 2006). However, the high costs often limit the widespread use of VPs. Real clinical learning is difficult because it is hard to provide the same experience for different learners because of lack of reproducibility of clinical encounters. VPs make repeated training possible and eliminate any subjectivity from the patients, which facilitates standardized practice and examination for learners (Cendan & Lok, 2012; Consorti, Mancuso, Nocioni, & Piccolo, 2012). VPs could also provide a safe and nonthreatening environment for students' history-taking training, which presents no risk of adverse outcomes (Sweigart, Burden, Carlton, & Fillwalk, 2014). In addition, learners can encounter a wide variety of VP cases (Baumann-Birkbeck, 2017).

Despite the advantages, the costs of VPs are often concerning. It has been reported that developing a VP scenario costs $10,000 to 50,000 on average (Cendan & Lok, 2012). More recently, Isaza-Restrepo, Gomez, Cifuentes, and Arguello (2018) also reported that the cost of the VP software and faculty time amounted to $49,000. In addition, the complicated training and required VP program maintenance need further funding. Therefore, at present, VPs are mainly used in education programs in medical colleges. Personal use of VPs has not yet been achieved.

Another fact worth noting is that the quality of VPs, such as validity and accuracy, has rarely been reported, although their effects have been widely discussed. Studies have shown that VPs can increase learners" self-efficacy, engagement, and performance (Harder, 2010, Barnett, Gallimore, Pitterle, & Morrill, 2016; Stultz, Forder, & Pakyz, 2017). However, there are limited studies on the quality of VPs. Guise, Chambers, Conradi, Kavia, and Valimaki (2012) reported that the content validity of VPs was tested by clinical experts and peer reviewed, but no detailed results were provided in their paper. Urresti-Gundlach et al. (2017) compared characteristics of VPs and population data from the German health care system, and they found that some patient information, such as unemployment, disability, and migration background, was almost nonexistent in the VPs. Maicher et al. (2017) tested several conversational VPs designed for history-taking training, and they found that the accuracy of the VP responses ranged from 79% to 86%. Based on these studies, it was found that there are no uniform methods and criteria for VP quality control assessment. Nevertheless, the professionalism, completeness, and accuracy of VP content should be guaranteed.

Sun Yat-sen University Nursing Education Platform (SYSU-NEP) is a software program that randomly provides different VPs to interact with subscribers for history-taking practice. It is in Chinese and the copyright is held by Sun Yat-sen university (Sun Yat-sen University, 2016). SYSU-NEP was developed on WeChat (Tencent, 2011), the most popular social media website in China. Similar to Facebook, WeChat can be logged onto for free using a personal computer, cell phone, or tablet that supports internet communication (Zeng, Deng, Wang, & Liu, 2016). Another innovation of WeChat is its varied functions, which are changing the lives of Chinese individuals. SYSU-NEP, which can be accessed via a public account on WeChat, has several advantages including its ease of use, accessibility, and capacity for facilitating autonomous learning. Unlike previous complicated and expensive VPs, SYSU-NEP was originally designed for personal use, so it is easily accessible and free of charge.

Earlier application showed that SYSU-NEP was helpful for identifying students' weaknesses regarding history taking, and it could improve their self-directed learning behavior (Liu et al., 2018). However, problems were also exposed. Some student queries were not matched to a correct VP response. This may be related to a design deficiency in the accuracy of the VP chat scripts. On the other hand, some student queries had no VP response matched. This may reflect problems with the content completeness of the VPs. Therefore, it is necessary to examine the accuracy and content completeness of the VPs in SYSU-NEP, which would be helpful for future modification of SYSU-NEP.

This study aimed to examine the quality of the VPs in SYSU-NEP regarding content completeness and accuracy.

The results of the study will be helpful to illustrate the reasons behind the emerging problems and provide evidence for improved design of conversational VPs on WeChat. The improvement of the VPs' quality could optimize users' experience and more fully exploit the advantages of using VPs in history-taking training. This study could also provide a reference and evidence for the future development of VPs.

Methods

Design

This study involved a retrospective data analysis of data from a pilot study. Data were derived from conversations between students and VPs that were automatically recorded by SYSU-NEP. The queries of subscribers without a response and the queries with an incorrect response were analyzed. The content completeness of each VP was defined as the percentage of student queries with a response. The accuracy was defined as the percentage of student queries with a correct response.

Pilot Study Procedure

The pilot study was conducted in the autumn semester from September 2016 to January 2017 when the Internal Medical Nursing curriculum was being taught because all the VPs in SYSU-NEP had internal medical diseases. The participants were recruited from the baccalaureate students in the third year of a four-year nursing program at School of Nursing, Sun Yat-sen university (69 in total). The inclusion criteria were as follows: (a) full-time third-grade baccalaureate nursing student; (b) passed the Health Assessment curriculum in the last semester; and (c) owned a mobile terminal and could access WeChat on a wireless network. Any student who had missed more than one week of the Internal Medical Nursing curriculum was to be excluded from the statistical analysis.

At the beginning of the autumn semester, training on the usage of SYSU-NEP was given to the students, which involved a 10-minute introduction. They were asked to select VPs with different organ system diseases according to the progress of the Internal Medical Nursing curriculum. It was emphasized that participation was voluntary, they could drop out at any time, and their performance on the history-taking software program would not influence their final grades for the Internal Medical Nursing curriculum because the use of SYSU-NEP was part of a research study rather than a formal component of their education program.

At the end of the autumn semester, 57 of 69 (83%) students participated in history-taking practice on WeChat. None of the 57 students was excluded from the final data analysis. They took part in 570 history-taking sessions.

Development of SYSU-NEP

The development of SYSU-NEP included two parts. One was the design of the VP chat scripts. The other involved the establishment of the software on a WeChat public account.

The original chat scripts were designed by clinical nursing specialists and reviewed by the study authors. The nursing specialists were required to select common clinical cases seen in internal medical departments, such as patients with chronic obstructive pulmonary disease, coronary heart disease, and peptic ulcer. Before drafting the chat scripts, essential preparation involved listening to and recording the bedside history-taking processes of different nurses and students with the same patient. The nursing specialists were then asked to write chat scripts according to a preset template. In the template, the nursing specialists needed to list as many queries as possible for each relevant VP response.

Functional Health Pattern (FHP), an acknowledged nursing model for nursing assessment, was the framework for establishing the VP constructs. It consists of 11 patterns (Gordon, 1994). Besides the 11 patterns of FHP, other essential components of history taking were also included for each VP. Thus, each VP included 17 parts: Beginning of the Conversation, Biographic Information, Present Conditions, Medical History, Family Medical History, Personal History, Health Perception—Health Management, Nutrition Metabolic, Elimination, Activity/Exercise, Sleep/Rest, Cognitive Perception, Roles and Relationships, Sexual—Reproductive, Coping/Stress and Self-Perception/Self-Concept, Values and Beliefs, and the End of the Conversation. The original chat scripts for each VP covered all 17 parts. Table 1 shows the structure and contents related to the VPs.

Adding a logical operation formula was a vital procedure for designing the chat scripts. This was implemented by the study authors. For each VP response, the keywords present in the various relevant subscriber queries were extracted and formed into a formula using the logical Boolean operator symbols "AND" and "OR." For the convenience of operation, "AND" was represented by "#" and "OR" was represented by a space. Thus, the various possible queries were replaced by a simple and effective logical operation formula, which represents all of the many possible query forms.

An engineer was in charge of the development of SYSU-NEP. SYSU-NEP contains three components: WeChat interface, Web operation management system, and subscriber identity authentication design. All the software codes were used on the Alibaba Cloud server (Alibaba Cloud Computing Co. Ltd, 2009). The WeChat interface can support online text, voice, and picture message interactions and simulate history-taking sessions. The Web operation management system was developed using Java language, and it can support the management of VPs and registered users and the statistical analysis of the history

Table 1 Structure and Contents of the Virtual Patients (Liu et al., 2018)

Structure	Main Contents
1. Beginning of a conversation	• Introduce oneself and the purpose of this history taking
2. Biographic information	• Patient's name, age, occupation, address, education level, race, admission form, and so forth
3. Present conditions	• Reason for admission, onset and duration of disease, main symptoms, diagnosis and treatment, other discomforts, and so forth
4. Past medical history	• Previous health status and diseases
5. Family history	• Health status of relatives
6. Personal history	• Patient's life experiences, working conditions, living conditions, and so forth
7. Health perception-health management	• Addictions (tobacco, alcohol, or drugs), allergy history, physical exercise, regular body examination, and so forth • For specific patients: self-management of blood pressure, blood glucose, or blood lipids, and so forth
8. Nutrition metabolic	• Weight changes, appetite, eating habits, and so forth • Wound, pressure injury, or drainage • For specific patients: swallowing disorder, food category, feeding method, and so forth
9. Elimination	• Urination (frequency, nature of urine, accompanying symptoms, incontinence, retention, external appliances, and so forth) • Defecation (frequency, nature of stool, accompanying symptoms, incontinence, constipation, and so forth) • Stoma (type, self-care ability, and so forth)
10. Activity/exercise	• Self-care regarding activities of daily living, activity tolerance, the influence of disease on activity, and so forth
11. Sleep/rest	• Sleep disorder and drugs for sleep
12. Cognitive perception	• Pain, sensory disorders regarding seeing, listening, touching, and so forth • Cognition of patients regarding disease
13. Roles and relationships	• People who live together, family relationship, and so forth
14. Sexual—reproductive	• Marriage and child-birth history, menstrual history
15. Coping/stress and self-perception/self-concept	• Health insurance • Stressors (medical cost, family, disease, working, learning, and so forth) and stress responses
16. Values and beliefs	• Religion and special needs
17. Ending of a conversation	• Express thanks

Table 2 Content Completeness and Accuracy for 13 Virtual Patients

VP*	Total Queries Asked	Queries Without Response Caused by VP Design	Content Completeness	Queries With Incorrect Response	Content Accuracy
VP1	1,059	180	83.0%	22	97.9%
VP2	1,035	28	97.3%	46	95.6%
VP3	951	212	77.7%	28	97.1%
VP4	1,209	265	78.1%	14	98.8%
VP5	1,645	223	86.4%	42	97.4%
VP6	1,089	199	81.7%	22	98.0%
VP7	2,648	333	87.4%	61	97.7%
VP8	1,006	217	78.4%	23	97.7%
VP9	554	65	88.3%	24	95.7%
VP10	562	148	73.7%	7	98.8%
VP11	490	113	76.9%	8	98.4%
VP12	287	19	93.4%	41	85.7%
VP13	777	22	97.2%	47	94.0%
Median interquartile range			83% 78.1%—88.3%		97.7% 95.7%—98%

Note. VP = virtual patients.

* VP1—VP4 had respiratory system diseases, VP5 had a urinary system disease, VP6 had an endocrine system disease, VP7 had a digestive system disease, VP8 had a hematological system disease, and VP9—VP12 had cardiovascular system diseases.

taking data. The WeChat subscriber identity authentication system ensures that access to the program is only available to authorized WeChat users.

Usage of SYSU-NEP

First, the subscribers need to log into the SYSU-NEP public account on WeChat and the software can then be used. There are multiple organ systems listed in the menu displayed at the bottom of the SYSU-NEP interface. Once an organ system is selected, a corresponding VP with a cartoon image is randomly assigned to the subscriber. The VP gender and some necessary usage instructions are also briefly provided in text form on the interface. The history-taking session duration is limited to 30 minutes. The subscriber can communicate with the VP through verbal or typed input. The software can recognize a subscriber's verbal or typed queries and display the VP's responses in text form. The conversation between the subscriber and VP can be stopped by verbal or typed input, stating ''Ending history-taking.'' The system also automatically gives a score according to the completeness of the history-taking session. The maximum score is 100.

Ethical Considerations

The study was approved by the Department of Education Administration and ethics committee of Sun Yat-sen university. All the students were informed that their participation in the study was on a voluntary basis. They could refuse to participate or drop out at any time. Their decisions not to participate in the study and their performance on SYSU-NEP would not influence their final performance in the Internal Medical Nursing Curriculum program.

Data Analysis

SPSS16.0 software (SPSS Inc., Chicago, IL) was used for statistical analysis. Frequencies and percentages were calculated to describe the total queries, queries without a response, and queries with an incorrect response for each VP. The content completeness of each VP was defined as the percentage of queries with a response. The accuracy was defined as the percentage of queries with a correct response.

After preliminary analysis, it was found that quite a number of the queries without a response were caused not by the VP design but by the students. Based on this, the queries without a response that were caused by poor VP design were identified and the content completeness and accuracy calculations were recalculated by two study authors together.

Because each VP consisted of 17 parts, for each VP, the constituent ratios (%) of the queries without a response and the queries with an incorrect response for each of the 17 parts were also computed. This was helpful for identifying the weaker parts in the chat scripts.

Table 3 Constituent Ratios of the Queries Without a Response for Each VP

	Number of Queries Without a Response and Constituent Ratio (%)																		
VP	Beginning	BI	PC	PMH	PH	FH	HP-HM	NM	Eli	A/E	S/R	CP	RR	S–R	C/S, SP/SC	VB	Ending		
VP1	14 (7.8%)	10 (5.6%)	96 (53.3%)	5 (2.8%)	0	0	7 (3.9%)	21 (11.7%)	0	5 (2.8%)	0	1 (0.6%)	1 (0.6%)	0	12 (6.7%)	0	8 (4.4%)		
VP2	7 (25.0%)	3 (10.7%)	15 (53.6%)	1 (3.6%)	0	0	0	1 (3.6%)	0	0	0	0	1 (3.6%)	0	0	0	0		
VP3	15 (7.1%)	15 (7.1%)	76 (35.8%)	16 (7.5%)	2 (0.9%)	5 (2.4%)	15 (7.1%)	27 (12.7%)	4 (1.9%)	14 (6.6%)	2 (0.9%)	3 (1.4%)	1 (0.5%)	0	3 (1.4%)	0	14 (6.6%)		
VP4	7 (2.6%)	9 (3.4%)	142 (53.6%)	17 (6.4%)	4 (1.5%)	8 (3.0%)	24 (9.1%)	26 (9.8%)	1 (0.4%)	9 (3.4%)	2 (0.8%)	2 (0.8%)	3 (1.1%)	0	5 (1.9%)	0	6 (2.3%)		
VP5	1 (1.5%)	7 (10.8%)	31 (47.7%)	7 (10.8%)	0	3 (4.6%)	0	9 (13.8%)	0	3 (4.6%)	3 (4.6%)	0	1 (1.5%)	0	0	0	0		
VP6	4 (2.7%)	11 (7.4%)	109 (73.6%)	4 (2.7%)	1 (0.7%)	5 (3.4%)	0	2 (1.4%)	0	5 (3.4%)	1 (0.7%)	1 (0.7%)	0	0	3 (2.0%)	0	2 (1.4%)		
VP7	2 (1.8%)	8 (7.1%)	51 (45.1%)	7 (6.2%)	0	0	9 (8.0%)	25 (22.1%)	0	2 (1.8%)	3 (2.7%)	0	1 (0.9%)	0	2 (1.8%)	0	3 (2.7%)		
VP8	4 (21.1%)	4 (21.1%)	9 (47.4%)	0	0	0	0	0	0	0	0	0	1 (5.3%)	1 (5.3%)	0	0	0		
VP9	5 (22.7%)	7 (31.8%)	9 (40.9%)	0	0	0	0	1 (4.5%)	0	0	0	0	0	0	0	0	0		
VP10	9 (4.0%)	6 (2.7%)	127 (57.0%)	29 (13.0%)	0	4 (1.8%)	3 (1.3%)	14 (6.3%)	0	20 (9.0%)	2 (0.9%)	0	3 (1.3%)	2 (0.9%)	1 (0.4%)	0	3 (1.3%)		
VP11	6 (3.0%)	15 (7.5%)	121 (60.8%)	12 (6.0%)	0	4 (2.0%)	0	16 (8.0%)	16 (8.0%)	4 (2.0%)	0	1 (0.5%)	1 (0.5%)	0	0	0	3 (1.5%)		
VP12	7 (3.2%)	7 (3.2%)	122 (56.2%)	36 (16.6%)	10 (4.6%)	3 (1.4%)	2 (0.9%)	13 (6.0%)	0	4 (1.8%)	3 (1.4%)	0	2 (0.9%)	1 (0.5%)	6 (2.8%)	0	1 (0.5%)		
VP13	24 (7.2%)	22 (6.6%)	142 (42.6%)	9 (2.7%)	0	15 (4.5%)	2 (0.6%)	72 (21.6%)	9 (2.7%)	19 (5.7%)	0	4 (1.2%)	1 (0.3%)	0	11 (3.3%)	0	3 (0.9%)		

Note. BI = biographic information; PC = present conditions; PMH = past medical history; PH = personal history; FH = family history; HP-HM = Health Perception–Health Management; NM = Nutrition Metabolic; Eli = Elimination; A/E = Activity/Exercise; S/R = Sleep/Rest; CP = Cognitive Perception; RR = Roles and Relationships; S-R = Sexual–Reproductive; C/S, SP/SC = Coping/Stress and Self-Perception/Self-Concept; VB = Values and Beliefs; VP = virtual patient.

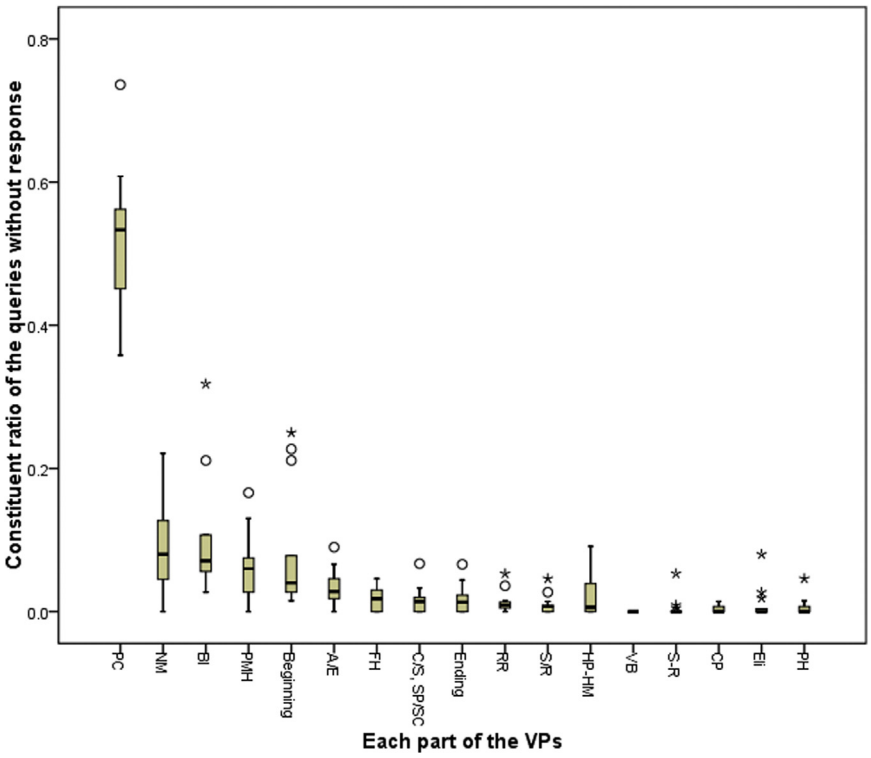

Figure 1 Queries without a response for each of the 17 VP parts. *Note.* VP = virtual patient; PC = Present Conditions; NM = Nutrition Metabolic; BI = Biographic Information; PMH = Past Medical History; A/E = Activity/Exercise; FH = Family History; C/S, SP/SC = Coping/Stress and Self-Perception/Self-Concept; RR = Roles and Relationships; S/R = Sleep/Rest; HP-HM = Health Perception−Health Management; VB = Values and Beliefs; S-R = Sexual−Reproductive; CP = Cognitive Perception; Eli = Elimination; PH = Personal History. The medians (among the 13 VPs) of the constituent ratio of the queries without a response are plotted for each VP part, from the highest median value to the lowest. The horizontal line in the boxes indicates the median, and the lower and upper bounds of the boxes indicate the interquartile range. The circles and asterisks indicate the extreme values (to different extents) of the data.

Results

Content Completeness and Accuracy of the 13 VPs

The content completeness and accuracy of the 13 VPs are shown in Table 2. There were a total of 13,312 queries. The number of queries generated for each VP ranged from 287 to 2,648 (mean: 1,024; SD: 605). The number of queries without a response caused by VP design ranged from 19 to 333 (mean: 156; SD: 100). The content completeness of the VPs ranged from 73.7% to 97.3% (median: 83%; interquartile range: 78.1%−88.3%), indicating that the VPs could answer 73.7%−97.3% of the queries raised by the students.

The numbers of queries with an incorrect response for each VP ranged from 7 to 61 (mean: 30; SD: 26). The content accuracy of the VP cases ranged from 85.7% to 98.8% (median: 97.7%; interquartile range: 95.7%−98%), indicating that the VPs could correctly answer 85.7%−98.8% of the queries raised by the students.

Constituent Ratios of Queries Without a Response for Each VP

To identify the weak points in the design of the VPs, the constituent ratios (%) of the queries without a response for each VP were computed (Table 3 and Figure 1). The part with the most queries without a response was Present Conditions (median: 53.3%; interquartile range: 45.1%−56.2%), indicating that a median of 53.3% (among the 13 VPs) of queries without a response were associated with Present Conditions. The other two parts with the highest constituent ratios of queries without a response were Nutrition Metabolic (median: 8%; interquartile range: 4.5%−12.7%) and Biographic Information (median: 7.1%; interquartile range: 5.6%−10.7%).

Constituent Ratios of Queries With an Incorrect Response for Each VP

The queries with an incorrect response were all caused by incorrect VP design. The constituent ratios (%) of the

Table 4 Constituent Ratio of Queries With an Incorrect Response for Each VP

VP	Number of Queries With an Incorrect Response and Constituent Ratio (%)																
	Beginning	BI	PC	PMH	PH	FH	HP-HM	NM	Eli	A/E	S/R	CP	RR	S—R	C/S, SP/SC	VB	Ending
VP1	0	3 (13.6%)	14 (63.6%)	0	0	1 (4.5%)	0	0	0	0	0	1 (4.5%)	3 (13.6%)	0	0	0	0
VP2	7 (15.2%)	6 (13.0%)	24 (52.2%)	1 (2.2%)	1 (2.2%)	4 (8.7%)	0	1 (2.2%)	0	1 (2.2%)	0	0	1 (2.2%)	0	0	0	0
VP3	0	5 (17.9%)	14 (50.0%)	0	0	8 (28.6%)	0	1 (3.6%)	0	0	0	0	0	0	0	0	0
VP4	0	2 (14.3%)	6 (42.9%)	0	0	5 (35.7%)	0	0	0	0	0	0	0	0	1 (7.1%)	0	0
VP5	0	2 (8.3%)	7 (29.2%)	1 (4.2%)	0	2 (8.3%)	9 (37.5%)	3 (12.5%)	0	0	0	0	0	0	0	0	0
VP6	0	1 (14.3%)	5 (71.4%)	0	0	1 (14.3%)	0	0	0	0	0	0	0	0	0	0	0
VP7	0	2 (25.0%)	3 (37.5%)	0	0	1 (12.5%)	0	0	0	0	0	0	1 (12.5%)	0	0	0	1 (12.5%)
VP8	4 (9.8%)	4 (9.8%)	17 (41.5%)	4 (9.8%)	0	4 (9.8%)	4 (9.8%)	0	1 (2.4%)	0	0	0	1 (2.4%)	1 (2.4%)	1 (2.4%)	0	0
VP9	22 (46.8%)	5 (10.6%)	8 (17.0%)	10 (21.3%)	0	1 (2.1%)	0	0	1 (2.1%)	0	0	0	0	0	0	0	0
VP10	0	1 (2.4%)	8 (19.0%)	29 (69.0%)	0	2 (4.8%)	1 (2.4%)	0	0	1 (2.4%)	0	0	0	0	0	0	0
VP11	0	0	13 (59.1%)	0	0	3 (13.6%)	6 (27.3%)	0	0	0	0	0	0	0	0	0	0
VP12	0	2 (8.7%)	13 (56.5%)	3 (13.0%)	1 (4.3%)	4 (17.4%)	0	0	0	0	0	0	0	0	0	0	0
VP13	0	2 (3.3%)	32 (52.5%)	0	1 (1.6%)	7 (11.5%)	0	7 (11.5%)	0	3 (4.9%)	3 (4.9%)	0	4 (6.6%)	0	2 (3.3%)	0	0

Note. BI = biographic information; PC = present conditions; PMH = past medical history; PH = personal history; FH = family history; HP-HM = Health Perception—Health Management; NM = Nutrition Metabolic; Eli = Elimination; A/E = Activity/Exercise; S/R = Sleep/Rest; CP = Cognitive Perception; RR = Roles and Relationships; S-R = Sexual—Reproductive; C/S, SP/SC = Coping/Stress and Self-Perception/Self-Concept; VB = Values and Beliefs; VP = virtual patient.

queries with an incorrect response for each VP were computed (Table 4 and Figure 2). The part with the most queries with an incorrect response was Present Conditions (median: 50%; interquartile range: 37.5%—56.5%), indicating that about 50% of queries with an incorrect response were associated with Present Conditions. The other two parts with the highest constituent ratios of queries with an incorrect response were Family History (median: 11.5%; interquartile range: 8.3%—14.3%) and Biographic Information (median: 10.6%; interquartile range: 8.3%—14.3%).

Discussion

The main purpose of this study was to examine the quality of the SYSU-NEP VPs in terms of content completeness and accuracy. The results showed that the content completeness and accuracy of the VPs were satisfactory but had room for improvement. The results regarding the constituent ratios for both the queries without a response and the queries with an incorrect response revealed that the Present Conditions part needed to be modified the most.

The ideal VP is supposed be able to answer any queries raised by nursing students. The content completeness reflected the actual content validity of the VP. First of all, high content completeness requires that the VP chat script design conforms to nurses' clinical thinking. To achieve this goal, the FHP, originally developed for nursing assessment, was used as the framework for establishing the VP constructs. Second, the contents of the VPs should reflect the actual clinical practice (Bateman, Allen, Samani, Kidd, & Davies, 2013). All the cases were designed by nursing specialists with rich clinical nursing and teaching experience according to specific design guidelines, including case selection, data preparation, and chat script drafting. The researchers were responsible for reviewing the chat scripts and adding logical operation formulate. The use of Boolean logical operators ensured that one response could be designed to correspond to queries in multiple different forms.

In the study, the median of the content completeness for the VPs was 83%, indicating that about 17% of queries had no response. There were several main reasons that accounted for the queries without a response. First, the logical operation formula failed to cover all the query forms. In Chinese, there can be different ways of expressing a single question. It was impossible for the VP developers to come up with all the possibilities for a single question in the first design. Therefore, the VPs need to be tested and modified continuously. The second reason is that the original chat scripts had no corresponding response, which was a design deficiency. For example, some students asked for the weight and height of the VP to assess nutrition status, but it was found that some VPs had no response regarding height. The third reason is that some students posed a query asking for information that was not related to

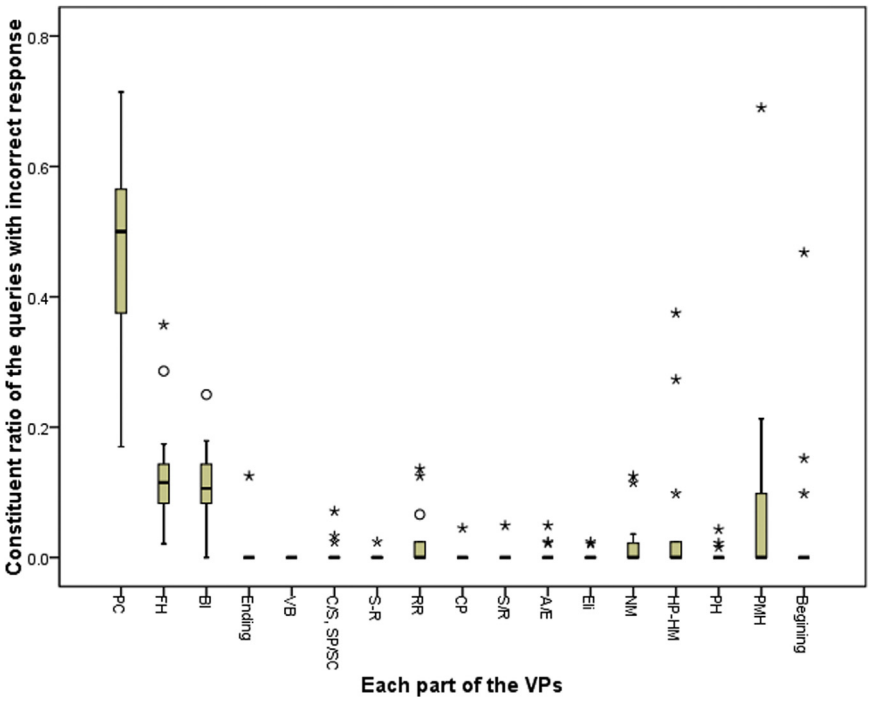

Figure 2 Queries with an incorrect response for each of the 17 VP parts. VP = Virtual Patient; PC = Present Conditions; FH = Family History; BI = Biographic Information; VB = Values and Beliefs; C/S, SP/SC = Coping/Stress and Self-Perception/Self-Concept; S-R = Sexual−Reproductive; RR = Roles and Relationships; CP = Cognitive Perception; S/R = Sleep/Rest; A/E = Activity/Exercise; Eli = Elimination; NM = Nutrition Metabolic; HP-HM = Health Perception−Health Management; PH = Personal History; PMH = Past Medical History. The medians (among the 13 VPs) of the constituent ratio of the queries with an incorrect response are plotted for each VP part, from the highest median value to the lowest. The horizontal line in the boxes indicates the median, and the lower and upper bounds of the boxes indicate the interquartile range. The circles and asterisks indicate the extreme values (to different extents) of the data.

the VP case because of being misled by the previous VP response, which led to the inability of the VP to answer. For example, a VP with asthma referred to having a runny nose in previous responses, and consequently, some students carefully asked about the nature of the nasal secretions. The misleading VP response contents had to be deleted as part of the improvement of the program.

In addition, it was found that many student queries without a VP response were caused not by the design of the VPs but by the students. Some student queries were just tests. Some students used obscure dialects. Some student queries lacked necessary subjects, predicates, or objects. In these cases, most students could adjust their inquiry approach and get correct responses. Sometimes, the problem was an incorrect query approach that should be avoided by the students, for example, inducible queries such as ''Did your abdominal pain occur after eating?'' rather than a more open question such as ''When does your abdominal pain occur?''

The content accuracy reflected the correctness of the responses. In the ideal situation, all queries should get correct responses from VPs. This can be divided into two procedures. First, the queries should be recognized by the system correctly. Second, the queries should be matched to the corresponding response. The first procedure was implemented by WeChat's own voice and text recognition system. The implementation of the second procedure mainly depended on the logical operation formulae. By using Boolean logical operators to connect keywords, it is possible for different query approaches to be matched to the sole correct response.

In this study, the median content accuracy of all the VPs was 97.7%, indicating that only about 2.3% of the queries were answered incorrectly. The results were better than the results reported by Maicher et al. (2017). The main reason for problems was that a logical operation formula sometimes incorrectly classified queries into a related but incorrect category and so the wrong response was given. For example, a query about sputum odor was matched by the logical operation formula to the response regarding the general nature of sputum (''It is like mucus''). These deficiencies in the logical operation formulae needed to be identified and rewritten accordingly.

The results showed that more than half the queries without a response and more than half the queries with an incorrect response occurred in the Present Condition part, indicating that this part needs to be analyzed and revised the most. Present Condition, also called Present Health Status or Illness, includes disease onset, main symptoms, evolution of the disease, course of treatment, and so forth. Whether from the perspective of VP designers, nursing educators, or nursing students, Present Condition is one of the most difficult parts of history taking because it varies greatly

depending on the type of disease. From the point of view of VP designers, a Present Condition part of a VP should cover the typical characteristics of the disease the VP has and the VP responses must lead the students in the correct direction. This requires focused and unambiguous VP responses, which need to be continuously tested and modified. For nursing students to conduct satisfactory Present Condition history taking, basic knowledge is needed about the structure of the Present Condition section, as is knowledge about different diseases and the corresponding symptoms and how to apply theoretical knowledge to practical interviewing is very important. Although the SYSU-NEP VPs all have very common diseases, it remains difficult for nursing students with little clinical experience to grasp all the key points in the Present Condition part.

Based on the analysis of the underlying reasons for problems with the Present Condition parts of different VPs, it was found that many queries raised by the students were not relevant for the VPs, despite the fact that there were many cues for relevant queries in the VPs' responses. On one hand, the problem may be caused by the students. The students participating in the study had just covered the Health Assessment and Internal Medical Nursing curricula. They had little clinical experience. As a result, their understanding of specific cases was insufficient for focused and complete history taking, which resulted in many of their queries not being closely related to the VP cases. On the other hand, the design of the VP cases should be improved. Some cues were not sufficiently apparent. Some information was useless and/or misleading. This means that the VPs require further testing and correction in future studies.

The advantages of VPs, including safety, repeatability, and objectivity, have been widely discussed in the field of medical education. However, the high cost and complicated training and maintenance requirements have limited the widespread use of VPs. In addition, quality assessments have rarely been reported, although the quality of VPs is the basis of their usage. In this study, the SYSU-NEP VPs were shown to have a high degree of content completeness and accuracy. High-quality VPs could guarantee the user experience of students and help them to gain more experience by providing safe, objective, and repeated history-taking encounters. VPs are an ideal complement to the current nursing history-taking training. In addition, SYSU-NEP is characterized by ease of use, accessibility, and low cost because of the convenience and popularity of WeChat. It overcomes the problems of previous VPs regarding high costs and complicated maintenance, and it makes personal use of VPs possible, which could facilitate students' autonomous learning. The VPs based on WeChat are an innovation in simulation education, combining the convenience of social media and the fidelity of simulation training, and they have great future prospects.

The perfection of any software requires continuous testing to be carried out. The SYSU-NEP VPs were only tested among 57 students. Although these students generated 570 history-taking sessions and 13,312 queries, this is far from enough. In the future, we will test the VPs using larger samples of users after modification of SYSU-NEP.

Conclusions

VPs used for history-taking practice should have high levels of content completeness and accurate responses. The results of this study showed that the content completeness and accuracy of the SYSU-NEP VPs were satisfactory but had room for improvement. The queries without a response and with an incorrect response were also identified, which is helpful for the continuous improvement of the VPs. Essentially, the current VPs could satisfy the needs of students regarding history-taking practice. We will modify the VPs according to the results of this study and update the VPs continuously, and more high-quality SYSU-NEP VPs will be developed in the future.

SYSU-NEP is a valuable and cost-effective tool for the development of nursing students' history-taking skills. Students could access the SYSU-NEP VPs for free via WeChat using various kinds of mobile terminals and gain more experience before experiencing real clinical encounters. VP quality is the basis of the usage of VPs. Only reliable and high-quality VPs have values that are deserved to be applied and broadened. The results of this study help provide evidence for the development and quality analysis of future VPs. The improvement of the VP quality also provides a basis for more efficient use of VPs in future history-taking training.

Acknowledgments

The authors would like to express special thanks to the software engineers Guocai Li and Li Ling who provided technical support for the development of the study software.

Authors' contributions: All the authors participated in the study design and implementation and gave final approval for the submitted paper. J.L. and J.Z. were responsible for the statistical analysis and the drafting and revision of the manuscript. H.H., T.L., and J.Z. designed the virtual patient chat scripts. K.L. was responsible for the quality control of the study and the review of the virtual patients and the manuscript.

Ethical approval: The study was approved by the Department of Education Administration and the ethics committee of Sun Yat-sen University.

References

Association of American Medical Colleges. (2017). Effective use of educational technology in medical education. In *Colloquium on educational*

technology: recommendations and guidelines for medical educators. Retrieved from http://www.ttuhsc.edu/som/curriculum/documents/For-OurTeachers/aamc_effective_use_of_educ_tech.pdf. (Accessed 27 November 2017).

Barnett, S. G., Gallimore, C. E., Pitterle, M., & Morrill, J. (2016). Impact of a paper vs virtual simulated patient case on student-perceived confidence and engagement. *American Journal of Pharmaceutical Education*, *80*(1), 16.

Bateman, J., Allen, M., Samani, D., Kidd, J., & Davies, D. (2013). Virtual patient design: Exploring what works and why. a grounded theory study. *Medical Education*, *47*(6), 595-606.

Baumann-Birkbeck, L. (2017). Appraising the role of the virtual patient for therapeutics health education. *Currents in Pharmacy Teaching and Learning*, *9*(5), 934-944.

Cendan, J., & Lok, B. (2012). The use of virtual patients in medical school curricula. *AJP: Advances in Physiology Education*, *36*(1), 48-53.

Consorti, F., Mancuso, R., Nocioni, M., & Piccolo, A. (2012). Efficacy of virtual patients in medical education: A meta-analysis of randomized studies. *Computers & Education*, *59*(3), 1001-1008.

Gordon, M. (1994). *Nursing Diagnosis: Process and Application*. New York, NY: McGraw-Hill Book Company.

Guise, V., Chambers, M., Conradi, E., Kavia, S., & Valimaki, M. (2012). Development, implementation and initial evaluation of narrative virtual patients for use in vocational mental health nurse training. *Nurse Education Today*, *32*(6), 683-689.

Harder, B. N. (2010). Use of simulation in teaching and learning in health sciences: A systematic review. *The Journal of Nursing Education*, *49*(1), 23-28.

Harless, W. G., Drennon, G. G., Marxer, J. J., Root, J. A., & Miller, G. E. (1971). CASE: A computer-aided simulation of the clinical encounter. *Journal of Medical Education*, *46*(5), 443-448.

Hege, I., Kononowicz, A. A., Tolks, D., Edelbring, S., & Kuehlmeyer, K. (2016). A qualitative analysis of virtual patient descriptions in healthcare education based on a systematic literature review. *BMC Medical Education*, *16*, 146.

Isaza-Restrepo, A., Gomez, M. T., Cifuentes, G., & Arguello, A. (2018). The virtual patient as a learning tool: A mixed quantitative qualitative study. *BMC Medical Education*, *18*, 297.

Kononowicz, A. A., Zary, N., Edelbring, S., Corral, J., & Hege, I. (2015). Virtual patients–what are we talking about? A framework to classify the meanings of the term in healthcare education. *BMC Medical Education*, *15*, 11.

Liu, T., Luo, J., He, H., Zheng, J., Zhao, J. J., & Li, K. (2018). History-taking instruction for baccalaureate nursing students by virtual patient training: A retrospective study. *Nurse Education Today*, *71*, 97-104.

Maicher, K., Danforth, D., Price, A., Zimmerman, L., Wilcox, B., Liston, B., …, & Rizer, M. (2017). Developing a conversational virtual standardized patient to enable students to practice history-taking skills. *Simulation in Healthcare*, *12*(2), 124-131.

Stultz, J. S., Forder, M., & Pakyz, A. L. (2017). Virtual pediatric patient activities with randomized scenarios as an instructional tool for pharmacy students. *The Journal of Pediatric Pharmacology and Therapeutics*, *22*(6), 444-452.

Sun Yat-sen University. (2017). *SYSU Nursing Education Platform*. Beijing: Copyright Protection Center of China. 2017SR273307

Sweigart, L., Burden, M., Carlton, K., & Fillwalk, J. (2014). Virtual simulations across curriculum prepare nursing students for patient interviews. *Clinical Simulation in Nursing*, *10*(3), e139-e145.

Tencent. (2011). WeChat. Retrieved from https://www.wechat.com/en/. (Accessed June 2018).

Urresti-Gundlach, M., Tolks, D., Kiessling, C., Wagner-Menghin, M., Hartl, A., & Hege, I. (2017). Do virtual patients prepare medical students for the real world? Development and application of a framework to compare a virtual patient collection with population data. *BMC Medical Education*, *17*(1), 174.

Zary, N., Johnson, G., Boberg, J., & Fors, U. G. (2006). Development, implementation and pilot evaluation of a web-based virtual patient case simulation environment–Web-SP. *BMC Medical Education*, *6*, 10.

Zeng, F., Deng, G., Wang, Z., & Liu, L. (2016). WeChat: A new clinical teaching tool for problem-based learning. *International Journal of Medical Education*, *7*, 119-121.