IFMSA
International Federation of Medical Students' Associations

Standing Committee on Medical Education

Manual

for the National or Local Officer on Medical Education

"We aim to achieve excellence in medical education throughout the world."

5th edition (August 2009)

IFMSA General Secretariat
c/o WMA B.P. 63, 01212 Ferney-Voltaire CEDEX-FRANCE
phone: +33 (450) 04 47 59 · fax +33 (450) 40 59 37
email: gs@ifmsa.org · www.ifmsa.org

Bank Account details
Account 58.52.12.090 Beneficiary IFMSA
Bank ABN-AMRO SWIFT ABNANL2A
IBAN NL94ABNA0585212090
Bank Address: Coolsingel 119 · PO Box 949
3000 DD Rotterdam · The Netherlands

Editor's note:

This document is meant for internal use of National or Local Officers on Medical Education of the National Members Organisations of the International Federation of Medical Education (IFMSA) only.

Beside the regular updates of this manual, up-to-date information on medical education issues can be found in the SCOME-wikipedia at http://wiki.ifmsa.org/scome

To receive additional copies of the SCOME-manual, please contact the Medical Education Director of IFMSA at scomed@ifmsa.org or send a letter to:

IFMSA General Secretariat
c/o WMA
B.P. 63
01212 Ferney-Voltaire CEDEX
France

Tel. +33 – 450 40 47 59
Fax +33 – 450 40 59 37

1st – 3rd edition compiled by Jan Hilgers, Medical Education Director 2005-2006 & Liaison Officer on Medical Education Issues 2006-2008.
4th - 5th edition compiled by Nikos Davaris, Medical Education Director 2008-2009

Written by members of the Standing Committee on Medical Education

Special thanks to:
Ana Babic, Elisa Berioja, Maja S. Basnov, Vasile Bintintan, Lars Choritz, Nikos Davaris, Becca Fenech, Vlad Gavrila, Alexander Geube, Torstein S. Hansen, Jan Hilgers, Petar Kresimir Hodzic, Jonas Johannink, Katja Kovac, Katharina Kulike, Sebastian Manoleasa, Özgür Onur, Aydogan Orhan, Camille Piguet, Emily Rigby, Dani Rodriguez Muñoz, Carl Savage, Tina Schweickert, Jacco Veldhuyzen, Hans Jacob Westbye, Vanessa Wennekes, Jan Westermann, Patrik Zamecnik

Contents

History

Medical Education should be a concern of every medical student as it shapes not only the quality of future doctors, but also the quality of healthcare. The International Federation of Medical Students' Associations (IFMSA) has a dedicated organ which aims to implement an optimal learning environment for all medical students around the world – the **S**tanding **C**ommittee **O**n **M**edical **E**ducation (SCOME).
It was one of IFMSA's first standing committees from the beginning of its foundation in 1951. IFMSA SCOME acts as a discussion forum for students interested in the different aspects of medical education in the hope of pursuing and achieving its aim.

Important moments in the history of SCOME

- First policy statement of IFMSA: 1951-1970 - Impact of Technology on Health Education
- Declaration on Primary Health Care and Medical Education, 1979
- Policy declaration on Primary Health Care, 1980
- Policy Declaration on Medical Education, 1980
- Resolution on Medical Education, 1983

Mission

Healthcare is changing at an unprecedented rate and at multiple fronts. Technology has revolutionized archaic diagnostic and therapeutic procedures. Medical science has increased our understanding of the body and created an explosion of new information. Patients are increasingly questioning and less trusting their doctors. But medical schools are not or only slowly introducing changes in their curriculum. Teachers at many medical faculties are not educated to teach; they are doctors and mostly lack knowledge of how to show their skills to their students.
We question that students educated in a so-called traditional curriculum are able to face the needs of healthcare in a modern society. Scientific data show that modern medical curricula are a lot more likely to teach students in an appropriate way in order to create doctors equipped with various skills and knowledge. Although there are a number of innovative approaches to teaching medicine, partly based on findings of cognitive science, change in medical curricula occurs slowly and at few medical schools. The need for change is either not recognized or ignored in many universities.
As medical students are directly exposed to medical curricula, they should rightfully be assumed to be experts on their educational system, and should therefore have an influence on the creation of new curricula. From IFMSA's experience, it is often the medical students who are the strongest proponents for adapting their education to the needs of their community.
Here SCOME enters the game. We try to promote modern medical education. Convinced by many positive examples we go on that mission by teaching and training students, teachers, and professors, exchanging experiences and spreading information.
As a global grassroots organization of medical students locally active in more than 94 countries worldwide, IFMSA has made meaningful contributions to improve medical education over the last decades.

On our way to improve medical education

Are medical students throughout the world acquainted with the subject of "medical education"? Do they recognize their role in the field of medical education and can they really make a change? Most of the students get involved in the process of medical education as passive participants fulfilling their duties, but not scrutinizing the educational process itself. In this way, they are missing the unique opportunity that they have as "consumers"/"clients", to give their opinion on the educational process and thus provide the data about that system from "inside". In general, students show lack of interest and awareness for this important issue. This is the main reason why motivating students is our first goal.
Currently, there is a "reform-pandemic" in medical education going on worldwide. Medical curricula are going through changes in most of the medical faculties throughout the world. Since it is difficult to predict the results of these changes and it takes a long period, they are usually drawn from the reform process itself. It is of great importance that students are actively involved in this process because they can very early inform the faculty authorities of the disadvantages which the new system might have. Students should promote vital feedback which is essential for the development of an efficient medical education system. Student organization should exchange experience and data which will enhance their role as active participants in the reform process. In that direction, the role of SCOME is not to represent a mere talking point, but a central coordinative unit which will guide medical students throughout the world towards a better medical education. SCOME is not meant to be only a discussion forum, but "headquarters" that will analyze the current situation in medical education, set up strategy for action and council students how to put that strategy in action in their own countries.

In most of the cases it is rather hard to improve our educational system. In most of the countries there is no tradition of integrating students in faculty development. Sometimes they even are not members of decision-making bodies within the schools or they are only a minority in those. So statements and proposals of students do not have a high value for stakeholders. This situation is well known to most of us. Why would you write this here? Rather: We must be aware of this well-known situation causing multiple problems. Our strategy has to be adapted to these circumstances. How?
In the last years we worked mainly in three fields:

a. Locally

The most promising strategy for change is a local approach. Even if students do not have a majority within the faculty boards, students could convince deans, professors, teachers, and stakeholders to develop their education. It may not be possible to change the whole curriculum at once; but small changes in each of the different subjects will slowly but steadily also improve the curriculum as a whole. In a constructive and cooperative way one can find many small solutions to make life easier. To get some ideas of how to approach see "Policy Statements" and "Concrete suggestions" below and exchange experiences with other National Officers on Medical Education (NOMEs) using the SCOME-wikipedia at wiki.ifmsa.org/scome!

To enable NOMEs and Local Officers on Medical Education (LOMEs) to facilitate improvements on the local level one of the main activities within SCOME are training sessions. Training covers all fields within medical education, like assessments and

exams, evaluation, teaching and learning systems, problem-based-learning, community-based-learning, computer-based-learning, policies of government and ministries. It is important to provide as many training sessions as possible. These sessions are held on general assemblies, pre-GA-Workshops, on regional and national meetings and on special international training workshops.
What is the task of SCOME on the local level?

- Provide some additional courses that can be useful for medical students
- Facilitate discussions between students and faculty
- Participating in the evaluating process
- Improve medical education
- Contribute to the SCOME projects and start new national ones
- Collect local students' opinions and try to implement them
- Represent local students in faculty's and university's boards

b. International Projects

We have different types of international projects:

- **Database projects**
  The main objective of these projects is to collect information (e.g. about curricula, residencies) and to provide it to all, mainly on behalf of the internet.
- **Research in Medical Education**
  We support and encourage students to do research on the field of Medical Education. Therefore we work together with the scientific students' conferences where we initiate medical education sessions and provide workshops. Furthermore we have our own research projects.
- **Courses**
  After students found a lack of a special topic in their curriculum they initiate courses. Students also invite guest speakers and experts themselves. If these courses lead to a success and the interest of the students is high enough, the medical school will accept to integrate them in the curriculum eventually. There are also many courses run by students as peer education.
- **Training**
  To improve our knowledge and skills we organize training workshops e.g. on the topic of the implementation of the Bologna process in medical education.

c. International lobby

There are some international organizations dealing with medical education. We try to co-operate with them and to represent the students' thoughts and wishes on the international level. In some cases these ideas find their way back to the national and local level. Actually we are working together with the "World Federation for Medical Education" (WFME) and its regional partners such as the "Association for Medical Education in Europe" (AMEE). It is the task of the NOMEs to find out which possibilities they have to work in a similar way on the national level with their national associations for medical education.

➔ *Concrete suggestions*

Many students find situations in their medical faculties, which they would like to change. Often they succeed in changing the curriculum: courses in sign-language, medical ethics or even the whole revised curriculum at the University of Berlin ("Charité"), Germany, are examples for these changes.

However, many students don't know how to implement changes. First, you have to define the problems. The next step will be to prioritize them. These are some guidelines that have produced positive experiences for students all over the world. You should pick up the ones appropriate for your country and situation. You can also modify them according to your needs.

In order to convince authorities, you can do the following things:

a) Attract faculties' attention by having debates, conferences including the ones that introduce new models, preferably the model you want to change into.
b) Approach student friendly faculty members, so that they will lobby on behalf of students to achieve positive changes.
c) Check what's going on in other schools and show the beneficial models to your authorities.
d) Show teachers that you want to work together with them, not against them.
e) Conduct objective evaluation studies, displaying the results scientifically, e.g. surveys, literature reviews.
f) Get a person (professor or not) in your school, to advice you on your ideas.
g) Get support (recommendation letter, motion etc.) from official bodies like IFMSA, WHO or any official local body, for your cause, including for example, student representation.

In order to change, you need to have large number of people working together, hence the importance of attracting attention of the students:

a) Find specific attracting tasks/ areas of change for students to work on so that after/ during these projects they will contribute to changes in a more permanent way. These tasks should be initially small, so as not to overwhelm the new members.
b) Student bodies, joint councils.
c) Have debates, conferences including the ones that introduce new models, preferably the model you want to change into.

You can also get the help of the community by showing the importance of medical education for their health. In order to attract attention of the community you can do the following:

a) Campaigns, media, press conference, posters, T-shirts, leaflets, publications.
b) Maybe strike…

During the process changing things, you can follow those strategies

a) Use university/ governmental regulations/laws pertaining to medical education, to your advantage.
b) Show the changes so that people will be motivated to join and further your work.

Steps for change

General approach for NOMEs

(These steps were developed at EOM 99, Slovenia, by the SCOME working group „Steps for change").

Remark: You should modify your actions according to local situation

1. Observe present situation. Use IFMSA recommendation to get the help from existing students' organizations and then try to get students' voice, if you don't have it already.
2. Once you have infiltrated into academic structure, find any relevant information about the structure and try to understand the way it functions (especially the responsibility line). In case you cannot do it, consult the local expert in the field (law, etc.)
3. Having knowledge about your own system, search through SCOME and other databases to find out more information on ME. Search also for schools that offer electives in ME (e.g.: McMaster (Canada), Maastricht (The Netherlands), etc.) or other schools whose systems or curricula can be most easily applied onto your present local system. Various databases can be found online.
4. Gather the Task Force group of at least 5 students that would invest some time in helping you within the forthcoming actions and events.
5. Set the stage at your School and do the preparations to process surveying
6. Let the Task Force Group to analyze the data and prepare the presentation. You should strive to do your best in obtaining recognition of your work from local, national and international institutions. Meanwhile, consult coordinator for Electives in ME, about possibilities for you to visit the chosen School. Ask him/her about application forms, and all the necessary information you might need. If needed search the Internet.
7. Work on fund-raising and contact relevant person at your School and ask if they are willing to help you with financial support.
8. Find appropriate time in your schedule to attend the chosen elective at the chosen school. If there aren't any gaps within your curriculum, exploit the possibility to have the course you will attend recognized by your local School.
9. Once you are there be very active and never stop asking. Collect the bunch of papers about the host system you will take back home.
10. Obtain the contact person and persuade at least one professor to come to your School and have a lecture during the Round table or Workshop you are going to organize. Moreover, find the local student officer willing to come to your country and present his/her own experiences documented by his personal academic achievements.
11. Try to establish some kind of agreement between your school and the hosting school.
12. While you are still there make assignments for Task Force group in order to fulfil prerequisites at the local School for the organization of Round table or Workshop on the issues you are concerned with.
13. Coming back home coordinate the Task Force group and continue preparations for Round table or Workshop. Set the tentative date accordingly with the schedule of the mentioned professor and student. Promise them board and lodging free of charge, and if necessary fund-raise for that purpose, or even ask your School officials to help you ensuring it.

14. Make and send invitations for all persons and officials that you consider to be important and who can make decisions concerning changes of the system at your School (University, Ministry of Education/Health/Science and Technology etc.)
15. Advertise the Round table or Workshop in local and state media. Invite the representatives of the media at the Media Conference to be held after the meeting. Don't forget to inform the students of your School.
16. At Round table or Workshop present the results of the surveys (better if you have international comparative study), whom you kept unveiled until that time. Wave in front of professors' eyes with documents that recognize your work and efforts either from local or international institutions (IFMSA, WFME, UN, EU etc). Let the invited professor and student have their show. Don't forget to find the person for professional simultaneous translation, if necessary.
17. Held the Media Conference together with invited professor and student about your efforts, actions, achievements and future plans.
18. After the Round table or Workshop start active lobbying among professors and relevant officials. Inform them about forthcoming events concerning Medical Education and persuade them to go there if possible.
19. Send more students to chosen example school or even to other schools having different systems in order to obtain additional information. Meanwhile, initiate foundation of test group of students willing to be taught in different manner.
20. In collaboration with already lobbied teaching staff, try to implement new system into either optional or compulsory curricula.
21. Make it sure that the students you have chosen are among the best ones and able to achieve the desired results with the new system. They should be open-minded, convinced that the change should be done, and consider the new system as a step forward.
22. Try to involve the teachers that support the new system and that are willing to do their best in order to achieve the best results.
23. Let the system work for a period of time, adjust what is not working and at the end of that period evaluate it. Make the comparison with the old system. Organize another Round table according to steps mentioned above, where to present your conclusions and discuss. You should invite the teachers that opposed the system before. Give or send the final report of the Round table to the members of the Faculty Council, and other persons you consider important.
24. Together with Task Force Group and supportive teaching staff design the Proposal of the Program for change.
25. With the positive results that support your position, try to persuade the Executive Board of the Faculty Council (Dean and Vice-Deans) to accept your arguments and present the Program for change in front of Faculty Council, as the voice of the Dean it surely more persuasive then the voice of the student. If they don't accept present the Program yourself.
26. During all steps continue to inform teaching staff and regular students about the whole process and lobby at the teaching staff in order to get majority when it comes to the voting phase. Try to involve ordinary students into this lobbying process as well.

IF YOU SUCCEED – CONGRATULATIONS!

IF YOU DO NOT - TRY OTHER WAYS AND CONTACT US.

Structure

In general

The **IFMSA Director on Medical Education** (SCOME Director) is elected each year by the National Member Organisations at the IFMSA General Assembly (GA) in August. She/he co-ordinates the work that is done by National Officers, project co-ordinators and others. The SCOME Director is responsible for the SCOME meetings at biennial IFMSA General Assemblies.

The SCOME Director can appoint members of the committee as coordinators for special tasks, if he/she is not able to fulfil them by him/herself.

To represent IFMSA and SCOME towards international associations in the field of medical education (e.g. the Association for Medical Education in Europe (AMEE) and the World Federation for Medical Education (WFME), the "**Liaison Officer on Medical Education Issues**" (LOMEi) was established in 2000. The Liaison Officer is member of the executive boards of the AMEE and the WFME. Another task of the Liaison Officer's work is to support the SCOME-D during the year, in meetings and at the GAs.

SCOME and IFMSA are in close contact to other non-governmental-organizations dealing with higher and medical education such as the "World Federation for Medical Education" (WFME), "United Nations Educational, Scientific and Cultural Organization" (UNESCO), the "Association for Medical Education in Europe" (AMEE), "Association for Medical Schools in Africa" (AMSA), "Pan-American Federation of Associations of Medical Schools" (PAFAMS), "Association for Medical Education in the Eastern Mediterranean Region" (AMEEMR), "South-East Asian Regional Association for Medical Education" (SEARAME) and "Association for Medical Education in the Western Pacific Region" (AMEWPR). The national and international communication between IFMSA and these organizations should go through the respective liaison officers.

Each National Member Organization of IFMSA elects a **National Officer on Medical Education** (NOME). Her/his task is to co-ordinate and to encourage local or national activities in the respective country. The NOMEs are also responsible for communicating with and reporting to the SCOME Director.
NOMEs are recommended to attend the international IFMSA meetings in March and August respectively. At these meetings, they network with other NOMEs, exchange ideas and attain new knowledge and motivation to bring back home to the Local and National Committees.

The **Local Officers on Medical Education** (LOME) are in charge of local improvement in Medical Education and related activities at the different local medical faculties of a National Member Organisation. They are elected locally and are responsible for tackling local problems. They are advised to form local working groups, whose work they co-ordinate. The LOMEs shall communicate with and report to the NOME.

The **Regional Assistants** are appointed by the Medical Education Director after consulting the respective Regional Co-ordinator.
The main tasks of the Regional Assistants are:

- To keep in touch with the national SCOME-groups in the region

- To provide SCOME-members of the region with his/her personal and professional support
- To encourage and assist the development of SCOME in the region
- To assist new-comers
- To encourage and maintain the co-operation within the region
- To inform SCOME-members about the latest developments in the region
- To identify problems in the region and work with SCOME-director in order to dissolve them
- To establish priorities in the region and work on the development of a regional plan of action
- To facilitate SCOME-sessions at the Regional Meetings and at IFMSA's General Assemblies
- To encourage and maintain the co-operation between the regions and to share common interests between the regions

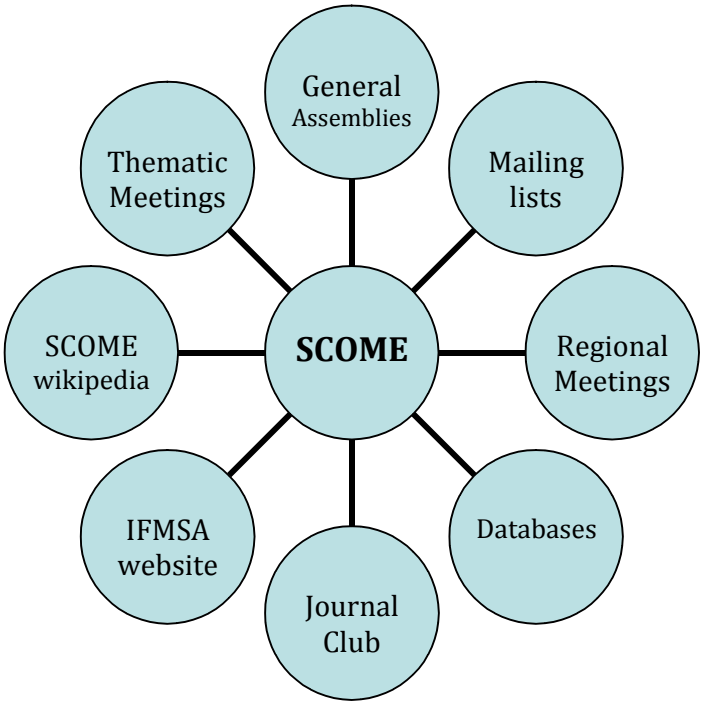
The regions of SCOME are:

- Africa
- The Americas
- Asia Pacific
- Eastern Mediterranean
- Europe

In the SCOME-wikipedia (wiki,ifmsa.org/scome) you can find which region your country belongs to.

Internal communication

As mentioned before our best tool to learn and improve our skills, find ideas and learn about experiences is communication within each other. **In many cases specific problems in different countries and medical schools are similar to each other.** A solution also might have been found somewhere before. Our task is not mainly to invent new solutions and ideas, but quite often it is worth listening to other solutions and to employ them in the own setting. To guarantee this exchange of experiences and knowledge, we created several settings and strategies:



a. General Assemblies

During the general assemblies every Standing Committee has its own working committee meetings. The biggest part of the time we use to exchange experiences and learn from each other. This is a forum, where everyone gets a chance to present his problems and solutions. We further work in training workshops, where a small group creates and prepares a training. To make this process more effective NOMEs should present their work in written reports and/or posters. Very important are also the informal meetings in between.

b. Regional meetings

To increase the communication between NOMEs and to integrate more LOMEs in the international work we encourage local committees to organize regional meetings. On those meetings the participants can exchange knowledge and information. Furthermore these meetings should be used as an opportunity to provide trainings in medical education and in general issues.

c. Thematic Meetings

Thematic Meetings can be organized in international, regional or subregional level to work on specific issues. For example IFMSA organizes the Bologna Process Workshops in order to form policy concerning the changes brought to Medical Education from the implementation of the Bologna Process in Europe.

d. Mailing lists

To enable communication between our March and August meetings we created two e-mailing-lists (see below).

e. Journal Club

The IFMSA-SCOME journal club regularly informs the SCOME-community about recently published articles related to Medical Education Issues and writes reviews of the articles. To contact the current co-ordinator please send an e-mail to scomed@ifmsa.org!

f. Databases

Sometimes it is more effective to contact certain NOMEs and not the whole mailing list. To get the contact information we created a contact database for the whole IFMSA (www.ifmsa.net). To get information from this homepage you need a login and a password. The President of the National Member Organizations provides these passwords to their national officers. To get it, please contact your NMO president. The SCOME-Director needs these pieces of information, too. He/she uses this information to prepare certificates, recommendation letters etc.

g. SCOME-homepage in IFMSA website

You can find information on projects and activities within SCOME and all contact information to the co-ordinators at the SCOME-website within ifmsa.org.
The aim of the website is to collect and provide as much information as possible in the field of medical education. You can access the SCOME-homepage at www.ifmsa.org (section: medical education).

h. SCOME-wikipedia

Since May 2006, SCOME has a new interactive element in its homepage.
It is a wikipedia, an open-source collection of information about medical education in general; projects, SCOME is running; reports and much more. In contrast to other wikipedias, IFMSA's SCOME-wikipedia aims to share ideas, to describe, and to discuss recent developments in the field of medical education worldwide rather than to provide scientific_articles.

Users shall describe interesting courses and concepts that are implemented in their local curriculum and be able to provide further information.
National Officers are invited to share projects their National Member Organisation (NMO) is running in the field of medical education. They shall also write articles about their NMO, the role of SCOME in their NMO and on local level, about the educational system in their country, admission criteria for studying medicine, about the different faculties in their country and any other thing they think is relevant to other NOMEs or LOMEs.

Also the SCOME-wiki may serve as a forum to share biannual or annual reports of the Standing Committees since it can easily be edited by any member of the NMO.

You can access the SCOME-wikipedia easily at wiki.ifmsa.org/scome.

First steps in the SCOME-wikipedia

1. On the top right corner, click on "Create an account or log in"
2. Please choose your real name as user name or your real name and the position that you currently have or have had (e.g. "Jan Hilgers, SCOME-D 2005-2006").
3. Choose a password and enter both your e-mail address and your real name.
4. Click the "Create the account"-button

Now that you have created an account, you can start editing and writing articles.

In the search field enter the search item (e.g. "Reports") and hit enter to find out if the article already exists.

On top of each article there are several different tabs:
- article,
- discussion,
- edit,
- history,
- move, and
- watch.

Clicking on the "article" tab will show you the article itself.
In the "discussion" tab you can find further questions users had for other users to discuss the content of the article.
Clicking on "edit" will show you the source code of the article to edit the text or the layout. You can add paragraphs, links or any additional information that you think is helpful for others.
The "history" tab will show you older versions of the article and who uploaded changes.
With the "move" tab you can move the article to another entry (e.g. from "Repotrs" which obviously is misspelled to "Reports").
If you click on "watch", the page will be added to your watchlist. Future changes to this page and its associated Talk page will be listed there, and the page will appear **in bold letters** in the "list of recent changes" to make it easier to pick out.

TIP: If you search "All pages" you will be forwarded to a list of all existing articles in the SCOME-wikipedia.
As next step you should get used to the syntax ("language") of the wikipedia system.
At this website http://en.wikipedia.org/wiki/Wikipedia:How_to_edit_a_page you will find an overview of the commands to edit and layout articles.

If you upload files (not pictures -> see the link above how to use pictures!) you will find them in the wiki when searching for "All files". To link the file to an article you should copy the whole URL of the file with an explanation text as you can find in the "Asia Pacific Regional Meeting 2005 – 2006" article.

The most practical way to learn about the function of the SCOME-wikipedia is to start writing and editing articles and to continue "learning by doing".

Writing and editing an article

Headlines are sub-dividing the article in sections. There are different kinds of subdivisions:
- "==Section==" generates a headline of the first level named "Section".
- "===Subsection===" generates a headline of the second level called "Subsection".
- "====Sub-subsection====" is a level 3 headline called "Sub-subsection".

The first **section** should be called "==General Information==" and be a summary of the article.

When the **name of the article** is mentioned for the first time in the article, it should be written in bold letters.

You can write words in **bold** letters by adding ''' before and after the words to be bold (e.g. if "**Reports**" shall be written bold, add '''Reports''').

If you want to write in *italics*, add '' before and after the words to be in italics (e.g. if "*Reports*" shall be written in italics, add ''Reports'').

You can also add **shortcuts to other articles of the SCOME-wikipedia**. If you, for example, want to link to the article called "Reports" in your article, add [[Reports]] to the text and it will appear like this "Reports" in the article.

If you want to link to the article without mentioning "Reports" but "Report" instead you can do so by adding [[Reports|Report]] and "Report" will appear in the article.

The sign "|" divides the original name of the article from the words you put in the article instead.

You can also add links to articles not existing yet. These links will appear in red letters instead of blue ones in the article.
You should only add one link to each article in every article. If "LOME" appears 20 times in your article, please only add a link to the article "Local Officer on Medical Education (LOME)" when the term is mentioned for the first time!

Also **linking to other websites** (or files beside images in the wikipedia) is possible.
Adding [http://www.bvmd.de] will create a link to the website of the German Medical Students' Association. But this link will appear like this "[1]" in the text.
To add some more information or to link words or phrases to other websites you must add an explanation. [http://www.bvmd.de German Medical Students' Association] will appear like "German Medical Students' Association" with the explanation of the link behind the space.

To add a **list**
- like
- this
- one,

you must add "*" in front of each new line.
The text in the edit format will then appear like
*like
*this
*one,

To add a **numbered list**
1. like
2. this
3. one,

you must add "#" in front of each line.
The text in the edit format will then appear like
#like
#this
#one,

A single empty line in the edit format creates a **new paragraph** 1,5 lines below the last line like

this.

If you add <br /> you can **switch into the next line** like
this.

If you want to **save your changes** to the article, please first push the "Show preview" button and go through the article. Read the article again! Is the layout as you wanted it to be? Are all the links working, can you see the tables and pictures?
Only if everything looks fine, push the "Save page" botton!

To use a **picture** that you have uploaded in an article, use [[image:*filename.jpg*]] or [[image:*filename.jpg|Description of the image*]] at the position where the picture shall be placed.
If you add "|thumb" ([[image:*filename.jpg|thumb|Description of the image*]] a small thumbnail of the picture will appear.

For **further information** on how to write and edit an article, please read the article: "How to edit an article" at the SCOME-wikipedia.

In the help section of the English wikipedia you can also find tutorials how to add pictures and tables to your articles.

Contact details

- IFMSA Director on Medical Education:
  Any question related to the work of the Standing Committee can be send to scomed@ifmsa.org
- Liaison Officer on Medical Education Issues (LOMEi or LO MedEduc):
  If you want to contact WFME, AMEE or any other international medical education institution or association, please contact the LOMEi at lme@ifmsa.org
- SCOME Regional Assistants
  Since there are no e-mail aliases available for Regional Assistants yet, e-mails to them go via scomed@ifmsa.org and are forwarded to the specific Regional Assistants
- PreGA workshops
  Since there are no e-mail aliases available for SCOME-D-Support person e-mails to them go via scomed@ifmsa.org

Mailing lists

SCOME-Server:
ifmsa-scome@yahoogroups.com
Every student interested in Medical Education is asked to subscribe on this server. To subscribe send a blank e-mail to:
ifmsa-scome-subscribe@yahoogroups.com

NOME-Server:
This server is a restricted server. To become a member to the mailing list, you should be a National Officer on Medical Education. If so, please contact the SCOME-Director; he/she will put you on the server.
On this mailing list, internal IFMSA-SCOME issues are discussed. These can be administrative, membership in IFMSA-SCOME or reporting issues.

Projects and workshops

Beside the various projects taking place at local level to improve medical education at the faculties, there are a number of international projects and workshops taking place under the umbrella of the IFMSA network.

You can find information about ongoing local or national projects and activities in the SCOME-wikipedia (Category: Project). Furthermore the reports of the NOMEs are available online there providing further information about work done in the NMOs.

Transnational projects

Curriculum Database (CDb)

The aim of the project (CDb) is to offer an opportunity to medical students to find information about the ways of studying and teaching medicine in other faculties and countries. The CDb contributed to the increasing need for information on different medical universities on a world-wide scale without the hassle of using multiple servers in different countries. Moreover the CDB offers opportunity for NOMEs to get a comparison of different curricula within their nations and across counties; thus using this data to help improve their curricula. One benefit of this project is that the information being provided includes both an official and students' point of view.
We think that, among students, there is a need of knowledge in the field of medical curricula. CDb will be a tool that is going to help:

- Student representatives to find out details about better and more developed medical education systems and,
- All medical students that intend to take part in a student exchange programme.

CDb will be on the SCOME homepage, so that any medical student from all over the world can access it. First it should consist of the information gathered from the NMOs that are IFMSA members. In the long run, we would like to extend the CDb to NMOs who are not involved in SCOME and to countries that are not IFMSA members. This process is subject to change if new and improved technical methods are found.

http://curriculumdatabase.osmcluj.ro/
Project co-ordinator: Ciprian Dospinescu, IFMSA-Romania
Contact: http://curriculumdatabase.osmcluj.ro/contact.php or cd@ifmsa.ro

Daisy project

"Daisy project – Margarita" is a pilot educational community based project which is proposed on a voluntary basis to medical students and it takes place in cooperation with International Association of Health Policy (IAHP).
As it is implied by the project's name, it is comprised by a central activity ("core") which is attended by all participants and 4 peripheral activities which are optional according to the students' preference.
"Training in communication skills" is the project's core. All participants attend the weekly sessions which are coordinated by a professional psychologist, specialized in the field of health. Some of the topics being discussed during the sessions are: patient-doctor communication, ways of dealing with uncooperative patients, announcement of bad news etc.
"Health Education Intervention in Secondary Schools" is the oldest of the peripheral activities. The medical students are trained throughout the year on health intervention applications, concerning topics such as STD, AIDS, contraception and general public health issues.
"Nurse Aid" is one of the most popular of the peripheral activities. The students are working as aids of previously trained and informed nurses following the workload of the day and the ward

they are allocated to. Thus, they are trained in nursing and clinical skills and in the same time they are exposed in personal contacts with patients, relatives and co-workers.
"Medicine in Community" gives the opportunity to students to get to know with Primary Health Care and preventive medicine. In cooperation with general practitioners who act as trainers, the students familiarize with the role of doctor as advisor and medical information source for the community.
"Research in Social Medicine" is the last of the peripheral activities of the "Daisy Project – Margarita". It enables students to train and practice on research techniques, by participating also in other current research projects.
The expected outcome includes:

- The exposure of medical students to the real working conditions in the field of community-based medicine and the application of theoretical knowledge in order to deal with public health problems.
- The increase of awareness and the development of skills concerning the communication with the patients and their relatives and the approach of the patient as psychosocial entity by medical students.
- The recognition of the doctor's role as health professional towards the direction of disease prevention and the promotion of health, as science researcher and as active citizen with social responsibility.
- The development of critical scientific spirit within the framework of inter-professional cooperation

International students will have the opportunity to be informed about the program and its outcome during the annual meetings of IFMSA and mailing lists, HelMSIC and the project's website, periodical publications etc.
All the students interested participating in the project have to fill in the application form. This is essential because after this date the program will be about to start running.
All medical students are eligible to participate in the program. However, there are some criteria to meet, due to limited number of places for participants. Students willing to participate have to complete an application form and write a motivation letter explaining why they want to participate in this program and what their expectations are. Students from all years (1st until 6th), who meet the requirements, can be accepted. The medical students are informed about their selection by written announcements, phone or e-mail.

http://www.helmsic.gr/en/projects/margarita.php
Contact: thessaloniki@helmsic.gr or daisy@helmsic.gr

Influence of Studies on Students' Health (ISSH)

The aim of the ISSH project is to determine the level of deterioration of health among medical and non-medical students as a result of stress-related factors and to develop programs to reduce or even prevent it. The organizational structure and curricula of a program as well as student workload can be considered major factors affecting the stress levels of medical students. Initially the project aimed to establish the level of correlation between countries, types of education and type of studies in relation to stress related diseases. The project was based on a 25-question survey which was statistically analyzed and compared between the various countries involved. The results show that appropriate changes need to be made in medical curricula in order to improve the quality of education and student life; there is a need for stress management at medical facilities.
The results obtained from the survey will not only be used to help change medical curricula and improve education through stress management initiatives, but will also serve to encourage a pro-active approach by the medical students themselves on these issues. To this end, we will organize stress management courses for interested students in addition to offering courses in time management.

A second version of the ISSH-questionnaire has been developed and distributed in 2007.

Every interested student, medical school or NMO can participate. If interested, you should contact the international coordinator who will then provide you with the necessary information on the current status of the project.

Contact: Mohammad Shalaby, m.s.shalaby@hotmail.com
or join the project's yahoogroup at issh@yahoogroups.com!

International Students' Network on Ageing and Health

The International Students Network on Ageing and Health was created to serve as an umbrella for all the IFMSA activities in the field of Ageing, for promoting all our projects, at the local and international level, and all the events where we have participated, proving our active involvement regarding this issue. The Network for Ageing and Health intends to actively involve students, professionals and educators interested in working in the field of ageing and health.
ISNAH is mainly our webpage (www.ageingnet.tripod.com) and our mailing list. Our activities are divided into four major components:

1. Information dissemination:
   To promote awareness of health care professionals on global population ageing and special needs of older people for example by a network homepage, publications, and interviews with experts. Build knowledge, have easy access to information for students on ageing and health, publishing for students, career development).
2. Community and Research projects:
   By starting projects on active ageing and volunteer services for older people on the grass-root level, make a research database, arrange clinical student exchanges. To stimulate intergenerational contact and to do this through local community based projects. To create Public Health Projects on community level to increase the health and status of older population. Make a research database of internships opportunities on ageing and health and publications. IFMSA project focus on older people, database on research on ageing and health by students, intergenerational contact, promoting old people friendly health care, co-operation with other actors in the field, grassroots level, ethics, healthy ageing, functional ability, volunteer projects).
3. Curriculum Development:
   To change the medical education and other health associated education to face up to the ageing population growth.
   To promote life course, interdisciplinary, health promotion, community based, focus on gender, on culture, on ethics in medical education on ageing and health.
4. Advocacy about Priority of Ageing issues towards Policy Makers: (UN principles of Older Persons, ageing on development agenda, HR, equal access to health care, gender issues)
   To advocate for the rights of older persons, including old age care in national health policies and to put these issues on the priority agenda.

To participate you just need to send an e-mail to one of the coordinators and you will be part of our mailing list. Then you can send us presentations of any project or event related to the field of Ageing and this way you will have the opportunity to share your experience with other people willing to contribute together with you at making some real changes regarding the life of the ageing population.

Project co-ordinator: Jesús Mateos del Nozal, IFMSA-Spain
Contact: isnah@ifmsa-spain.org

Residency Database (RDb)

The Residency Database project's aim is to facilitate international medical students and young doctors to reach information concerning the residency system and application procedure to different countries of the world. Furthermore, it provides to the Residency Database site's

visitors with the chance to compare the advantages and the disadvantages of the many different countries' residency systems, along with the opportunity for further research in this field.

The whole idea is about the construction of a Residency Database (RDb), where every medical student and young doctor will find information about the residency system, the financial state and the application procedure for a residency position in many different countries. The importance of this project lies to the fact that a continuous growing number of medical scientists from all over the world desire to specialize or sub-specialize in a medical field outside his/her own country and they face a lot of difficulties in finding the proper source of information. The Residency Database project will facilitate these people to have an easy and quick access to the relevant information.

Project's structure
There are two groups of people who work on the realisation of the project: the Project's Participants and the International Coordinating Team.

What is the Project's Participants role?
A Project Participant can be any of you who would like to fill out the Residency Database (RDb) Questionnaire with the information concerning your country. The RDb Questionnaire is the cornerstone of the RDb project and it requires information and references concerning the residency system, the financial state and the application procedure for a residency position in your country. After filling out the RDb Questionnaire and sending it to us, it will be evaluated for its adequacy of the provided information and the stating of references by the International Coordinating Team. Then, the provided information will be uploaded to the Residency Database web-site.

What does the International Coordinating Team do?
1. Finalise the RDb questionnaire
2. Gather the completed questionnaires
3. Evaluate the received data
4. Promote the project

If you are interested in learning more or participating in the project (by completing the Residency Database Questionnaire with the required information concerning your country), please send us an e-mail to: residency_database@yahoo.com . Your contribution to the project is not only highly desired but also necessary for its realisation!

How can I participate in the International Coordinating Team?
If you are interested in playing an active role to the RDb project's materialisation, please send an e-mail to ifmsa-rdb-subscibe@yahoogroups.com in order to be subscribed in our working list and to participate in the discussion. There your opinions count!

http://residency-database.helmsic.gr/
Contact: rdb@helmsic.gr, rdb@ifmsa-spain.org

Think Global

Global health education aims for students to have an understanding of the broad determinants of health and healthcare delivery. Think Global aims for all future healthcare professionals to have an understanding of global health. The project will work with students involved in the IFMSA (International Federation of Medical Students' Associations) and provide them with opportunities to learn about global health in the context of their clinical and extracurricular activities.
The aims of the project are
1. For students of health related subjects to develop the knowledge and skills necessary to work effectively in a global society

2. To promote the importance of global health education and advocate for increased opportunities for students to learn about global health within their curriculum
3. To enable students to organise events and activities for their peers to learn about global health education on a local and national level
4. To build a network of students concerned by global health
   For all students to have access to information on global health issues and the education opportunities available to undergraduates

What is global health?
There is much debate about the definition of global health (or international health as it is sometimes known) and there are many different interpretations of the term. We believe that global health is a broad discipline that develops students' understanding of the local, national and international determinants of health and healthcare delivery.

Through studying global health, students examine the wider influences of health such as poverty, debt, globalisation, healthcare financing, human rights, famine, environment, violent conflict and the movement of populations. Global health draws from a number of disciplines including politics, economics, sociology, demography, anthropology, epidemiology and philosophy.

Why is global health education important for students of health related subjects?
The health of the world's populations is governed by a number of different yet interrelated factors. Societal factors are increasingly acknowledged as important determinants of the health of individuals and populations, but this acknowledgement is often not reflected in the scope of training that future health professionals receive. Global health, in some small way, aims to make up for this disparity. Students who have studied global health are better equipped to understand the root causes in addition to the clinical manifestations of ill health.

Global health teaching will also help to foster a generation of health professionals who are committed to health for all, as enshrined at the International Conference on Primary Health Care, Alma-Ata, in 1978. We see health professionals as having a commitment not only to their patients but also to the health of society as a whole. The growing number of health professionals who are committed to global health equity can form a powerful group of advocates for health for all.

Globalisation is changing the structure of societies and the way in which decisions about health are taken. Many societies are becoming more multicultural, and global health teaching helps medical students to understand both the reasons for increased population movement and the social, economic and cultural factors underlying patients' ill health.

Decisions about health and healthcare are also increasingly made through global agreements such as TRIPS and GATS, and it is important for health professionals to understand the influence of such global policies on their work.

Globalisation means that medical students and health professionals are increasingly likely to work outside their own countries. Global health helps them to understand the different societies and health systems where they may work, and thus adapt better and more quickly to their new surroundings.

What do students need to know to work effectively in a global society?
The content of medical curriculum differs within and between countries and it is hard to establish learning objectives that are applicable to all students involved in IFMSA. 'Think Global' has identified the following as the core areas necessary for students' to work effectively in global society.

To understand how global health influences patient and community health and clinical medicine
To understand how to relate global issues to patient cases
To know how to access information on global health issues and education
To appreciate the role of doctors in advocacy for health
To understand the influence of cultural background on patient and community health

If you are interested in being involved in Think Global, please email the project coordinator,
http://www.globalhealtheducation.org/
Contact: thinkglobal@ifmsa.org

International Meetings and Conferences

Bologna-Follow up workshops

Since April 2003 IFMSA and the European Medical Students' Association (EMSA) have organised six medical students' conferences on the implementation of the Bologna directives to medical education.
In summer 2004 IFMSA and EMSA agreed on the *Megève paper*, the first position paper ever on the Bologna process in medicine. This paper has been used by other organisations when working on their own statements since.
In summer 2005 a conference on "Quality Assurance" took place in Copenhagen, Denmark, where more than 40 students from 15 countries participated in. Outcome of this workshop was a policy paper on quality improvement in medical education.
In 2006, the follow-up conference took place in Bristol (UK) and a consensus outcome-based "European Core Curriculum – the Students' Perspective" was written. It has been published in the "Medical Teacher" one of the leading international journals in the field of medical education.
In 2007, the meeting took place in Amsterdam (The Netherlands) and participants were discussing the Bachelor and Master structure for the study of medicine. Consensus points were identified and another policy paper written.
After some difficulties after the last meeting, this year's conference has been organised by the IFMSA only. In July 2008 we have met in Berlin (Germany) and focused on the issue of "International quality labels as a way to improve mobility".
In 2009, IFMSA and EMSA jointly organized a Bologna Process Workshop in Cordoba (Spain) with the topic "Bologna Process in Medical Education: beyond 2010"

The outcome documents of the above meetngs can be found in the "Policy Statements" section in this manual from page 31 on.

World Healthcare Students' Symposium (WHSS)

The World Healthcare Students' Symposium (WHSS), is a great new initiative. The four most prominent health care student associations, EPSA, IFMSA, EMSA and IPSF as well as international nursing students' associations have joined forces to put together this event. Health care students from all over the world are invited to attend this symposium, the aim of which is to discuss future cooperation between both health care professions and health care students. WHSS represents the largest international meeting between health care students working towards a common goal.

The vision of the students participating in the project is a future in which health care professionals worldwide cooperate with their colleagues in other health care disciplines for the benefit of their patients.

In 2005 the first symposium was held in Malta. The Maltese Medical Students' Association and the Maltese Pharmaceutical Students' Association organised a symposium which educated and inspired 250 future health care professionals, from pharmacy, medicine and nursing, to be activists and advocates for this vision.
The "Statement of beliefs" written at the conference in Malta can be found in the "Policy statements" section of this manual. The second symposium took place in November 2007 in Albufeira, Portugal. The third one will take place in November 2009 in Egypt.

Reports

Since the August Meeting 2006 in Zlatibor (Serbia) it is compulsory for National Officers to hand in reports of the activities and projects of SCOME in their NMO in order to remain "active" status in the SCOME-Database at www.ifmsa.net. The annual reports from the Medical Education Director, the Regional Assistants, and the National Officers as well as reports from the General Assemblies' SCOME-sessions and the Regional Meetings can be found at the SCOME-wikipedia (search item "Reports").

Since 2007-2008, efforts have been made to improve the reporting process and make optimal use of the data collected. Two forms have been created, the "Work in Medical Education Report Form - MERF" and the "Projects Report Form - PRF". Data collected through the above forms is available online in the SCOME website and Wikipedia.

Two times per year, before each General Assembly, the IFMSA Director on Medical Education informs the National Officers on Medical Education about the report they should send and the deadlines set. The mailing list used for this purpose is the **NOME-Server**. If you are a NOME, contact the SCOME Director through scomed@ifmsa.org to add you there.

Policy statements

IFMSA position paper on student assessment

This IFMSA statement seeks to delineate the problems medical students are facing on an international level regarding assessment in/of medical education

Objectives:

1. To outline the skills a medical student should acquire during his or her medical studies in order to become a good physician.
2. To outline the problems in assessment of medical students that hinder the acquisition of such abilities.
3. To define the pedagogical role of assessment in medical education.
4. To find alternative methods or improve the ones already existing in order to overcome problems of objective 2) above.
5. To support any other proposals that would help to achieve improvements in students' assessment.

I. IFMSA considers it necessary for a medical student to acquire the following skills during his or her medical education:
- To be able to assimilate, integrate and apply medical information in the manner most profitable for the patient and society.
- To be able to bear in mind the humanitarian and ethical aspects of any of his or her decisions.
- To be able to perform a meticulous clinical examination.
- To be able to attach the due importance to the doctor-patient-relationship.
- To be able to undertake efficient interaction with other members of the medical professions.

II. The traditional methods of assessment in medical education confront us with the following problems:
- They do not allow enough space for the development of the full individual in each medical student.
- Instead of promoting the students' ability to learn actively and solve problems, some assessment methods rather induce a passive attitude in the student. Reproduction or, even worse, recognition of information is given more importance then analysis and problem solving (MCQ).
- On the other hand assessment methods with direct teacher-student-contact can never guarantee full objectiveness. Furthermore we find that the lack of standardization between medical faculties in different countries limits the mobility of students.

III. IFMSA believes that student assessment should transcend its present dimension of solely passing or failing students to one that is more pedagogically oriented. This should mean that the assessment would be a platform for motivation without undue competition: The student have the possibility to see that what will be required from him or her in any kind of exam is of relevance to his or her future work as a physician. Moreover there should be a feedback for both student and medical teacher providing both with information on the level the student has reached in his or her medical education. Assessment should allow the student to view the patient in his or her entirety (i.e. without labels of medicine, surgery, etc.).

IV. IFMSA considers the following as possible solutions to the above problems:

IMPROVING EXISTING METHODS OF ASSESSMENT

1. Oral exams should be performed as comprehensive exams, viewing the patient in his or her entirety. Objectivity should be increased by using exam commissions instead of single examiners and keeping minutes of each exam.
2. Essays should be corrected according to a standardized answer sheet. They should be patient-centred.
3. Practical exams should comprise basic clinical skills.
4. MCQs should never be used as sole method of assessment. They should only be used provided there is continuous evaluation and feedback.
5. Assessment should always be based on a variety of methods. All these methods should follow a standardized protocol in an effort to maximize the objectivity of the method. Anonymity should be safeguard as far as possible. Students should receive feedback about their performance in all exams undertaken.

INTRODUCING ALTERNATIVE METHODS OF ASSESSMENT

IFMSA is of the opinion that alternative ways of assessment in medical education should play a much larger role than they have so far had. Among the methods to be taken in to consideration should be:

1. OSCE (Objective structured clinical examination), involving the testing of various practical skills via a number of stations, each having a checklist to assess the performance of the student.
2. Continuous assessment and feedback on wards, contributed by all members of the team.
3. Continuous assessment using the same set of questions throughout the whole curriculum (students from different semesters would expected reach different levels).
4. Paper cases with several steps (each subsequent page would give further information on the “paper patient”).
5. Assessment of communication skills using video cameras.
6. Utilization of the group process in tutorial groups as a means of assessment in order to strengthen collaboration and reduce competition among students.
7. Introducing quality assessment of curricula and medical teaching staff by the medical students themselves that has official and substantial bearing on the rewarding of teaching posts.
8. Rotation of examiners in a regional group of medical schools or presence of external examiners, so as to improve objectivity.

GENERAL COMMENTS

Preparation for medical assessment should allow enough free time to develop the full social and cultural potential in each medical student. Students should be provided with guidelines and a framework of studying. The methods of assessment and the minimum requirements for passing should be made available to the students at the beginning of the course of studies. Flexibility concerning the sitting for exam sessions should be guaranteed.

Assessment should lead to one universal degree for medical doctors.

The better the assessment of medical students is, the better is the quality of future medical care to be provided for the whole society.

IFMSA Recommendations on Implementation of the Continuous Medical Education in Medical Curricula

Adopted by the participants of the 5th IFMSA Workshop on Medical Education: Life Long Learning, The Former Yugoslav Republic of Macedonia, October 1999

Definition of Continuous Medical Education (CME)

Medical education never ends. It doesn't stop upon graduation from medical school. The needs of the society in which we live are changing, and so is the information available for medical education. Our ultimate goal is to produce better and more competent doctors who are able to adapt themselves to the needs that the future brings. It is not possible to acquire all the necessary medical knowledge in the short period of university studies. That means that the medical schools' most important task is to prepare future doctors to work in any kind of changing environment. The principal is to learn how to learn.

Task Description

At this moment, the reality is that medical schools all over the world to give their students all of the knowledge available in the medical field during their undergraduate studies. This fact, however, does not guarantee that future doctors will be competent enough to approach a patient after graduation. The result of this kind of teaching are overloaded curricula, which still cannot teach ALL of the knowledge and skills needed. The problem that medical schools today do not prepare students with adequate skills on how to continuously learn, how to find and select and judge the newest available medical information, how to cope with new technologies, how to deal with the changing environment concerning communication skills, law, community needs and so forth. The question to ask, then, is how to balance the importance of theoretical medical knowledge with the clinical skills needed to be a doctor.

IFMSA's Wishes and Recommendations

IFMSA specifically recommends the fostering of self-directing learning skills, critical thinking skills, interviewing skills and communication skills. These communication skills should emphasize not only strong doctor-patient but also strong doctor-doctor and doctor-community relationships. Teamwork in this world is a growing need, as is peer-education and evaluation. Other important goals that we should strive to promote include the knowledge to use new technologies, management skills, practical skills, basic research skills (knowledge about scientific methods and research), and skills how to use all available information services (including the internet and libraries).
Medical students need to learn how to select and judge the available information. Future doctors can only set good priorities if they have the goals of the community in mind. We should specifically be educated on how to listen to society. In developing the core curricula, we must realize that it is and must be dynamic. What is "core" today may not be what is "core" in 20 years or more.

Policy/Position on the Impact of Technology on Health Education

Following a round table discussion, 50th General Assembly Meeting, 2001 Aalborg

We recognize that technology impacts health care education, research, and science educators in the areas of research, classroom teaching and distance education. While the overall effect is not yet fully assessable, the presence of technology in so many different aspects of the profession makes it important to more clearly recognize and appreciate its current and potential role.

IFMSA recognizes the following things:

- While there is no assessing body to monitor the presence of technology in this field, and information technology is fast creating an affect, IFMSA feels that the sense of direction of the impacts it creates has not been spared from the chaos and distress that accompanies this unprecedented era.
- Biomedical knowledge and clinical information about patients are essentially unmanageable by traditional paper-based methods, and to a growing conviction that the process of knowledge retrieval and expert decision making are as important to modern biomedicine as the fact base on which clinical decisions or research plans are made.
- Information technologies can be educators' tools in finding creative ways that encourage students to self-test, self-question, and self-regulate learning in helping them to create solutions to complex problems.
- Information technologies are providing new opportunities for linking medical schools around the world for sharing computer-based learning materials. Information technologies open a wide horizon for acquiring and expending medical knowledge originated in any part of the world without limitations of time, space or distance.
- Information technologies have lead to the improvement of evidence-based medicine.

IFMSA urges for a creation of an international independent monitoring board by international organization such as the World Health Organization (WHO) and the Word Federation of Medical Education (WFME) to lead the sense of direction of technology in the right path.

IFMSA stresses that the use of computers and information technology in medical education should be regarded as an additional tool and must never be a goal in itself but part of flexible learning. On the contrary clinical medical education should always be centred on direct patient contact and bedside education. While we urge for direct patient contact we believe that using simulations would also benefit the student in training.

IFMSA will work with different organizations and institutions world wide in developing a comprehensive online resource that wouldn't contain an overload of information and that can be monitored for content following international standards.

IFMSA understands the advantages produced by information technologies in data retrieval and research management and urges that this be geared to serve the needs at the international level.

IFMSA will communicate to all international organizations, national organizations and local organizations urging to ensure that the best possible training is given to the students by the educators while they integrate the use of computers into the system as different teaching methods need different approaches. Traditional methods in some cases are proven to be more effective and these methods shouldn't be replaced in order to just keep abreast of the technology and careful consideration and study should be done before replacing traditional methods. IFMSA also urges strongly the integration of the technology as part into the education.

IFMSA would like all the educators to take the students into consideration while developing or planning for new information technology, as students are the best resources. Students worldwide are thereby requested to take an active role while any developments to this effect come in place.

IFMSA will through its network work on linking medical schools and organizations for sharing computer based learning material but would like an international organization to be a part of it to monitor the standards thus creating an International self study resource with no boundaries in information and which will provide equal opportunity to countries that cant afford or keep abreast with the technologies.

IFMSA while recognizing that information technologies have improved the evidence-based study strictly urges that Technologies should not estrange us from our humanity or the noble profession. We believe that medicine is an art by itself.

IFMSA believes that information technology is educating the patient and urges for the creation of a course in the medical curriculum of how to handle a patient who has obtained his knowledge good or bad through the technology.

IFMSA stresses and urges all students and everyone in this profession to ensure that the ethical and moral aspects are safeguarded.

ALL IN ALL, IFMSA RECOGNIZES THE IMPORTANCE OF IMPACT OF TECHNOLOGY ON HEALTH EDUCATION AND ENCOURAGES MEDICAL STUDENTS, EDUCATORS AND INTERNATIONAL MONITORING ORGANIZATIONS TO TAKE THE INITIATIVE TO TAKE ROLE IN THE CURRENT PHASE TO DIRECT TECHNOLOGY IN THE RIGHT PATH AS WE DO NOT KNOW HOW THIS WOULD BE IN THE FUTURE.

Policy/Position on Implementing International Standards in Basic Medical Education

Following a round table discussion, 50th General Assembly Meeting, Aalborg 2001

We recognize that to have quality development in basic medical education implementing international standards is vital.

IFMSA views the current situation in medical education as follows:

- Basic Medical Education courses conducted in about 1600 medical schools worldwide varies from one school to another. But only a little number of these medical schools worldwide are subject to external evaluation and accreditation procedures.
- These result in a very different level of medical knowledge, skills and behaviour acquired by graduates of medical schools.
- Globalization is helping to produce a new vision of cooperation for common goals and specific advantages without precluding the local culture, language and various requirements responsive to local realities.
- Medicine itself is universal and requires a universal identity to work on it. We will be doctors for all.
- There is clearly no global system that provides the implementation of international standards.

IFMSA describes international standards in basic medical education as follows:

- IFMSA defines the word "standards as both a goal (what should be done) and a measure of progress toward that goal (how well it was done).
- IFMSA keeps in mind that "'implementing international standards' does not imply uniformity of medical schools or a threat to the fundamental principles that medical education has to address the specific needs in a given social and cultural context."
- IFMSA describes the report of World federation For Medical Education (WFME) on Defining International Standards in Basic Medical Education as a reference point for international standardization.

- IFMSA states that providing globalization by collaborating with international organizations such as WFME is a must in medical education.
- IFMSA urges all medical students to do their best for their own education.
- IFMSA describes taking role in implementation of international standards in basic medical education as one of the ways to achieve the best in medical education.

IFMSA urges medical students all around the world to take the initiative to reach international standards locally, nationally and internationally as follows:

- IFMSA aims to take a step further for the implementation of the international standards in basic medical education by,
- Helping professionals on medical education to investigate problems associated with implementation of international standards in basic medical educations and adapt strategies.
- Helping to raise awareness on the international standards in collaboration with international organizations.
- IFMSA urges medical students focus more on the international standards by organizing forums, workshops and training programs where recommendations of the professionals for the stage of implementation locally, nationally and internationally be presented.
- IFMSA suggests that the report of WFME on defining international standards in basic medical education be translated into different languages making it possible for everyone involved in medical education to understand
- IFMSA urges medical students to work in collaboration with International organizations to introduce the report on international standards to local, national and international authorities.

IFMSA advocates all national and local authorities in medical education to get involved in the stage of implementing international standards in basic medical education as follows:

- IFMSA calls upon all national and local authorities in medical education to view these standards as a way for individual faculties to get integrated with international recommendations and as a method to measure themselves.
- IFMSA calls upon all national and local authorities in medical education to implement international standards in their own curriculum in synthesis with their regional needs.

ALL IN ALL, IFMSA RECOGNIZES THE IMPORTANCE OF IMPLEMENTING INTERNATIONAL STANDARDS IN BASIC MEDICAL EDUCATION AND ENCOURAGES MEDICAL STUDENTS TO TAKE THE INITIATIVE TO TAKE ROLE IN THE IMPLEMENTATION PHASE AS THEY ARE THE ONES TO CONTINUE THIS IN THE FUTURE.

The Bologna Declaration and Medical Education

A Policy Statement from the Medical Students of Europe, Megève, France, July 4th, 2004

Outcome of the third workshop on the Bologna process organized by the International Federation of Medical Students' Associations (IFMSA) and the European Medical Students' Association (EMSA)

Adopted by the IFMSA General Assembly in August 2004

Summary:
Most points of the Bologna process are welcomed by the medical students of Europe. Medical education is in many ways in a special position when it comes to implementing the changes, and we would like to emphasize the importance of three points:
1. A common system for quality assurance of medical education in Europe would increase mobility and improve the quality of tomorrow's physicians.
2. We are concerned about the negative implications of a two-cycle structure on medical education. Harmonization of medical education in Europe is crucial whatever system exists.
3. Student involvement is essential at all levels of the process.

Background:
The Bologna Declaration of June 1999 established the following objectives:
1. Adoption of a system of easily readable and comparable degrees
2. Adoption of a system essentially based on two main cycles, undergraduate and graduate.
3. Establishment of a system of credits – such as the ECTS system
4. Promotion of mobility by overcoming obstacles to the effective exercise of free movement
5. Promotion of European co-operation in quality assurance
6. Promotion of the necessary European dimensions in higher education

These objectives are, according to the Declaration, to be attained "within the framework of our institutional competencies and taking full respect of the diversity of cultures, languages, national education systems and of University autonomy."

Two subsequent meetings were held where additional points were added:

Prague communiqué, May 19th, 2001:
7. Integrate life long learning into the overall strategy
8. Higher education institutions and students
9. Promoting the attractiveness of the European Higher Education Area

Berlin communiqué, September 19th, 2003:
10. Establish a European research area

Our point of view:
1. **Adoption of a system of easily readable and comparable degrees**
   The medical degree is already easily readable and comparable within the EU through the Medical Education Directive EC 93/16. This can further be improved through implementation of the Diploma supplement.

2. **Adoption of a system essentially based on two main cycles, undergraduate and graduate**
   We are concerned about negative consequences in implementing a two-cycle structure in medical education. Current efforts to update the medical curriculum recognise that the early integration of basic and clinical science is essential to produce better doctors. This provides a meaningful context in which to integrate current research with basic care. It is also supported by adult learning theory, which acknowledges the difference between having factual knowledge and being able to apply it to a real-life situation. The implementation of a two-cycle structure must not be allowed to cement the traditional

division between the basic sciences and clinical sciences, as described in the Flexner Report of 1910.
In those countries with a two-cycle structure for medicine, students should be required to have a Bachelor of Medicine or bachelor degree with academic equivalence to enter the Master of Medicine, to ensure the quality of those who graduate as physicians. Without a European consensus on implementing the two-cycle structure in medicine, two degree systems will result, seriously hampering easy readability and mobility. Some medical curricula teach subjects over several years. The implementation of the two cycle structure in such curricula will lead to an artificial separation of these subjects, limiting mobility. This must be avoided by introducing guidelines for bachelor and master content. One model for this is described in the idea of a European *Core Curriculum* in medical education, as mentioned by the British General Medical Council in 1993 and defined by AMEE Education Guide no 5.
At the same time, we recognise the value of having a unified degree structure for higher education in Europe. For medical education, we recognize potential improvements in flexibility and mobility, and more opportunities to choose a master degree.

3. **Establishment of a system of credits – such as the ECTS system**
   Establishment of ECTS can easily be done in most European countries, and has already been implemented at several European medical schools. We require that a European grading system must be researched and evidence-based to determine the most appropriate manner in which to assess medical students. A correct and consistent implementation of ECTS and the grading system is of great importance for mobility and quality of assessment throughout Europe.

4. **Promotion of mobility**
   Mobility is desirable on all levels of medical studies, from individual courses or clinical clerkships, as in today's Erasmus program, to entire degrees. The recognition of common guidelines for the content in the degrees would increase mobility. The Lisbon Convention has established a means to get degrees and courses recognized, and this is an important step to increase mobility.

5. **Promotion of European co-operation in quality assurance**
   We urge the ministers to agree on a system for quality assurance in Europe. The task of creating this system should be given to independent experts. For medical education this could, for example, be AMEE. Student involvement in this process is absolutely necessary. Quality Assurance can be achieved through the establishment of common guidelines for the content of the degrees and an adoption of, for instance, the WFME Global Standards for Quality Improvement. A common European system for accreditation of medical schools would establish and maintain high educational quality and provide a means for comparison between different medical schools. We welcome harmonization, but preserving the diversity of the individual medical schools in Europe is of utmost importance.

6. **Promotion of the necessary European dimensions in higher education**
   We recognise that the cultural diversity of Europe is currently reflected in the way medicine is taught in different countries. We hope that the future European medical education is based on a holistic view of the complex world we are living in and reflects the fast changing environment and growing knowledge base of tomorrow's physicians. More focus on language learning would enhance communication in the profession and improve mobility.

7. **Integrate life long learning into overall strategy**
   The healthcare environment is rapidly changing making continuous professional development essential after graduation. The role of medical schools in preparing their graduates for this process cannot be stressed enough. We see the utilization of modern teaching methods and self-directed learning as setting the foundation for life long learning.

8. **Higher education institutions and students**
   The recognition of students as *"competent, active and constructive partners"* is a step forward in increasing the quality of medical education. We welcome this invitation of the ministers for more active student participation which we hope will be welcomed and implemented *at all levels.* We feel strongly about our education and that of the generations to come. We are the key to shaping tomorrow's education. We will, after all, be tomorrow's teachers.

9. **Promoting the attractiveness of the European Higher Education Area**
   Through establishing a common European system for quality assurance and safe-guarding easily readable and comparable degrees, Europe will be more attractive for both European and non-European students.

10. **Establish an European research area**
    In our knowledge-based society, research is one of the pillars of the modern university. We see the potential benefits of the establishment of a European research area and appreciate its importance in academia.

In conclusion, we strongly welcome most points of the Bologna process, which encourages flexibility, mobility and quality assurance. We are concerned about the negative implications of the two-cycle structure on medical education. However, not implementing the two-cycle structure should not be an excuse not to implement the rest of the Bologna process. We emphasise the importance of common European guidelines for the content of medical degrees. The integration of the basic sciences and clinical worlds from day one is paramount to our success as future physicians.

We look forward to active participation in Europe's drive towards the highest quality medical education possible.

Quality Assurance in Medical Schools

Moving from Quality Assurance to Quality Improvement
Quality Assurance Workshop, EMSA/IFMSA, Copenhagen (Denmark), July 6-10, 2005

Executive Summary:
The medical students of Europe are strongly committed to supporting the quality movement in medical education. However, we challenge the European Higher Education Area to set their sights higher. There is no guarantee that quality improvement naturally follows upon quality assurance. Rather, quality assurance is a first step towards the implementation of quality improvement. The move from quality assurance to quality improvement must be consciously and systematically implemented. This effort begins with involving stakeholders and widespread dissemination of evaluation results and continues with the establishment of a common core curriculum, the systematic use of improvement tools and the universal understanding that the ultimate goal of medical education is to improve the health of our citizens.

Background:
In May of 2005, the European ministers of education met in Bergen to discuss the further development of the Bologna Declaration and left the meeting committed to quality assurance in higher education. They adopted the standards and guidelines for quality assurance proposed by ENQA (European Network on Quality Assurance) for the European Higher Education Area, as well as the creation of a register for quality assurance agencies. They also agreed that students must be given a more active role in the implementation of the Bologna process. We, as the medical students of Europe, applaud these decisions to assure the quality of higher education. However, we would also like to challenge the ministers of education, the European Higher Education Area and our own medical universities to not stop at quality assurance but move to quality improvement in order to keep up with the changing needs of healthcare.
In July of 2005, students from IFMSA (International Federation of Medical Students' Associations) and EMSA (European Medical Students' Association) and EMS Council (European Medical Students Council) met in Copenhagen for a follow-up workshop to the Megève policy statement to reach consensus on quality assurance in medical education. We define "*Quality"* as the "characteristics of a function, process, system or object that is fulfilled when compared to predefined goals or standards." "*Quality assurance"* therefore is a "way to warrant that the predefined standards are met." "Q*uality improvement"* is defined as "a continuous process to review, critique and implement changes."

Quality Assurance must be implemented in all medical universities
We agree with the ministers in their decision to support ENQA in creating standards for accreditation agencies around Europe. This makes the quality of the accreditation agencies comparable. However, there may be a place for professional organisations in the accreditation process.

WFME Standards
We support the implementation of certain baseline criteria to which all medical schools must adhere. Certain aspects of medical education are universal, regardless of the university where one is educated. The guidelines set forth by the WFME in their Global Standards for Basic Medical Education from 2003 should be the starting point.

Course evaluations
A prerequisite for the accreditation process is validated and regular course evaluations, something which, at the present moment, is far from universal in Europe. We see this as an absolute requirement.

Quantitative and qualitative aspects of evaluation
The evaluation process should include both quantitative and qualitative reviews which allow for feedback in a constructive way. The focus should be on the development process which forms the basis for learning and improvement.

Transparency of the evaluation process
Results must be published, circulated and used as a basis for decisions on improvement.

Core Curriculum
Due to the increasing movement of physicians throughout Europe, it is in the interest of quality of patient care to establish a common *Core* Curriculum within medical education. The Core Curriculum would become a minimum standard for all physicians throughout Europe, regardless of where they were educated. A core curriculum in no way limits the

individual autonomy of any medical school. It still allows all countries and regions opportunities for individualization in curricular decisions and pedagogical methods.

Licensing Exam
Another quality assurance measure that has been discussed within Europe is the implementation of a common licensing exam. This is not something we support at this time. The degree of variance over Europe between the current medical education systems is too great to be measured in a standardized examination without having first established standards such as a common core curriculum.

Student Involvement
Students must be involved in all aspects of quality assurance. Not only are we customers of our education, but we are adamant that the education we receive should help us serve our future patients optimally. For this reason, students should not only contribute data but also be included in analysis and dissemination of the results. Awareness that evaluation has an effect on the curriculum is an important motivator for everyone involved.

Stakeholders
In many countries the evaluation only assesses the relationship between students and teachers. Other stakeholders, such as medical professionals and patient groups, should be included in the evaluation as well. They are valuable informants regarding which competencies a medical graduate should have.

Quality improvement must be a consequence of Quality Assurance
In order to improve medical education in a systematic and effective way, quality assurance is a first step on the way to quality improvement. Rather than determining the level of quality at a fixed point in time, quality improvement is a continuing and dynamic process to review, critique and change in order to make medical education better. Improvement must be built into medical schools as a continuous process that exists at all levels of a medical school, from the individual course to the entire program. It is not sufficient to simply be aware of the current state – rather it is the ability to improve and develop that determines success.

Define the mission
To start moving towards quality improvement, the faculty, students and staff need to actively define and express the mission and goals of the medical school. The mission must be reflected in all educational interactions.

Make improvement a natural part of the existing system.
Medical schools often make curriculum changes without continuous reflection on the consequences regarding the education of tomorrow's doctors. This is something that we cannot accept given the large body of knowledge that exists in the area of improvement science which allows us to evaluate the changes. There is the difference between change and actual improvement.

Conclusion
Quality assurance is a process which sets minimum standards of quality in education and is a requirement of the Bologna Process. Quality assurance requires transparency of process where results must be published and disseminated widely. In creating a competitive Higher Education Area, quality assurance is an essential factor. ENQA and the WFME standards are important building blocks in this effort.

However, if Europe is to create a competitive Higher Education Area that is to last, we must move from quality assurance to quality improvement. This will not happen without a conscious effort because improvement does not follow naturally upon quality assurance. Constructive feedback and improvement tools are prerequisites. If we succeed, we will have in place systems to continuously review, critique, and implement changes. The ultimate goal of medical education is to improve the health of society. We should always remember that quality improvement of education is quality improvement of health care.

WorldMaPS Statement of Beliefs

Summary:
The participants of the above symposium agreed on several principles:
1. Healthcare should be patient centred and multidisciplinary.
2. Healthcare professionals must have appropriate knowledge, good communication skills, be team players and have an empathic approach.
3. Healthcare education must be practical, multidisciplinary and state of the art.

The symposium:
The first joint symposium of world healthcare students was organised by IFMSA (International Federation of Medical Students' Associations), IPSF (International Pharmaceutical Students' Federation), EPSA (European Pharmaceutical Students' Association) and EMSA (European Medical Students' Association), and took place in Malta from the 7th to 12th November 2005. The symposium brought together 230 medical, pharmacy and nursing students from 42 countries in an international forum. The conference was intended to generate understanding and discussion between the professions, develop skills and awareness of concepts in multidisciplinary settings, and to create student advocates for a cooperative multidisciplinary approach to patient centred care to optimise health outcomes.

Explanation:
Patient Care
The participants of the symposium agreed that good patient care takes into consideration the individual needs of the patient. There needs to be effective communication within the healthcare professions and with the patients themselves. This conference considers the STEEP principles described by the Institute of Medicine in "Crossing the Quality Chasm: A New Health System for the 21st Century" as an appropriate model for the delivery of excellent patient care. The STEEP principles are Safe, Timely, Efficient, Effective, Equitable, and Patient-centred.

Healthcare professionals
Healthcare professionals knowledge should at all times be relevant, current and evidence based. Healthcare professionals need good communication skills in order to be a team player in a multidisciplinary environment. A priority for all healthcare professionals is to be patient focused which requires an interactive and empathic approach.

Education
Healthcare education should be practical with maximum exposure to clinical settings from the beginning of the curricula. It should be interactive with a variety of teaching methods including problem based learning. Healthcare education should mirror the multidisciplinary working of healthcare teams which includes learning together in
order to gain an understanding of other professions. Healthcare education should be state of the art. It should include the latest evidence-based practice and be delivered according to the latest developments in education.

In order to achieve all the above, a motivational learning environment must be created where members of the healthcare team are working together as equals from the very beginning of their careers.

European Core Curriculum – the Students' Perspective

Note: You can find the original document with all the references and a list of the participants of the conference in the SCOME-wikipedia! The full text is also printed later in this manual from page 110 on.

Summary

Medical Students of Europe have written an outcome-based core curriculum identifying 9 domains with 76 learning outcomes for graduates of European medical schools.

Introduction

Over the last few years, in innovative medical education, focus has shifted from acquisition of knowledge towards the achievement of concrete learning outcomes. Society and stakeholders are now more interested in the final product of the educational programme rather than the processes used to reach them.
More than 40 medical students' representatives from 15 European countries met for the 5th IFMSA/EMSA Bologna follow-up conference in Bristol (UK) to discuss and write an outcome-based "European Core Curriculum from the Students' Perspective".

Explanation

Participants agreed on the demand for a common outcome-based "European Core Curriculum from the Students' Perspective".
They have identified 9 domains with 76 learning outcomes to be covered in the course of medicine in medical schools of Europe.
The domains are:

- Clinical skills
- Communication
- Critical thinking
- Health in society
- Life-long learning
- Professionalism – attitudes, responsibilities, and self-development
- Teaching
- Teamwork
- Theoretical knowledge.

The core curriculum also includes a preamble explaining the goals and objectives of IFMSA/EMSA in writing the core curriculum.

The core curriculum will serve medical students all over Europe as a framework to be adjusted for specific national and local needs. It serves as a common basis aiming to improve the quality of education, healthcare and mobility, therefore furthering the establishment of a European Higher Education Area.

Considering that the dynamic nature of the medical field needs to be matched by education provided, the development and evolution of the core curriculum will not end with the adoption of this policy statement. We therefore do not include it in this form

but attached to this statement. In case there will be major changes of the core curriculum, the plenary will be asked to re-adopt the core curriculum.

UNESCO Guidelines for Quality Provision in Cross-border Higher Education

Summary

The Standing Committee on Medical Education (SCOME) of the International Federation of Medical Students' Associations (IFMSA) broadly accepts the suggestions put forward by the United Nations Educational, Scientific and Cultural Organization (UNESCO) on Cross-border Higher Education, Paris 2005.

Introduction

Cross-border education provides a global perspective and a unique opportunity to obtain a holistic outlook on one's future career. We believe that quality provision is vital in the medical field due to the impact of medical education on healthcare provision.

Explanation

The Standing Committee on Medical Education (SCOME) of the International Federation of Medical Students' Associations (IFMSA) broadly accepts the suggestions put forward by the United Nations Educational, Scientific and Cultural Organization (UNESCO) on Cross-border Higher Education, Paris 2005. Cross-border education provides a global perspective and a unique opportunity to obtain a holistic outlook on one's future career. We believe that quality provision is vital in the medical field due to the impact of medical education on healthcare provision.

Development of quality assurance mechanisms is important in medical education, as reflected in the IFMSA policy statement on „Quality Assurance in Medical Schools" (Copenhagen, 2005) and our involvement in the development of the WFME Global Standards in Basic Medical Education. We welcome the movement to extend quality provision to cross-border education.

As active partners in the field of Medical Education, we embrace the inclusion of guidelines for student bodies. This recognition acknowledges our role as integral to the provision of high quality education and the consequent impact on healthcare. Student bodies should be at the forefront of such initiatives, as they represent the receivers of this form of education, and are best equipped to provide input. The IFMSA has an international perspective, which places us in an ideal situation to work on the provision of quality cross-border education.

WHO/WFME Guidelines for Accreditation of Basic Medical Education

Summary

The Standing Committee on Medical Education (SCOME) of the International Federation of Medical Students' Associations (IFMSA) supports the WHO/WFME Guidelines for Accreditation of Basic Medical Education.

We think that accreditation of medical faculties is important to improve standards, assuring the quality of medical education and also establishing basic criteria to increase the mobility for medical students and physicians.

Introduction

In 2004 the WHO and the WFME have launched a strategic partnership to improve medical education worldwide by setting up an international task force on accreditation. The results of this task force's work have been formulated in a set of guidelines for accreditation of basic medical education institutions and programmes.

Explanation

The Standing Committee on Medical Education (SCOME) of the International Federation of Medical Students' Associations (IFMSA) supports the WHO/WFME Guidelines for Accreditation of Basic Medical Education. We think that accreditation of medical faculties is important to improve standards, assuring the quality of medical education and also establishing basic criteria to increase the mobility for medical students and physicians.

Still, we believe that special attention should be drawn to the following:

1. The purpose of accreditation should be to help medical schools improve their standards in medical education. If a medical school does not meet the standards, a deadline should be set and suggestions given on how to meet the standards. If the medical school still does not meet the standard, then it should have consequences for the school.
2. Student participation in the process of accreditation is important because they are receiving and participating in the education, and therefore they have a unique perspective to offer on medical education. They are the ideal group to ask for feedback on education and for suggestions for improvement.
3. The duration of the school's program should be the maximum time frame between two accreditation circles.
4. An international framework should supervise/ensure that the process of accreditation is carried out according to the same standards.

The Bachelor and Master structure in Medicine

Note: You can find the complete document with all the references in the SCOME-wikipedia!

The implementation of a Bachelor/Master structure is one of the most controversial aspects of the Bologna Process in Medicine. Increased engagement of all stakeholders is essential to ensure that the quality of medical education within Europe and consequentially, patient safety does not suffer. Given the conditions and prerequisites outlined in this statement of beliefs, medical students of Europe conceive that the implementation of the Bachelor/Master structure is possible. However, one must be aware of possible negative consequences if implemented without proper consideration and care.

This statement of beliefs expresses the opinion of the European medical students as discussed at the 6th IFMSA/EMSA Bologna Process Conference. It aims to serve as basis for further discussion and debate on the Bachelor/Master structure for medicine and raise awareness throughout the profession about the action lines as laid out in the Bologna Declaration and its follow-up documents.

We believe it may be possible to implement the Bachelor/Master structure in an integrated curriculum. However, it is vitally important that the implementation of the Bachelor/Master structure does not negatively impact either upon integrated or upon non-integrated curricula.

In order to achieve harmonisation of medical education in Europe it is necessary to agree on core learning outcomes to be achieved by graduation. These common core outcomes would

constitute the European Core Curriculum in accordance with relevant European regulations such as directive 2005/36/EC to be achieved by all European graduates. Local academic traditions and priorities should however be encouraged and these additional curricular elements should be clearly defined.

A European Core Curriculum is a prerequisite if the Bachelor/Master structure is to be implemented in Medicine. In addition, to secure patient safety in the context of student mobility between Bachelor/Master cycles, assessment of student competencies needs to be evidence based.

A Bachelor/Master system may enable students from non-medical Bachelor courses to enter Master of Medicine training. These students however would need to demonstrate the core competencies of a Bachelor of Medicine before entering the Master of Medicine course

Whilst it should be possible to enter the Master of Medicine after successfully completing Bachelor courses other than medicine, it must be stressed that the study of medicine should be considered as a continuum. Therefore, the study of medicine should be considered as a whole – Bachelor of Medicine and Master of Medicine together.

Implementation of transparent internal and external quality assurance measures in compliance with generic and profession specific quality standards is needed to achieve quality improvement of education. This is the key to building mutual trust, recognition of qualifications and ensuring the safety of European patients. However, any quality assurance procedure should not be unnecessarily costly, withdrawing precious resources from the actual education.

Financial consequences of implementing a Bachelor/Master structure must be considered. In particular, students should not be subject to increased tuition fees associated with procedures required to implement the system.

The Bachelor/Master structure should not be an obstacle to improve, develop and reform medical curricula. In itself it does neither contradict nor negatively impact on integrated curricula. It rather is its thoughtless implementation that may lead to adverse effects on educational outcomes and patient safety.

IFMSA's specifications to the WFME Global Standards for Basic Medical Education

Note: You can find the complete document with all the references in the SCOME-wikipedia!

The International Federation of Medical Students' Associations (IFMSA) supports the "Basic Medical Education WFME Global Standards for Quality Improvement". The Students' specifications are an adaptation of this document seen from the students' perspective. We, as medical students, are strongly committed to support the quality movement in medical education. IFMSA has developed the "Students' specifications" to the Global Standards and adopted them as the "Basic Medical Education WFME Global Standards for Quality Improvement – Students' specifications".

The IFMSA represents medical students from all over the world, as recognized by the United Nations (UN) and the World Health Organisation (WHO). It brings together more than a million medical students from over 90 countries on all continents. Its mission is to offer

future physicians a comprehensive introduction to global health issues as well as to improve the education of tomorrow's doctors. We believe that Quality Improvement is essential in order to achieve these objectives.
The IFMSA, through its Standing Committee on Medical Education (SCOME), has extensively dedicated its work to the implementation of Quality Assurance and Quality Improvement. This work has resulted in a continuous expression of the students' perspectives through several policy statements.

As mentioned in the Copenhagen Paper, we define "Quality" as the "characteristics of a function, process, system or object that is fulfilled when compared to predefined goals or standards". "Quality Assurance" therefore is a "way to warrant that the predefined standards are met". "Quality Improvement" is defined as "a continuous process to review, critique and implement changes".
To achieve Quality Assurance and Quality Improvement, we believe that student involvement is essential in all aspects and at all levels. This fact is also stressed by the WFME, which encourages active involvement of students in the Global Standards.

The IFMSA considers global standards as a starting point on the way to bring quality in medical education to the same level all over the world. This was also the aim of the World Federation for Medical Education, the international representation of all medical teachers and medical teaching institutions, in developing these standards. The IFMSA emphasizes the need for global standards because of the fact that more than 1600 faculties all over the world are teaching basic medical education with only a few of these having established systems of Quality Assurance and Quality Improvement. We want to stress that certain aspects of medical education are universal, regardless of the medical school where one is educated. Furthermore, it should always be remembered that quality improvement of education is quality improvement of health care.

Therefore we welcome the initiative of the WFME. With the Students' specifications, IFMSA wants to add important aspects from the students' point of view. Since the 56th IFMSA August Meeting held in 2007 in Canterbury, UK, the Standing Committee on Medical Education of the IFMSA has worked on developing these specifications. At the 57th March Meeting held 2008 in Monterrey, Mexico, the "Basic Medical Education WFME Global Standards for Quality Improvement – Students' specifications" have been adopted by the IFMSA General Assembly.

IFMSA Policy Statement on the Timescale for Implementation of the Bologna Process

The IFMSA, as the organisation that represents medical students worldwide, acknowledges that the Bologna Process is a change that will have a significant impact on medical education. The nature of that impact highly depends on its proper implementation. As a result of that, the IFMSA has been working extensively to formulate recommendations on the Bologna action lines over the last five years and will continue in their work in the future.

Therefore, and as a response to concerns raised by medical students across Europe the IFMSA:

1) Acknowledges that the current timescale for the implementation of the Bologna Process in full across the whole of Europe is unlikely to be met with regards to medical education unless necessary and immediate action is taken where relevant.

2) Calls for the relevant governmental bodies in those countries who are yet to formally put in place necessary protocols for implementation (including those countries who are yet to decide whether to implement Bologna in medical education or not) to take immediate action by consulting with the relevant stakeholders to ensure that when the process is implemented it is done so with as little disruption as possible.
3) Urges those governments where the process is to be implemented to ensure that quality of education is assured and believes that implementation should be well planned and in the best interests of medical education.

Clearly this policy statement is only relevant to those nations who have signed the Bologna Declaration.

Undergraduate Mobility in Medical Education in the European Higher Education Area

Note: You can find the complete document with all the references in the SCOME-wikipedia!

Curricula

- *European Core Curriculum – learning outcomes.* We strongly recommend the implementation of a European core curriculum as suggested by EMSA and IFMSA. As stated in the "European Core Curriculum for Medicine – the students' perspective", we support the idea of defining expected learning outcomes of medical education. This can assist mobility even in the core parts of the curriculum by facilitating comparison of curricula and courses.
Aiming for common core goals can enhance mobility by increasing trust between institutions. Hence, every medical school should specify learning outcomes for their curriculum and make these publicly available. We support the faculties' autonomy and their right to define their curriculum within the framework of the European Core Curriculum and following the Students' Specifications on the WFME Global Standards for Basic Medical Education.
- *Recognition.* When comparing and recognizing courses provided at other universities, achieved learning outcomes should be considered more important than teaching and assessment methods, duration of study or title of course. The learning outcomes a student expects to achieve at the host university should be agreed upon prior to the commencement of the programme. Meeting these learning outcomes must lead to automatic recognition at the student's home university. Furthermore, upon finishing the course, students should be provided with supplementary documents stating which of the previously specified learning outcomes have been achieved. Additional achieved learning outcomes should also be stated. We encourage higher education institutions to use the tools for comparability implemented by European institutions, such as the Europass and the European Qualifications Framework.
- *Decision making.* Decisions about recognition should be made in best interest of the student, the student's study progress and the quality of medical education. This decision should be made by the academic staff and students in charge of mobility. The decision should be based upon predefined criteria. The students applying for recognition must have the right to ask for the decision to be revised.
- *Assessment.* Assessment should preferably be done at the host university following the course taken, but students should be offered the flexibility to be assessed at the home university if this serves to avoid adverse effects on the student's academic, professional and personal progress.
- *Flexibility.* Universities should allow and facilitate medical students to combine courses from different periods of studies to facilitate the achievement of the expected learning outcomes. They should also provide incoming students with guidance during their study period in order to help them in meeting all expected learning outcomes. If the host faculty cannot enable the students to fulfil the expected requirements, the home institution should actively support the students in achieving the required learning outcomes.

- *Electives.* We strongly recommend every medical faculty to offer elective parts within their medical curriculum. Students should be allowed to take electives at their own university or elsewhere.
- *Research.* As stated in the "European Core Curriculum for Medicine – the students' perspective", students should be involved in research work. It should be possible for students to do this at another university.

Language

High standards of language skills are particularly important in medicine, since medical education includes contact with patients. Therefore, we believe that:
- *Language courses.* Universities should offer language courses in order to enhance mobility.
- *Requirements.* Appropriate minimum language standards, as required by a host faculty, should be communicated by the home faculty. This should occur prior to the commencement of a period of study abroad.
- *Appropriate language skills.* If a faculty allows part of the curriculum to be taught in a non-native language, it should be ensured that students enrolled in that course have appropriate skills in the language of delivery.
- *Native language courses.* We encourage host faculties to offer native language courses, including medical vocabulary, in order to help the incoming students to gain the appropriate skills required to communicate with patients and health care professionals.
- *Online medical dictionaries.* We encourage the further development of online multilingual dictionaries of medical terms.

The role of universities and faculties

- *Information to students.* Faculties should recognize that student mobility is a part of the academic development and therefore encourage students to take parts of their studies at another institution. In order to give students the best opportunity to participate in mobility programmes, faculties should communicate mobility opportunities to the students early in their studies. Faculties should also provide students with up-to-date information about their partner universities, as advised in the ECTS guide, and about the possibilities of recognition of courses. We strongly encourage all faculties to participate in initiatives aimed at creating an accurate and up-to-date database containing relevant information about medical schools, e.g. the Avicenna Directories.
- *Cooperation between faculties.* Universities are encouraged to seek new possibilities for students to study at another university. Faculties are also encouraged to establish different types of cooperation, including bilateral agreements and networks with other
universities. This will help improve mobility and foster the mutual trust between the faculties within the network. These networks should always be open to the incorporation of new faculties. Mobility of students enrolled in a faculty member of a network should also be possible to universities that are not members of that network. Networks should not be seen as a tool to reduce the diversity and autonomy of faculties or to standardize educational systems and curricula, but as a tool for all member faculties within them to learn from each other. We also encourage the creation of
mobility programmes, such as ERASMUS, specifically applied to medicine and to the different kinds of medical mobility. Faculties are encouraged to improve mobility of academic staff to enhance academic cooperation and gain trust.
- *Protocol of procedures.* Faculties should provide a protocol of procedures helping students to arrange a period of studies at another university. To enhance the contact between students and faculties in already existing networks, we also recommend the
creation of web based forums to facilitate the communication between everyone participating in mobility.
- *Quality and quantity.* We support the idea of unilateral mobility programmes, as long as the increase in numbers of students does not negatively affect the quality of education. Faculties

must always ensure that an increase in quantity would not compromise the quality of medical education.
- *Application.* All application and selection processes should be transparent, fair, easily accessible and comprehensible. Application requirements should be adequate for the type of exchange, especially considering whether the student will have contact with patients or not.
- *International office.* To maximize the academic achievements of the exchanges there should be a centralized international office at every faculty. This office should be responsible for informing and assisting incoming and outgoing students as well as faculty members in all practical and academic aspects. Local students could be mentors for incoming and outgoing students, thereby increasing the motivation of local students to become more mobile themselves. Furthermore, this office should encourage and assist members of staff to establish exchange programmes and cooperation between faculties. The staff of the international office should include students. Students should take an active part within the selection and application process, including participating in the definition of the selection criteria.

ECTS
- *Correct implementation of ECTS.* We strongly encourage the correct implementation of ECTS credits across Europe with consistency between faculties according to the ECTS guide of the European Commission. It is recommended that all faculties use the ECTS framework and its nomenclature. This should be done for both local and incoming students.
- *Transparency of ECTS Framework.* Transparency of the framework and its validation are imperative. European or national bodies should supply training and recommendations to help faculties implement the ECTS framework properly according to the ECTS guide. This could be done through, among others, providing more information and guidelines about the importance, meaning and implementation of the ECTS framework. The guidelines should be kept up-to-date and accessible in relevant languages for both students and faculties. Student organizations can assist in the wide distribution of information concerning the ECTS framework amongst their peers.
- *Supplementary documents*. Faculties are strongly encouraged to provide supplementary documents indicating the student's achieved learning outcomes and study progress at any time during studies to increase mobility and transparency. Including learning outcomes in supplementary documents is a prerequisite for using the ECTS point system as a functional measure. In this case, the ECTS credit would provide quantitative information and the supplementary documents would provide information about the content of the studies. This would increase comparability and thus promote mobility. The TUNING project has developed a format for this purpose.
- *Students' involvement.* Students should be involved in the implementation of the ECTS framework, including the definition of ECTS points and in the evaluation of students' workload estimation.

Quality assurance
In order to improve quality of mobility, internal quality assurance systems should be implemented in each faculty. The quality of these systems should be assured by national accreditation bodies. The national accreditation bodies should be accredited by a common European accreditation institution. It is essential for the improvement of quality in mobility that this quality assurance scheme is regularly evaluated as well as transparent and accessible.
We recommend the use of:
- *Standards and Guidelines.* We strongly recommend the implementation of international standards and guidelines within the European Higher Education Area, e.g. WFME Global Standards for Basic Medical Education and the Students' Specifications, to these, to assure transparency and to have a common reference level.

- *Databases.* We encourage the establishment of a database providing accurate information about medical schools. In order to achieve this, we welcome the idea of the elaboration of the Avicenna Directories.
- *Quality assured assessment methods.* To increase recognition between medical schools and therefore increase mobility, faculties' chosen assessment methods should be quality assured, transparent and based on best evidence.
- *Internal quality assurance.* The university and/or faculty should be responsible for the evaluation of both its mobility programmes and its international office or other relevant departments responsible for mobility. The evaluation should involve both students (local, incoming and outgoing) and academic staff. Incoming students should give feedback about the relevant medical education programme at the host university and, by this, promote innovation and improvement. Outgoing students should be encouraged to give recommendations to improve medical education at their home faculty. Evaluation, recommendation and feedback results should be publicly available and accessible.
- *External quality assurance.* Accreditation systems ensure the quality of the programme, and thus could improve the recognition of achieved learning outcomes by the home faculty. National or regional accreditation systems for medicine should cooperate on a European level.

IFMSA Statement on Patient Safety in undergraduate curricula.

We, the students of the International Federation of Medical Students' Associations (IFMSA),

Acknowledge that the lack of appropriately designed and implemented systems to ensure patient safety represents a severe threat to patients all over the world.
Believe that poor undergraduate training in patient safety negatively affects the competences of medical students and future physicians in delivering safe patient healthcare.
Welcome the ongoing work and dedication to this topic of the World Health Organization (WHO) and the World Alliance for Patient Safety (WAPS), and the creation of their recommendations for the WHO Patient Safety Curriculum Guide for medical schools.

Because of the aforementioned, we,

- encourage WHO and WAPS to continue with their work on patient safety.
- recommend medical schools take active part in ensuring patient safety, by implementing patient safety topics into the curriculum because of the following:
  1- It fits under the quality control of the physicians' job
  2- It concerns every specialty and every physician, and therefore it should be included in core curricula
  3- It increases patient care and health at minimal costs when compared to the financial cost of the results of poor patient safety.
- encourage medical schools to promote a culture of safety in the healthcare setting as a necessary requisite for the effective implementation of patient safety in curricula.
- recommend that national and regional bodies follow the recommendations of the WHO and to actively work on them, creating legislation to ensure the safety of patients and that patient safety issues are covered in medical school curricula.

International partners

World Federation for Medical Education (WFME)

www.wfme.org

Structure and Objectives

The World Federation for Medical Education was founded in 1972 and has its office in Copenhagen, Denmark . The federation serves today with the purpose of being an umbrella organisation for its regional associations for medical education, following the regional structure of the World Health Organisation. Four other institutions related to the field of medical education are also members of the WFME Executive Council.

The regional associations are:

- **AMSA** - Association for Medical Schools in Africa
- **PAFAMS** - Pan-American Federation of Associations of Medical Schools
- **AMEEMR** - Association for Medical Education in the Eastern Mediterranean Region
- **AMEE** - Association for Medical Education in Europe
- **SEARAME** - South-East Asian Regional Association for Medical Education
- **AMEWPR** - Association for Medical Education in the Western Pacific Region

The four institutions are:

- **WHO** – World Health Organisation
- **WMA** – World Medical Association
- **ECFMG** – Educational Commission for Foreign Medical Graduates
- **IFMSA** – International Federation of Medical Students' Associations

WFME aims at enhancing the quality of medical education world-wide, taking initiatives with respect to new methods, new tools, and new management. It covers all phases of medical education (graduate education, specialist training and continuing medical education).
The general objective of WFME is "*to strive for the highest scientific and ethical standards in medical education, taking initiatives with respect to new methods, new tools, and management of medical education*".

Projects and Activities

WFME is undertaking a number of different activities, e.g. the Guidelines for the use of Information and Communication Technology in Medical Education.

The process of implementing the WFME Programme on Global Standards in Medical Education, as documented in the *Trilogy of Global Standards for Quality Improvement of Medical Education*, is progressing:

- Pilot Studies have been expanded from the Standards in Basic Medical Education to the Standards in Postgraduate Medical Education and Continuing Professional Development (CPD) of Medical Doctors.
- The number of medical schools and other educational institutions which are using the WFME Standards in programme development is rapidly increasing. Also, the number of authorities or agencies, which are incorporating the Standards in national and regional standard setting and systems of accreditation, is growing.

- A manual for WFME advisors has now been developed. An advisor corps, representing all regions, with the purpose of assisting medical schools in utilising the WFME Global Standards, is now ready.
- Based on the results of a task force meeting on accreditation, WHO and WFME have now defined guidelines for accreditation of basic medical education institutions and programmes.
- A new programme for Promotion of Accreditation of Basic Medical Education has been developed. The idea is to offer assistance to institutions and agencies regarding the various steps of an accreditation procedure. Interested institutions, organisations and agencies are invited to take advantage of this programme. Contact should be made to the WFME Office.

The WHO – WFME Strategic Partnership to improve medical education is now working in close collaboration with the WHO Regional Offices. Concrete examples are a process of supporting medical education reforms in the CIS countries, development of accreditation systems in the Eastern Mediterranean region and capacity building of health manpower in Sub-Saharan Africa and quality development of medical education in e.g. Latin America, South East Asia and the Western Pacific Region.

A statement on the Bologna Process and its relationship to medical education has been developed jointly by WFME and AMEE. The organisations endorse the purpose of the Bologna Declaration and support that medical education as a part of higher education should be fully involved in the Bologna Process. However, the specificity of medical curricula and the current situation of European medical schools must be considered, and it is the opinion that the two-cycle division in a Bachelor and a Master degree would invalidate endeavours to integrate basic and clinical sciences in the medical curriculum.

A Task Force under the EU project MEDINE, organised by WFME and Association of Medical Schools in Europe (AMSE), is working on a proposal for definition of European Standards in medical education. The Task Force had its first meeting in January 2006 and has just conducted a survey on recognition/accreditation systems in medical education in Europe. The European Specifications of the Global Standards Programme are resulting from this task force's work.

WFME is now working with WHO about changes of the WHO Directory of Medical Schools to a comprehensive Database on Health Professions Education Institutions, comprising not only medical schools, but also educational institutions for dentistry, public health, physiotherapy, pharmacy, midwifery and nursing. It is also part of the new development to include qualitative information about institutions and programmes such as accreditation issues. The database is called AVICENNA directories and is administered by the University of Copenhagen (Denmark).

For more information about these activities, please visit the WFME website
http://www.wfme.org

International Standards in Basic Medical Education
This project has defined standards to outline minimum requirements of medical education institutions worldwide.
The project has three main intentions:

1. to stimulate medical schools to formulate their own plans for change and for quality improvement in accordance with international recommendations
2. to establish a system of national and/or international assessment and accreditation of medical schools to assure minimum quality standards for medical school programmes

3. to safeguard practice in medicine and medical manpower utilisation, and its increasing internationalisation, by well-defined international standards of medical education

The standards are divided into two levels:
1. Basic standards - should be fulfilled by all institutions involved in ME
2. Standards for quality development - serves as an incentive for development and a leverage for improvement

Standards are defined in these two levels for each of the following areas of work and administration/planning of the medical school:
1. Mission and objectives
2. Educational program and principles
3. Assessment of educational outcomes
4. Students
5. Academic staff/faculty
6. Educational resources
7. Monitoring and evaluation of programs and courses
8. Governance and administration
9. Continuous renewal of the medical school

You can find the WFME Global Standards for Quality Improvement in Basic Medical Education at http://www.wfme.org in different translations. The English version can also be accessed at the SCOME-wikipedia.

3rd World Conference on Medical Education
In March 2003 WFME has organised the 3rd World Conference on Medical Education in Copenhagen, Denmark. The theme for the event has been "Global Standards in Medical Education for Better Health Care".
The aim was to stimulate a "*debate among decision-makers in medical education and health care about the complex question of introducing generally accepted global standards in medical education in order to promote the quality of health care delivery systems*".

WHO/WFME strategic partnership to improve medical education
Following the WFME World Conference in Medical Education: *Global Standards in Medical Education for Better Health Care*, Copenhagen, Denmark, March 2003, the World Health Organization (WHO) and WFME decided to establish a Joint Policy on Promotion of Health Systems Performance Through Improvement of Health Professions Education. As a result, a WHO/WFME strategic partnership to improve medical education was formulated in January 2004. The partnership agreement is available on www.who.int/hrh/links/partnership/en/print.html or www.wfme.org.

A brochure outlining the partnership can be ordered from the WFME office. An action plan was finalized in June 2004 for the WHO-WFME joint policy covering the period 2004-2006.

Publications
Since 1997 the WFME is affiliated with the journal "Medical Education". "Medical Education" is one of the leading publications in the field of medical education.

IFMSA in WFME

Representation
At the WFME Executive Council (EC) meeting in Vienna in 1997 it was decided that IFMSA should be represented in the EC by its president for a period of 2-3 years for

continuity's sake. Given the fact that turnover rate is rather high in IFMSA it was agreed that the IFMSA-representative could also be a past president. Alternatively, the current or immediate past SCOME Director is eligible to take over this function.
The EC meeting in Copenhagen in September 2000 was the first time that IFMSA attended WFME activity.
The present IFMSA representative on the WFME Executive Council is the Liaison Officer on Medical Education Issues, Robbert Duvivier (IFMSA – The Netherlands).

Collaboration
The collaboration between IFMSA and WFME in the year 2001 was fruitful, first and foremost in relation to the IFMSA August Meeting, where WFME contributed with valuable expertise to the roundtable discussions "Implementing International Standards in Basic Medical Education" and "Impact of Technology on Health Education". Concrete outcomes of these discussions are two policy statements that will serve as a basis for IFMSA activities in these fields in the coming years. The workshop "Future of Medical Education" was organised under the patronage of WFME.
The "3rd World Conference on Medical Education" was a concrete example of the valuable collaboration between the two organisations. IFMSA has been invited as one of the partners in preparing workshops, presentations and speeches for this important event, and thus had the opportunity of participating in shaping policies and opinions in the very field that dominates our life as medical students and future doctors - namely our everyday medical education.
In 2005 a Bologna follow-up workshop has been organised in Copenhagen, Denmark, where two members of the WFME EC gave presentations on the topic of quality assurance and quality improvement in the scope of the Bologna process in medical education.
In its two annual General Assembly meetings and numerous international training workshops IFMSA offers a good venue for the members of WFME to share their ideas with medical students, and to get direct input from the next generation of physicians.
IFMSA aims at working actively together with the regional associations for medical education in the future development of medical training, and the organisation wishes to be involved in the work of the regional associations to the largest extend possible.

Association for Medical Education in Europe (AMEE)

www.amee.org

The Association for Medical Education in Europe (AMEE) was founded in 1972 in Copenhagen (Denmark) to foster communication among medical educators and to help promote national associations for medical education throughout Europe. It is the European regional association of the World Federation for Medical Education (WFME). Several European national medical education associations are corporate members of AMEE. Over the past decade AMEE has developed steadily both in size and in the sphere of its activities and is now a worldwide association with members and contacts in over 90 countries.

AMEE helps teachers, doctors, researchers, administrators, curriculum developers, assessors and students keep up to date with developments in the rapidly changing world of medical and healthcare professions education. AMEE's activities include the annual conferences, Publications including Medical Teacher and AMEE education guides, courses such as ESME, FAME and RESME, projects including BEME and MedEdCentral, and Special Interest Groups.

Each year since 1973 AMEE has organised an annual conference in a European city. AMEE conferences now regularly attract over 1800 participants from around the world and the event has become the major gathering for all interested in medical and healthcare professions education to get together to network, share ideas and hear the best of what's happening in medical education throughout the world. It's not necessary to be an AMEE member to attend the conferences, although members do receive a discount on registration.

AMEE's Education Guides are designed as practical, how-to-do-it guides on important topics such as problem-based learning, outcome-based education, portfolios in student assessment, and a wide range of other topics. In the BEME Guide Series, the reports of the Best Evidence Medical Education systematic reviews are published. AMEE members receive a discount on purchases of Guides.

The ESME Programme has been designed in the context that all doctors in any branch of medicine or field of practice are likely to have some teaching responsibilities for undergraduates, postgraduates, peers, other healthcare workers or patients. ESME provides an entry-level teaching qualification for teachers who are engaging in medical education for the first time, or who have been given some new responsibilities or assignment relating to teaching. ESME courses are offered at major medical education meetings, including AMEE.

FAME, a basic level course in assessment, is being offered for the first time at AMEE 2007. It is designed for those with responsibility for assessing undergraduate medical students, graduate trainees and practising doctors. The course will also include selected aspects of program evaluation.

RESME is a course that introduces participants to some essential principles and methods of research in medical education, including phrasing a research question, methodology of research and research designs.

AMEE is a founder member of the BEME Collaboration (www.bemecollaboration.org) which aims to promote Best Evidence Medical Education through the dissemination of information that assists evidence-based decisions, the publication of high-quality systematic reviews in medical education and the creation of a culture of the use of evidence to inform practice. Five BEME systematic reviews have now been published and nine more are in progress.

MedEdCentral (www.MedEdCentral.org) is an online medical education resource currently under development, that includes medical education terminology, publications, institutions, medical schools, associations and individuals. The database may be accessed by anyone. Built around the Wiki principle, registered users may contribute to many areas of the site, and may suggest additions to others.

AMEE's newest project includes setting up Special Interest Groups (SIGs) on key topics in medical education, to provide a forum for the exchange of ideas and information. It is intended that anyone may register to join a SIG group and participate in online discussion or face-to-face in groups at AMEE conferences.

Members of AMEE are also allowed to participate in the annual conference at a reduced rate. The annual membership fee for student members of AMEE is £ 39 with access to the archives of the journal "Medical Teacher".

Other organizations

Association for the Study of Medical Education (ASME)

www.asme.org.uk/

The Association for the Study of Medical Education (ASME) seeks to improve the quality of medical education by bringing together individuals and organisations with interest and responsibilities in medical and healthcare education.

The values of ASME are:
- Education and learning are central to the delivery of high quality healthcare
- Education must be an important component in the strategies of Governmental and other healthcare organisations
- Good healthcare educators are central in planning, delivering and evaluating high quality healthcare
- Individual members of ASME should be supported and developed
- High quality research is necessary for the development of healthcare education
- Vision, innovation and leadership in healthcare education are to be fostered

ASME seeks to
- Promote high quality research in to medical education
- Provide opportunities for developing medical education
- Disseminate good evidence based educational practice
- Inform and advise Governmental and other organisations on medical education matters
- Develop relationships with other organisations and groupings in healthcare education

As such as AMEE, ASME also offers special membership fees to students. The annual fee for students is £ 30.
Membership includes free personal copies of "Medical Education", "Clinical Teacher" and the "ASME Bulletin".

Comité permanent des Médecins Européens (CPME)

CPME is the international umbrella organization of the National Medical Associations in the Countries of the European Union/European Economic Area seated in Brussels (Belgium). EMSA is active an active partner and represents the interests of students. To

focus the activities there is no liaison officer from IFMSA, but the liaison officer of EMSA is in close contact to IFMSA and reports about CPME activities.

European Medical Students´ Association (EMSA)

EMSA is an organization that aims to integrate all medical students in geographical Europe through activities organised for and by medical students. EMSA has a committee on Medical Education, which is co-operating with us. The most important joint project in the last years have been the follow-up conferences on the Bologna Process in Medicine.

European Students Conference (ESC)

ESC is a scientific conference for medical students and young doctors taking place annually in Berlin (Germany). A scientific board selects students, who present their research projects and results in various thematic areas.

Publications and journals in the field of medical education

Beside the publications of each countries association for medical education there is a growing number of international publications on medical education.

The following list shows the most important ones:

- **Academic Medicine (Acad Med)**
  Academic Medicine, a peer-reviewed monthly journal, serves as an international forum for the exchange of ideas and information about policy, issues, and research concerning academic medicine, including strengthening the quality of medical education and training, enhancing the search for biomedical knowledge, advancing research in health services, and integrating education and research into the provision of effective health care. It is the journal of the Association of American Medical Colleges.

- **Advances in health sciences education: theory and practice (Adv Health Sci Educ Theory Pract)**
  Advances in Health Sciences Education is a forum for scholarly and state-of-the art research into all aspects of health sciences education. It will publish empirical studies as well as discussions of theoretical issues and practical implications. The primary focus of the Journal is linking theory to practice, thus priority will be given to papers that have a sound theoretical basis and strong methodology.

- **Clinical Teacher**
  The Clinical Teacher has been designed with the active, practising clinician in mind. It aims to provide a digest of current research, practice and thinking in medical education presented in a readable, stimulating and practical style. The journal includes sections for reviews of the literature relating to clinical teaching bringing authoritative views on the latest thinking about modern teaching. There are also sections on specific teaching approaches, a digest of the latest research published in *Medical Education* and other teaching journals, reports of initiatives and advances in thinking and practical teaching from around the world, and expert community and discussion on challenging and controversial issues in today's clinical education. ASME

members receive Clinical Teacher as part of their membership subscription.

- **Education for health (Educ Health)**
- **Journal of Nursing Education (J Nurs Educ)**
- **Journal of Health Education / Association for the Advancement of Health Education (J Health Educ)**
- **Medical Education (Med Educ)**

Medical Education seeks to be the pre-eminent journal in the field of education for health care professionals, and publishes material of the highest quality, reflecting world wide or provocative issues and perspectives. It is published on behalf of the Association for the Study of Medical Education (ASME).
The journal welcomes high quality papers on all aspects of medical education including;

- undergraduate education
- postgraduate training
- continuing professional
- development interprofessional education

ASME members receive Medical Education as part of their membership subscription.

- **Medical Teacher (Med Teach)**

Medical Teacher, the journal of the AMEE, is a peer-reviewed journal, listed in Medline and published ten times a year. It publishes reports of innovation and research in medical education, case studies, commentaries and practical guidelines and BEME Guides, as well as a range of popular features to help readers keep up to date with the rapidly developing area of medical and healthcare professions education. AMEE members receive Medical Teacher as part of their membership subscription.

- **Teaching and learning in medicine (Teach Learn Med)**

The most important ones out of these are "Medical Teacher", "Medical Education", "Academic Medicine" and "Advances in Health Science Education"
The one to read the easiest probably is "The Clinical Teacher", which summarizes latest developments in the field of medical education.

Some national journals are

- British Journal of Medical Education (Br J Med Educ) -> United Kingdom
- GMS Zeitschrift für Medizinische Ausbildung -> Germany, Switzerland, Austria
- Pédagogie Médicale -> France and French-speaking countries
- Tijdschrift voor Medisch Onderwijs -> The Netherlands

Professionalizing Medical Education

In recent year there has been a trend towards professionalizing medical education. Master degrees in Medical Education (MME) have been established and research of best practice in medical education has been introduced.

By the end of 2005 there are 21 programmes offering a master's degree in medical or health sciences education in the English-speaking world (including the programme in Maastricht, Netherlands).

These programmes differ a lot in content and length so that it until now still remains difficult to compare MME-graduates from different programmes. The "Degrees of Difference Report" by the Association for the Study of Medical Education (ASME), which was published in 2006 gives a more detailed insight of the provision of Masters and Doctorate in medical and healthcare education in the UK recording and highlighting the similarities and differences in the offered programmes.
The research has revealed wide variation in the programmes of study, and in particular in the dissertation element of each programme. Copies can be purchased from the ASME office in Edinburgh.

Beside these programmes offered in English there are a couple of other programmes available in other languages, for example the MME programme of the University of Bern (Switzerland) or the German programme MME-D which is a joint project of many German universities and is organised by the University of Heidelberg.

Background information on Medical Education Issues

Bologna process

- ***Introduction***

The purpose of the **Bologna process** is to create the European Higher Education Area (EHEA) by harmonising academic degree standards and quality assurance standards throughout Europe for each faculty and its development. The name is based on the fact that the process was proposed at the University of Bologna with the signing, in 1999, of the Bologna declaration by ministers of education from 29 European countries in the Italian city of Bologna. This was opened up to other countries, and further governmental meetings have been held in Prague (2001), Berlin (2003), Bergen (2005), London (2007); the last meeting took place in Leuven (Belgium) in summer 2009.

The Council of Europe and UNESCO have jointly issued the "Lisbon recognition convention" on recognition of academic qualifications as part of the process, which has been ratified by the majority of the countries party to the Bologna process.

- ***History***

>>Sorbonne 1998

In May 1998 the ministers in charge of higher education of France, Italy, the United Kingdom and Germany signed the so-called Sorbonne Declaration on the "harmonisation of the architecture of the European Higher Education System" at the Sorbonne University in Paris. Other European countries later subscribed to the Declaration.

The Sorbonne Declaration focused on

- a progressive convergence of the overall framework of degrees and cycles in an open European area for higher education
- a common degree level system for undergraduates (Bachelor's degree) and graduates (Master's and doctoral degree)
- enhancing and facilitating student and teacher mobility (students should spend at least one semester abroad); removing obstacles for mobility and improving recognition of degrees and academic qualifications

>>Bologna 1999

In June 1999, 29 European ministers in charge of higher education met in Bologna to lay the basis for establishing a European Higher Education Area (EHEA) by 2010 and promoting the European system of higher education world-wide. In the Bologna Declaration, the ministers affirmed their intention to:

- adopt a system of easily readable and comparable degrees
- adopt a system with two main cycles (undergraduate/graduate)
- establish a system of credits (such as ECTS)
- promote mobility by overcoming obstacles
- promote European co-operation in quality assurance
- promote European dimensions in higher education

>>Prague 2001

Two years after the Bologna Declaration, the ministers in charge of higher education of 33 European signatory countries met in Prague in May 2001 to follow up the Bologna Process and to set directions and priorities for the following years.
In the Prague Communiqué the ministers

- reaffirmed their commitment to the objectives of the Bologna Declaration
- appreciated the active involvement of the European University Association (EUA) and the National Unions of Students in Europe (ESIB)
- took note of the constructive assistance of the European Commission
- made comments on the further process with regard to the different objectives of the Bologna Declaration
- emphasised as important elements of the European Higher Education Area:
    - lifelong learning
    - involvement of students
    - enhancing the attractiveness and competiveness of the European Higher Education Area to other parts of the world (including the aspect of transnational education)

>>Berlin 2003

When ministers met again in Berlin in September 2003, they defined three intermediate priorities for the next two years: quality assurance, the two-cycle degree system and recognition of degrees and periods of studies. In the Berlin Communiqué, specific goals were set for each of these action lines.
- Quality assurance
Ministers stressed the need to develop mutually shared criteria and methodologies and agreed that by 2005 national quality assurance systems should include:

- A definition of the responsibilities of the bodies and institutions involved
- Evaluation of programmes or institutions, including internal assessment, external review, participation of students and the publication of results
- A system of accreditation, certification or comparable procedures, international participation, co-operation and networking

- The two-cycle system
Ministers asked for the development of an overarching framework of qualifications for the European Higher Education Area. Within such frameworks, degrees should have different defined outcomes. First

and second cycle degrees should have different orientations and various profiles in order to accommodate a diversity of individual, academic and labour market needs.

- Recognition of degrees and periods of studies

Ministers underlined the importance of the Lisbon Recognition Convention, which should be ratified by all countries participating in the Bologna Process. Every student graduating as from 2005 should receive the Diploma Supplement automatically and free of charge.

- The third cycle

Ministers also considered it necessary to go beyond the present focus on two main cycles of higher education to include the doctoral level as the third cycle in the Bologna Process and to promote closer links between the European Higher Education Area (EHEA) and the European Research Area (ERA). This added a tenth action line to the Bologna Process:

- Doctoral studies and the synergy between EHEA and ERA.

Ministers charged the Follow-up Group with organising a stocktaking process in time for their summit in 2005 and undertaking to prepare detailed reports on the progress and implementation of the intermediate priorities set for the period.

>>Bergen 2005

In Bergen in May 2005, the Ministers responsible for higher education in the 40 participating countries to the Bologna-process have met for a mid-term review and for setting goals and priorities towards 2010. They confirmed their commitment to coordinating their policies through the Bologna-process to establish the European Higher Education Area (EHEA), and they committed themselves to assisting the new participating countries to implement the goals of the process.

>>London 2007

Montenegro was the 46th country signing the Bologna Declaration. Mobility has been identified to be one of the most relevant topics until the next conference in Belgium in 2009.

>>Leuven/Louvain-la-Neuve 2009

Ten years after the historical Bologna Declaration that structurally reshaped European higher education, another Ministerial Conference was held in Leuven/Louvain-la-Neuve (Belgium).

- ***Structure***

1) Ministerial Conferences

Bologna 1999
Bergen 2001
Berlin 2003
Prague 2005
London 2007
Leuven/Louvain-la-Neuve 2009
Budapest/Vienna 2010 (anniversary meeting)
Bucharest 2012

2) Bologna Follow-up Group

Oversees the process between the ministerial conferences and is composed of:

- representatives of the 46 countries participating in the process of creating the EHEA;
- European Commission as additional full member;
- eight consultative members, namely Council of Europe, UNESCO's European Centre for Higher Education, European University Association, European Association of Institutions in Higher

Education, European Students' Union, European Association for Quality Assurance in Higher Education, Education International Pan-European Structure, and BUSINESSEUROPE.

The Bologna Follow-up Group (BFUG) meets at least once every six month, is chaired by the country holding the Presidency of the European Union and is supported by a Bologna Secretariat, currently provided by the Benelux countries (as host of the next ministerial conference). The host of the next ministerial conference also acts as vice-chair of the BFUG.

3) Working Groups

At its meeting in October 2007, the Bologna Follow-up Group adopted a **work programme** for the time leading to the next ministerial meeting in April 2009 and established working or coordination groups on the following topics: data collection, employability, European higher education in a global setting, lifelong learning, mobility, qualifications frameworks, social dimension, and stocktaking.

4) Bologna Seminars

Valuable input for working groups, Bologna Follow-up Group and ultimately ministerial conferences comes from a number of official Bologna Seminars that are organised on a variety of issues all over Europe. Those Seminars usually serve the dual purpose of policy development and dissemination and are open to a wide range of participants involved in higher education and higher education policy-making.

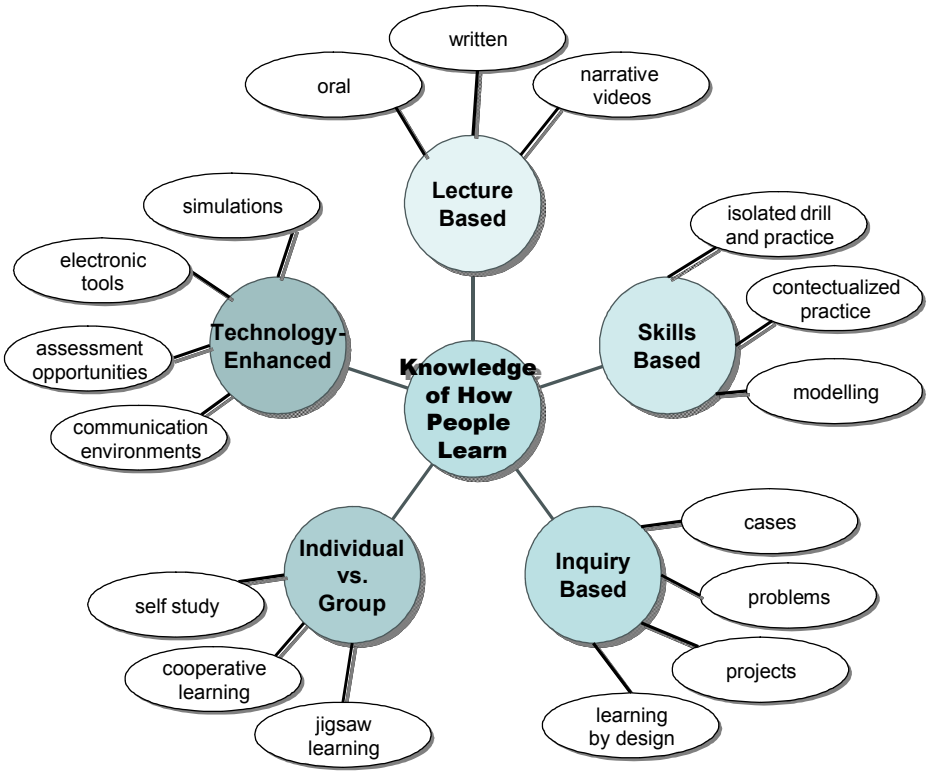
- ***Further information***

Further information regarding the Bologna process can be found at the SCOME-wikipedia (search item “Bologna Process” plus related articles)! You can also access the category “Bologna Process” to easily find all articles related to it.
There you can also find the original versions of the “Lisbon Convention”, the “Sorbonne Declaration”, the “Bologna Declaration”, the “Prague Communiqué”, the “Berlin Communiqué”, the “Bergen Communiqué” and the “London Communiqué”.

You can also find information at the official Bologna Process homepage:
http://www.ond.vlaanderen.be/hogeronderwijs/bologna/

Teaching methods

Both practical and theoretical knowledge is important in the medical education. There are different teaching methods being used for giving the students the best opportunities to learn. This short review shall give you an overview on different methods of the delivery of medical education.
It may not cover all aspects but aims to give you some information on the mostly used teaching methods in medical education.



From: Bransford, J.D., A.L. Brown, and R.R. Cocking: How People learn: Brain, Mind, Experience, and School. Washington, D.C., National Academy Press: 2000.

Lectures

What is it?	• The lectures are lead by (most often) one lecturer • They are hold in lecture rooms for a large group of students • One-way communication from lecturer to students • Little or no student participation (except questions) • Presentation of core material, patients • Used to deliver factual knowledge
Group size:	• Whole academic years.
Advantages:	• Students can gain much information quickly. • Cheaper teaching method for the faculties compared to other teaching methods.
Disadvantages:	• Little or no student participation/feedback. • No registration monitored – skiving students.

	• Visual aids far away – can be difficult to see. • Can be difficulties keeping up with the lecturer during the session. • Easy to fall asleep ☺

How to improve them:	• Clear structure
	• Define goals and objectives
	• Define expected learning outcome
	• Define lesson plan
	• Staff development:
	  ○ Presentation skills
	  ○ Active lecturing

References	• Cantillon P. Teaching large groups. BMJ 2003;326;437 • Bligh DA. What's the use of lectures? San Francisco: Jossey-Bass, 2000. • Brown G, Manogue M. AMEE medical education guide No 22: Refreshing lecturing: a guide for lecturers. Medical Teacher 2001; 23:231-44.

Problem-based learning

What is it?	• Standardized small group sessions with discussion of pre-made problems/cases. • During regular meeting the participants discuss the keywords and ideas they extract from the PBL-case. • For each meeting they decide on a learning objective. • Between meetings, he students do individual reading/learning on the objective. • In the next meeting, the learning issues are beeing discussed again. • They can do presentations or discussions of their home-work. • The meetings are facilitated by a facilitator.
Group size:	• Small, can be 6-8 or 12-15 (scribe, tutor, chair, group members)
Advantages:	• More understanding on issues • Problem-based – relevant to future career • Learn more – not easy to forget • Two-way communication • Students interests considered more
Disadvantages:	• Good facilitators are needed to get good PBL • Facilitator may not be student-based
References:	• Wood DF. Problem based learning. BMJ 2003;326;328-330 • Davis MH, Harden RM. AMEE medical education guide number 15: problem-based learning: a practical guide. Med Teacher 1999;21:130-40. • Norman GR, Schmidt HG. Effectiveness of problem-based learning curricula: theory, practice and paper darts. Med Educ 2000;34:721-8. • Albanese M. Problem based learning: why curricula are likely to show little effect on knowledge and clinical skills. Med Educ 2000; 34:729-38. • http://edweb.sdsu.edu/clrit/PBL_WebQuest.html

	• http://www.unimaas.nl/pbl/

Clinical rotations (Practical teaching)

What is it?	• Visiting hospital departments/wards • Students get the opportunity to work with patients, equipment and cases. • Supervised by a physician • Rotating between departments/ward – related to teaching topic
Group size:	• Typically 6-8
Advantages:	• Direct patient contact • Physical examination training • Real-life medicine
Disadvantages:	• Group sizes often too large
Resources:	• Spencer J. Learning and teaching in the clinical environment. BMJ 2003; 326;591-594. • Cox K. Planning bedside teaching. (Parts 1 to 8.) Med J Australia 1993;158:280-2, 355-7, 417-8, 493-5, 571-2, 607-8, 789-90, and 159:64-5. • Parsell G, Bligh J. Recent perspectives on clinical teaching. Med Educ 2001;35:409-14. • Ramani S. Twelve tips to improve bedside teaching. Med Teach 25(2), 2003;112-115. • http://www.henryfordhealth.org/1497.cfm

Clerkships (Practical teaching)

What is it?	• Visiting hospitals or GPs for a longer period. • Can be in both rural or urban areas • Students follow the physicians in their work and get close contact with patients • Physicians supervise the students
Group size:	• 1 student
Advantages:	• Long period – continuous work • Students have more responsibility and opportunities • Often good supervising
Disadvantages:	• Too many students, too few patients • Patients prefer qualified doctors
References	• See above at clinical rotations

Small-group teaching

What is it?	• Small student groups with one facilitator
Group size:	• Various
Advantages:	• Practise on everyday situations • Very relevant to career
Disadvantages:	• Sometimes not taken seriously
References	• Jaques D. Teaching small groups. BMJ 2003; 326;492-494 • Habeshaw T, Habeshaw S, Gibbs G. 53 interesting things to do in

	your seminars and tutorials. Bristol: Technical and Educational Services, 1992. • Tiberius R. Small group teaching: a trouble-shooting guide. London: Kogan Page, 1999. • Orlander JD. Twelve tips for use of a white board in clinical teaching. Med Teach 2007; 29; 89-92

Courses

What is it?	• Practical courses in lab. techniques, histology, anatomy demonstrations, biochemistry etc. • Teaching led by a tutor • Tutor assistants may be present
Group size:	• Various
Advantages:	• Assistants present to help students • Hands-on work/experiences • Useful for future careers
Disadvantages:	• Sometimes a lack of time

Self-directed learning (SDL)

What is it?	• Students are given a topic/questions to complete • The work is done individually, and does not include teaching • Completed for own notes • Not checked or assisted by any tutor
Group size:	• 1 student
Advantages:	• Independent learning
Disadvantages:	• No feedback • No supervising of progress

Written essays (SDL)

- includes research work and special study modules (SSM)

What is it?	• Extended written essay • Completed over a number of weeks • Individual work or completed in small groups • Tutor supervises the work and the progress
Group size:	• Mostly 1-2 students
Advantages:	• Pursue own interests • Academic writing • Chance to produce scientific article
Disadvantages:	• Takes a lot of time – some think this time could be used more valuably

Electives (SDL)

What is it?	• Work in a hospital, supervised by physicians • Teaching of clinical and practical skills • Many students travel abroad to do this work
Group size:	• 1 student
Advantages:	• Students choose their own elective

	• Time off study – a kind of break • See medicine in foreign countries – broader perspective on medicine
Disadvantages:	• The cost – expensive to fundraise your travel • A long preparation time

Principles of assessment

Assessment *n.*
The process of documenting, usually in measurable terms, knowledge, skills, attitudes and beliefs.

Note: These paragraphs will – hopefully – provide you with real precious knowledge about the statistical backgrounds of testing theory… at least if you have an interest in statistics and a general understanding about it. It reads complicated and if you are not really really interested in the subject, it might be better for you to skip it. However, having read it and remembering the most important concepts might help you when discussing the quality of assessment in your faculty a lot.
Most of the references are available online at the website of the National Council on Measurement in Education (NCME) at http://www.ncme.org/pubs/items.cfm.

Formative and summative assessment

Formative assessment is carried out in order to intervene with intention to improve future performance or learning habits of students. In contrast to this, summative assessment is carried out in order to make decisions as good/bad, pass/fail, ready to move forward/repeat a programme.

Bob Stake explained the difference between summative and formative: "Formative assessment is usually contrasted with summative assessment in the following way: 'When the cook tastes the soup, that's formative evaluation. When the guests taste the soup, that's summative evaluation'"

In general, students' educational achievements are assessed for many purposes:
- To assure minimal predetermined qualifications
- To identify students who have achieved a level required for promotion to the next level or who need to repeat the programme
- To select the best students for a given programme
- To allow students monitor their own learning
- To provide information regarding student level of achievement
- To generate performance profiles of students' strengths and weaknesses

Traditionally, the first three of these are associated with summative assessment and the last three with formative assessment.

For selection and promotion purposes, summative assessment is the preferred approach. For feedback purpose to students, teachers or the school, a formative assessment is the appropriate method.

However, a summative assessment system can contain a formative component by providing feedback to students on strengths and weaknesses as well.

Forms of assessment most suitable for formative methods are portfolios, objective structured clinical examinations (OSCE), modified essay questions (MEQs) and multiple choice questions (MCQs).

Measurement of assessment

Classical test theory

Classical test theory predicts outcomes of testing such as the difficulty of items or the ability of test-takersThe aim of classical test theory is to understand and improve the reliability of psychological tests.

Classical test theory may be regarded as roughly synonymous with “true score theory”. The term “classical” refers not only to the chronology of these models but also contrasts with the more recent psychometric theories, generally referred to collectively as item response theory

True and error scores

Classical test theory is based on the decomposition of observed scores (which are ordinal, but analyzed as interval) into true and error scores. The theory views the observed score *x* of person *i*, denoted as x_i, as a realization of a random variable *X*. The person is characterized by a probability distribution over the possible realizations of this random variable. This distribution is called a "propensity distribution". The true score of person *i*, t_i, is axiomatically defined as the expectation of this propensity distribution. This definition formally stated as $\varepsilon(X_i) = t_i$

Secondly, the so-called error score for person *i*, E_i, is defined as the difference between *i*'s observed score and his true score: $E_i = X_i - t_i$

Note that X_i and E_i are random variables, but t_i is a constant. Also note that it directly follows from these definitions that the error score has expectation zero:

$$\varepsilon(E_i) = \varepsilon(X_i - t_i) = \varepsilon(X_i) - \varepsilon(t_i) = t_i - t_i = 0$$

Relation to population

The above equations represent the assumptions that classical test theory makes at the level of the individual person. However, the theory is never used to analyze individual test scores; rather, the focus of the theory is on properties of test scores relative to populations of persons. Hence, the next step is to introduce a population-sampling scheme into the structure of classical test theory. When we assume that people are randomly sampled from a population, the true score becomes a random variable too, so that we get the equation: $X = T + e$

Classical test theory is concerned with the relations between the three variables *X*, *T*, and *E* in the population. These relations are used to say something about the quality of test scores. In this regard, the most important concept is that of reliability. The reliability of the observed test scores *X*, which is denoted as ρ^2_{XT}, is defined as the ratio of true score variance σ^2_T to the observed score variance σ^2_X :

$$\rho^2_{XT} = \frac{\sigma^2_T}{\sigma^2_X}$$

Because the variance of the observed scores can be shown to equal the sum of the variance of true scores and the variance of error scores, this is equivalent to

$$\rho^2_{XT} = \frac{\sigma^2_T}{\sigma^2_X} = \frac{\sigma^2_T}{\sigma^2_T + \sigma^2_E}$$

This equation, which formulates a signal-to-noise ratio, has intuitive appeal: The reliability of test scores becomes higher as the proportion of error variance in the test scores becomes lower and vice

versa. The reliability is equal to the proportion of the variance in the test scores that we could explain if we knew the true scores. The square root of the reliability is the correlation between true and observed scores.

Reliability
Note that reliability is not, as is often suggested in textbooks, a fixed property of tests, but a property of test scores that is relative to a particular population, and computed for this sample. This is because test scores will not be equally reliable in every population or even every sample.

For instance, as is the case for any correlation, the reliability of test scores will be lowered by restriction of range. Thus, IQ-test scores that are highly reliable in the general population will be less reliable in a population of college students and even less reliable in a sample of sophomores. Also note that test scores are perfectly unreliable for any given individual *i*, because, as has been noted above, the true score is a constant at the level of the individual, which implies it has zero variance, so that the ratio of true score variance to observed score variance, and hence reliability, is zero. The reason for this is that, in the classical test theory model, all observed variability in *i*'s scores is random error by definition. Classical test theory is relevant only at the level of populations and samples, not at the level of individuals.

Cronbach's α

Cronbach's α is defined as

$$\alpha = \frac{N}{N-1}\left(\frac{\sigma_X^2 - \sum_{i=1}^{N}\sigma_{Y_i}^2}{\sigma_X^2}\right),$$

where *N* is the number of components (items or testlets), σ_X^2 is the variance of the observed total test scores, and $\sigma_{Y_i}^2$ is the variance of component *i*.

Alternatively, the standardized Cronbach's α can also be defined as

$$\alpha = \frac{N \cdot \bar{r}}{(1 + (N-1) \cdot \bar{r})}$$

where N is the number of components (items or testlets) and $\bar{r}$ is the average of all (Pearson) correlation coefficients between the components.

Cronbach's alpha will generally increase when the correlations between the items increase. For this reason the coefficient is also called the internal consistency or the internal consistency reliability of the test.

Reliability cannot be estimated directly since that would require one to observe the true scores, which according to classical test theory is impossible. However, estimates of reliability can be obtained by various means. One way of estimating reliability is by constructing a so-called "parallel test". A parallel test is a test that has the property that, for every individual, it yields the same true score and the same observed score variance as the original test. If we have parallel tests x and x', then this means that
$\varepsilon(X_i) = \varepsilon(X_i')$

and
$\sigma_{E_i}^2 = \sigma_{E_i'}^2$

Under these assumptions, it follows that the correlation between parallel test scores equals reliability.

$$\rho_{XX'} = \frac{\sigma_{XX'}}{\sigma_X \sigma_{X'}} = \frac{\sigma_T^2}{\sigma_X^2} = \rho_{XT}^2$$

The estimation of reliability by the use of parallel tests is cumbersome, because parallel tests are very hard to come by. In practice the method is rarely used. Instead, researchers use a measure of internal consistency known as Cronbach's α. Consider a test consisting of *k* items u_j, $j = 1, \ldots, j, \ldots, k$. The total test score is defined as the sum of the individual item scores, so that for individual *i*

$$X_i = \sum_{j=1}^{k} U_{ij}$$

Then Cronbach's alpha equals

$$\alpha = \frac{k}{k-1} \frac{\sum_{j=1}^{k} \sigma_{U_i}^2}{\sigma_X^2}$$

Cronbach's α can be shown to provide a lower bound for reliability under rather mild assumptions. Thus, the reliability of test scores in a population is always higher than the value of Cronbach's α in that population. Thus, this method is empirically feasible and, as a result, it is very popular among researchers.

Conclusion

As has been noted above, the entire exercise of classical test theory is done to arrive at a suitable definition of reliability. Reliability is supposed to say something about the general quality of the test scores in question. The general idea is that, the higher reliability is, the better. Classical test theory does not say how high reliability is supposed to be. In the literature a value over .80 appears to be deemed 'acceptable'; a value over .90 is 'good'. Values between .70 and .80 are seen as mediocre but still defensible; values below .70 are bad.

It must be noted that these 'criteria' are not based on reasonable arguments but the result of convention. Whether they make any sense or not is unclear.

Alternatives

Classical test theory is by far the most influential theory of test scores in the social sciences. In psychometrics, the theory has been superseded by the more sophisticated models in Item Response Theory (IRT). IRT models, however, are catching on very slowly in mainstream research. One of the main problems causing this is the lack of widely available, user-friendly software; also, IRT is not included in standard statistical packages like SPSS, whereas these packages routinely provide estimates of Cronbach's α. As long as this problem is not solved, classical test theory will probably remain the theory of choice for many researchers.

Further information

A comparison of CTT and IRT can be found in the National Council on Measurement in Education's (NCME) series "Instrucional Topics in Educational Measurement (ITEMS)":

- Hambleton R.K., Jones R.W. "Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development"
- Harvill L.M. "Standard Error of Measurement"
- Kolen M.J. "Traditional Equating Methodology"

Generalisability theory

Generalisability theory (G Theory) is a statistical framework for conceptualizing, investigating, and designing reliable observations.

The G Theory compares with the Classical test theory (CTT) where the focus is on determining the error of the measurement. Perhaps the most famous model of CTT is the equation $X = T + e$, where X is the observed score, T is the true score, and e is the error involved in our measurement. Although e could represent many different types of error (i.e., rater error, instrument error), CTT only allows us to estimate one type of error at a time.
Although CTT is suitable in the context of highly controlled laboratory conditions, variance is a part of everyday life. In field research, for example, it is unrealistic to expect that the conditions of measurement will remain constant.

Generalisability, or G, theory extends beyond CTT by recognizing that many different sources of error may affect our measurement (and that it may benefit us to examine them at the same time). The advantage of G theory, therefore, lies in the fact that researchers can estimate what proportion of the total variance in the results is due to the individual factors that often vary in assessment, such as setting, time, items, and raters.

In G theory, sources of variation are referred to as "facets". Facets are similar to the "factors" used in analysis of variance, and may include persons, raters, items/forms, time, and settings among other possibilities. The usefulness of data gained from a G study is crucially dependent on the design of the study. Therefore, the researcher must carefully consider the ways in which he/she hopes to generalize any specific results. Is it important to generalize from one setting to a larger number of settings? From one rater to a larger number of raters? From one set of items to a larger set of items? The answers to these questions will vary from one researcher to the next, and will drive the design of a G study in different ways.

In addition to deciding which facets the researcher generally wishes to examine, it is necessary to determine which facet will serve as the object of measurement (e.g. the systematic source of variance) for the purpose of analysis. The remaining facets of interest are then considered to be sources of measurement error. In most cases, the object of measurement will be the person to whom a number/score is assigned. Ideally, nearly all of the measured variance will be attributed to the object of measurement (e.g. individual differences), with only a negligible amount of variance attributed to the remaining facets (e.g., rater, time, setting).

The results from a G study can also be used to inform a decision, or D, study. In a D study, we can ask the hypothetical question of "what would happen if different aspects of this study were altered?" For example, a soft drink company might be interested in assessing the quality of a new product through use of a consumer rating scale. By employing a D study, it would be possible to estimate how the consistency of quality ratings would change if consumers were asked 10 questions instead of 2, or if 1,000 consumers rated the soft drink instead of 100. By employing simulated D studies, it is therefore possible to examine how the generalisability coefficients (similar to reliability coefficients in CTT) would change under different circumstances, and consequently determine the ideal conditions under which our measurements would be the most reliable.

Another important difference between CTT and G theory is that the latter approach takes into account how the consistency of outcomes may change if a measure is used to make absolute versus relative decisions. An example of an absolute, or criterion-referenced, decision would be when an individual's test score is compared to a cut-off score to determine eligibility or diagnosis (i.e. a child's score on an achievement test is used to determine eligibility for a gifted program). In contrast, an example of a relative, or norm-referenced, decision would be when the individual's test score is used to either (a) determine relative standing as compared to his/her peers (i.e. a child's score on a Reading subtest is used to determine which reading group he/she is placed in), or (b) make inter-individual comparisons (i.e. comparing previous versus current performance within the same individual). The type of decision that the researcher is interested in will determine which formula should be used to calculate the generalisability coefficient (similar to a reliability coefficient in CTT).

For the typical clinical exam, CTT does not work well: not all candidates can see all patients or not all candidates can be seen by all examiners. Therefore variability due to examiners and clinical scenarios exists as well as case-specificity: some candidates doing better with some types of case than others (imagine you having learned at a gastroenterology department being examined with a cardiology case). Here, G theory generalises the CTT to include such components. A measure equivalent to reliability ("generalisability") can be calculated to find out how similar a candidate's mark would be with different examiners and different scenarios or cases.

Further information
A good summary of the generalisability theory can be found in the National Council on Measurement in Education's (NCME) series "Instrucional Topics in Educational Measurement (ITEMS)":

- Brennan R.L. "Generalizability Theory"

Item response theory
Item response theory (IRT) is a body of theory describing the application of mathematical models to data from questionnaires and tests as a basis for measuring abilities, attitudes, or other variables. It is used for statistical analysis and development of assessments, often for high stakes tests such as the state exams. At its most basic level, it is based on the idea that the probability of getting an item correct is a function of a latent trait or ability. For example, a person with higher intelligence would be more likely to correctly respond to a given item on an intelligence test.

Formally, IRT models apply mathematical functions that specify the probability of a discrete outcome, such as a correct response to an item, in terms of person and item parameters. Person parameters may, for example, represent the ability of a student or the strength of a person's attitude. Item parameters include difficulty (location), discrimination (slope or correlation), and pseudoguessing (lower asymptote). Items may be questions that have incorrect and correct responses, statements on questionnaires that allow respondents to indicate level of agreement, or patient symptoms scored present/absent.

Among other things IRT theory provides a basis for evaluating how well assessments work, and how well individual questions on assessments work. In education, Psychometricians apply IRT in order to achieve tasks such as developing and refining exams, maintaining banks of items for exams, and equating for the difficulties of successive versions of exams (for example, to allow comparisons between results over time).

IRT is often referred to as "latent trait theory", "strong true score theory", or "modern mental test theory" and is distinguished from Classical test theory.

Overview
IRT models are used as a basis for statistical estimation of parameters that represent the 'locations' of persons and items on a latent continuum or, more correctly, the magnitude of the latent trait attributable to the persons and items. For example, in attainment testing, estimates may be of the magnitude of a person's ability within a specific domain, such as reading comprehension. Once estimates of relevant parameters have been obtained, statistical tests are usually conducted to gauge the extent to which the parameters predict item responses given the model used.

Stated somewhat differently, such tests are used to ascertain the degree to which the model and parameter estimates can account for the structure of and statistical patterns within the response data, either as a whole, or by considering specific subsets of the data such as response vectors pertaining to individual items or persons. This approach permits the central hypothesis represented by a particular model to be subjected to empirical testing, as well as providing information about the psychometric properties of a given assessment, and therefore also the quality of estimates.

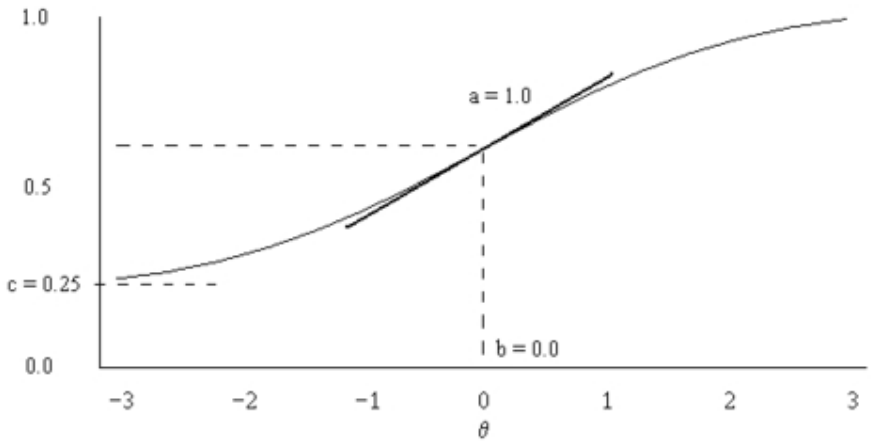
From the perspective of more traditional approaches, such as classical test theory, an advantage of IRT is that it potentially provides information that enables a researcher to improve the reliability of an assessment. This is achieved through the extraction of more sophisticated information regarding psychometric properties of individual assessment items.

IRT models are often referred to as “latent trait models”. The term “latent “is used to emphasise that discrete item responses are taken to be “observable manifestations” of hypothesized trait, construct, or attribute, not directly observed, but which must be inferred from the manifest responses. Latent trait models were developed in the field of sociology, but are virtually identitical to IRT models.

The other major body of psychometric theory of relevance to IRT is classical test theory. For tasks that can be accomplished using classical test theory, IRT generally brings greater flexibility and provides more sophisticated information. Some applications, such as computerized adaptive testing are enabled by IRT and cannot reasonably be performed using only classical test theory.

The Item Characteristic Curve

The performance of an item in a test is described by the “item characteristic curve” (ICC). The curve gives the probability that a person with a given ability level will answer the item correctly. Persons with lower ability (<0.0) have less of a chance, while persons with high ability are very likely to answer correctly.



Much of the literature on IRT centres on item response models for the ICC. A given model describes the probability of a correct response to the item as a function of a person or ability parameter (or, in the case of multidimensional item response theory, a vector of person parameters). This probability depends on one or more item parameters for the item response function (IRF). For example, in the three parameter logistic (3PL) model, the probability of a correct response to an item *i* is:

$$p_i(\theta) = c_i + \frac{(1 - c_i)}{1 + e^{-a_i(\theta - b_i)}}$$

where θ is the person (ability) parameter and a_i, b_i, and c_i are the item parameters.
The item parameters simply determine the shape of the IRF and in some cases have a direct interpretation. The figure to the right depicts an example of the 3PL model of the ICC with an overlaid conceptual explanation of the parameters. The parameter b_i represents the item location which, in the case of attainment testing, is referred to as the item difficulty. It is point on θ where the IRF has its maximum slope. The example item is of medium difficulty, since b_i=0.0, which is near the centre of the distribution. Note that this model scales the item's difficulty and the person's trait onto the same continuum. Thus, it is valid to talk about an item being about as hard as Person A's trait level or of a person's trait level being about the same as Item Y's difficulty, in the sense that successful performance of the task involved with an item reflects a specific level of ability.

The item parameter a_i represents the discrimination of the item: that is, the degree to which the item discriminates between persons in different regions on the latent continuum. This parameter characterizes the slope of the IRF where the slope is at its maximum. The example item has a_i=1.0, which discriminates fairly well; persons with low ability do indeed have a much smaller chance of correctly responding than persons of higher ability.

For items such as multiple choice items, the parameter c_i is used in attempt to account for the effects of guessing on the probability of a correct response. It indicates the probability that very low ability individuals will get this item correct by chance, mathematically represented as a lower asymptote. A four-option multiple choice item might have an IRF like the example item; there is a 1/4 chance of an extremely low ability candidate guessing the correct answer, so the c_i would be approximately 0.25. This assumes that all options are equally plausible, because if one option made no sense, even the lowest ability person would be able to discard it.

The two parameter logistic model (2PL) is equivalent to the 3PL model with c_i = 0. The 2PL model is appropriate for testing items where guessing the correct answer is highly unlikely, such as write-in tests.

Logistic and Normal IRT Models
An alternative formulation constructs IRFs based on the cumulative normal probability distribution function, these are sometimes called "normal ogive models". For example, the formula for a two-parameter normal-ogive IRF is:

$$p_i(\theta) = \Phi(\frac{\theta - b_i}{\sigma_i})$$

The normal-ogive model derives from the assumption of normally distributed measurement error and is theoretically appealing on that basis. Here b_i is, again, the difficulty parameter. The discrimination parameter is σ_i, the standard deviation of the measurement error for item i, and comparable to $1/a_i$.

With rescaling of the ability parameter, it is possible to make the 2PL logistic model closely approximate the cumulative normal ogive. Typically, the 2PL logistic and normal-ogive IRFs differ in probability by no more than 0.01 across the range of the function. The difference is greatest in the distribution tails, however, which tend to have more influence on results.
The latent trait/IRT model was originally developed using normal ogives, but, at the time this was considered computationally demanding. The logistic model was proposed as a simpler alternative, and has enjoyed wide use since.

Latent Traits and Factors
The person parameter θ represents the magnitude of "latent trait" of the individual. The estimate of the person parameter is derived from the individual's total score on the assessment, which is a weighted score when the model contains item discrimination parameters. The latent trait is the human capacity or attribute measured by the test. It might be a cognitive ability, physical ability, skill, knowledge, attitude, personality characteristic, etc. In a one dimensional model such as the one above, this trait is analogous to a single factor in factor analysis

IRT Models
Broadly speaking, IRT models can be divided into two families: unidimensional and multidimensional. Unidimensional models require a single trait (ability) dimension θ. Multidimensional IRT models model response data hypothesized to arise from multiple traits. However, because of the greatly increased complexity, the majority of IRT research and applications utilize a unidimensional model.

IRT models can also be categorized based on the number of scored responses. The typical multiple choice item is dichotomous; even though there may be four or five options, it is still scored only as

correct/incorrect (right/wrong). Another class of models apply to polytomous outcomes, where each response has a different score value. For example, the polytomous Rasch model is a generalisation of the Rasch model that applies to data in two or more ordered categories. A common example of this Likert-type items, e.g., "Rate on a scale of 1 to 5."

Dichotomous IRT models are described by the number of parameters they make use of. The 3PL is named so because it employs three item parameters. The two-parameter model assumes that the data has minimal guessing, but that items can vary in terms of location (b_i) and discrimination (a_i). The one-parameter model assumes that there is minimal guessing and that items have equivalent discriminations, so that items are only described by a single parameter (b_i).

Information
One of the major contributions of item response theory is the extension of the concept of reliability. Traditionally, reliability refers to the precision of measurement (i.e., the degree to which measurement is free of error). And traditionally, it is measured using a single index defined in various ways, such as the ratio of true and observed score variance. This index is helpful in characterizing a test's average reliability, for example in order to compare two tests. But IRT makes it clear that precision is not uniform across the entire range of test scores. Scores at the edges of the test's range, for example, generally have more error associated with them than scores closer to the middle of the range.

Item response theory advances the concept of item and test information to replace reliability. Information is also a *function* of the model parameters. For example, according to Fisher information theory, the item information supplied in the case of the Rasch model for dichotomous response data is simply the probability of a correct response multiplied by the probability of an incorrect response, or,

$$I(\theta) = p_i(\theta)q_i(\theta).$$

The standard error of estimation (SE) is the reciprocal of the test information of at a given trait level, is the

$$\mathrm{SE}(\theta) = 1/\sqrt{I(\theta)}.$$

Thus more information implies less error of measurement. For other models, such as the two and three parameters models, the discrimination parameter plays an important role in the function. The item information function for the two parameter model is

$$I(\theta) = a_i^2 p_i(\theta)q_i(\theta).$$

In general, item information functions tend to look bell-shaped. Highly discriminating items have tall, narrow information functions; they contribute greatly but over a narrow range. Less discriminating items provide less information but over a wider range.

Plots of item information can be used to see how much information an item contributes and to what portion of the scale score range. Because of local independence, item information functions are additive. Thus, the test information function is simply the sum of the information functions of the items on the exam. Using this property with a large item bank, test information functions can be shaped to control measurement error very precisely.

Characterizing the accuracy of test scores is perhaps the central issue in psychometric theory and is a chief difference between IRT and CTT. IRT findings reveal that the CTT concept of reliability is a simplification. In the place of reliability, IRT offers the test information function which shows the degree of precision at different values of theta.

These results allow psychometricians to (potentially) carefully shape the level of reliability for different ranges of ability by including carefully chosen items. For example, in a certification situation in which a test can only be passed or failed, where there is only a single "cutscore," and where the actually passing

score is unimportant, a very efficient test can be developed by selecting only items that have high information near the cutscore. These items generally correspond to items whose difficulty is about the same as that of the cutscore.

Scoring
After the model is fit to data, each person has a theta estimate. This estimate is their score on the exam. This "IRT score" is computed and interpreted in a very different manner as compared to traditional scores like number or percent correct. However, for most tests, the (linear) correlation between the theta estimate and a traditional score is very high (often it is .95 or more). A graph of IRT scores against traditional scores shows an ogive shape implying that the IRT estimates separate individuals at the borders of the range more than in the middle.

It is worth noting the implications of IRT for test-takers. Tests are imprecise tools and the score achieved by an individual (the observed score) is always the true score occluded by some degree of error. This error may push the observed score higher or lower.
Also, nothing about these models refutes human development or improvement. A person may learn skills, knowledge or even so called "test-taking skills" which may translate to a higher true-score.

A comparison of classical and Item Response theory
Classical test theory (CTT) and IRT are largely concerned with the same problems but are different bodies of theory and therefore entail different methods. Although the two paradigms are generally consistent and complementary, there are a number of points of difference:
IRT makes stronger assumptions than CTT and in many cases provides correspondingly stronger findings; primarily, characterizations of error. Of course, these results only hold when the assumptions of the IRT models are actually met.

Although CTT results have allowed important practical results, the model-based nature of IRT affords many advantages over analogous CTT findings.

CTT test scoring procedures have the advantage of being simple to compute (and to explain) whereas IRT scoring generally requires relatively complex estimation procedures (note that in the Rasch model the total score for a person is the sufficient statistic of the person parameter).
IRT provides several improvements in scaling items and people. The specifics depend upon the IRT model, but most models scale the difficulty of items and the ability of people on the same metric. Thus the difficulty of an item and the ability of a person can be meaningfully compared.

Another improvement provided by IRT is that the parameters of IRT models are generally not sample- or test-dependent whereas true-score is defined in CTT in the context of a specific test. Thus IRT provides significantly greater flexibility in situations where different samples or test forms are used. These IRT findings are foundational for computerized adaptive testing.
It is worth also mentioning some specific similarities between CTT and IRT which help to understand the correspondence between concepts. Lord (1980) showed that under the assumption that θ is normally distributed, discrimination in the 2PL model is approximately a monotonic function of the point-biserial correlation. In particular:

$$a_i \cong \frac{\rho_{it}}{\sqrt{1-\rho_{it}^2}}$$

where ρ_{it} is the point biserial correlation of item *i*. Thus, if the assumption holds, where there is a higher discrimination there will generally be a higher point-biserial correlation.
Another similarity is that while IRT provides for a standard error of each estimate and an information function, it is also possible to obtain an index for a test as a whole which is directly analogous to Cronbach's alpha, called the "separation index". To do so, it is necessary to begin with a decomposition

of an IRT estimate into a true location and error, analogous to decomposition of an observed score into a true score and error in CTT. Let

$$\hat{\theta} = \theta + \epsilon$$

where θ is the true location, and ε is the error association with an estimate. Then SE(θ) is an estimate of the standard deviation of ε for person with a given weighted score and the separation index is obtained as follows

$$R_\theta = \frac{\operatorname{var}[\theta]}{\operatorname{var}[\hat{\theta}]} = \frac{\operatorname{var}[\hat{\theta}] - \operatorname{var}[\epsilon]}{\operatorname{var}[\hat{\theta}]}$$

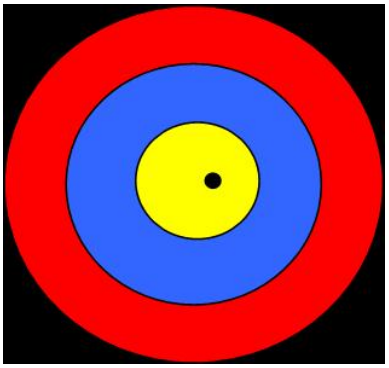
where the mean squared standard error of person estimate gives an estimate of the variance of the errors, ε_n, across persons. The standard errors are normally produced as a by-product of the estimation process (see, for example, Rasch model estimation). The separation index is typically very close in value to Cronbach's alpha (Andrich, 1982).

Further information

In the National Council on Measurement in Education's (NCME) series "Instrucional Topics in Educational Measurement (ITEMS)" several publications focus on the IRT and its use in educational measurement:

- Harris D. "Comparison of 1-, 2-, and 3-Parameter IRT Models"
- Cook L.L., Eignor D.R. "IRT Equating Methods"
- Hambleton R.K., Jones R.W. "Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development"
- Ackermann T.A., Gierl M.J., Walker C.M. "Using Multidimensional Item Response Theory to Evaluate Educational and Psychological Tests"
- Clauser B.E., Mazor K.M., "Using Statistical Procedures to Identify Differentially Functioning Test Items"
- Harvill L.M. "Standard Error of Measurement"

Validity



In psychology, validity has two distinct fields of application. The first involves test validity, a concept that has evolved with the field of psychometrics but which textbooks still commonly gloss over in explaining that it is the degree to which a test measures what it was designed to measure. The second involves research design. Here the term refers to the degree to which a study supports the intended conclusion drawn from the results. In the Campbellian tradition, this latter sense divides into four aspects: support for the conclusion that the causal variable caused the effect variable in the specific study (internal validity), support that the same effect generalizes to the population from which the sample was drawn (statistical conclusion validity), support for the intended interpretation of the variables (construct validity), and support for the generalization of the results beyond the studied population (external validity).

Introduction

An early definition of test validity identified it with the degree of correlation between the test and a criterion. Under this definition, one can show that reliability of the test and the criterion places an upper limit on the possible correlation between them (the so-called validity coefficient). Intuitively, this reflects the fact that reliability involves freedom from random error and random errors do not correlate with one another. Thus, the less random error in the variables, the higher the possible correlation

between them. Under these definitions, a test cannot have high validity unless it also has high reliability. However, the concept of validity has expanded substantially beyond this early definition and the classical relationship between reliability and validity need not hold for alternative conceptions of reliability and validity.

Within classical test theory, predictive or concurrent validity (correlation between the predictor and the predicted) cannot exceed the square root of the correlation between two versions of the same measure — that is, reliability limits validity.

Test validity can be assessed in a number of ways and thorough test validation typically involves more than one line of evidence in support of the validity of an assessment method (e.g. structured interview, personality survey, etc). The current Standards for Educational and Psychological Measurement cover various types of validity evidence for a single summative validity judgment. These include construct related evidence, content related evidence, and criterion related evidence which breaks down into two subtypes (concurrent and predictive) according to the timing of the data collection.

Construct related evidence involves the empirical and theoretical support for the interpretation of the construct. Such lines of evidence include statistical analyses of the internal structure of the test including the relationships between responses to different test items. They also include relationships between the test and measures of other constructs. As currently understood, construct validity is not distinct from the support for the substantive theory of the construct that the test is designed to measure. As such, experiments designed to reveal aspects of the causal role of the construct also contribute to construct validity evidence.

Content related evidence involves the degree to which the content of the test matches a content domain associated with the construct. For example, a test of the ability to add two-digit numbers should cover the full range of combinations of digits. A test with only one-digit numbers, or only even numbers, would not have good coverage of the content domain. Content related evidence typically involves subject matter experts (SME's) evaluating test items against the test specifications.

Criterion related evidence involves the correlation between the test and a criterion variable (or variables) taken as representative of the construct. For example, employee selection tests are often validated against measures of job performance. Measures of risk of recidivism among those convicted of a crime can be validated against measures of recidivism. If the test data and criterion data are collected at the same time, this is referred to as concurrent validity evidence. If the test data is collected first in order to predict criterion data collected at a later point in time, then this is referred to as predictive validity evidence.

Face validity is an estimate of whether a test appears to measure a certain criterion; it does not guarantee that the test actually measures phenomena in that domain. Indeed, when a test is subject to faking (malingering), low face validity might make the test more valid.
In contrast to test validity, assessment of the validity of a research design generally does not involve data collection or statistical analysis but rather evaluation of the design in relation to the desired conclusion on the basis of prevailing standards and theory of research design.

Internal validity
Internal validity is an inductive estimate of the degree to which conclusions about causes of relations are likely to be true, in view of the measures used, the research setting, and the whole research design. Good experimental techniques in which the effect of an independent variable on a dependent variable is studied under highly controlled conditions, usually allow for higher degrees if internal validity than, for example, single-case designs.

Eight extraneous variables can interfere with internal validity:

1. History, the specific events occurring between the first and second measurements in addition to the experimental variables
2. Maturation, processes within the participants as a function of the passage of time (not specific to particular events), e.g., growing older, hungrier, more tired, and so on.
3. Testing, the effects of taking a test upon the scores of a second testing.
4. Instrumentation, changes in calibration of a measurement tool or changes in the observers or scorers may produce changes in the obtained measurements.
5. Statistical regression, operating where groups have been selected on the basis of their extreme scores.
6. Selection, biases resulting from differential selection of respondents for the comparison groups.
7. Experimental mortality, or differential loss of respondents from the comparison groups.
8. Selection-maturation interaction, etc. e.g., in multiple-group quasi-experimental designs

External validity
The issue of External validity concerns the question to what extent one may safely generalize the (internally valid) causal inference (a) from the sample studied to the defined target population and (b) to other populations (i.e. across time and space).

Four factors jeopardizing external validity or representativeness are:
1. Reactive or interaction effect of testing, a pretest might increase
2. Interaction effects of selection biases and the experimental variable.
3. Reactive effects of experimental arrangements, which would preclude generalization about the effect of the experimental variable upon persons being exposed to it in non-experimental settings
4. Multiple-treatment interference, where effects of earlier treatments are not erasable.

Ecological validity
This issue is closely related to external validity and covers the question to which degree your experimental findings mirror what you can observe in the real world (ecology= science of interaction between organism and its environment). Ecological validity is whether the results can be applied to real life situations. Typically in science, you have two domains of research: Passive-observational and active-experimental. The purpose of experimental designs is to test causality, so that you can infer A causes B or B causes A. But sometimes, ethical and/or methological restrictions prevent you from conducting an experiment (e.g. how does isolation influence a child's cognitive functioning?) Then you can still do research, but it's not causal, it's correlational, A occurs together with B. Both techniques have their strengths and weaknesses. To get an experimental design you have to control for all interfering variables.

That's why you conduct your experiment in a laboratory setting. While gaining internal validity (excluding interfering variables by keeping them constant) you lose ecological validity because you establish an artificial lab setting. On the other hand with observational research you can't control for interfering variables (low internal validity) but you can measure in the natural (ecological) environment, thus at the place where behaviour occurs.

Construct validity
Construct validity refers to the totality of evidence about whether a particular operationalisation of a construct adequately represents what is intended by theoretical account of the construct being measured. (Demonstrate an element is valid by relating it to another element that is supposively valid.) There are two approaches to construct validity- sometimes referred to as 'convergent validity' and 'divergent validity'.

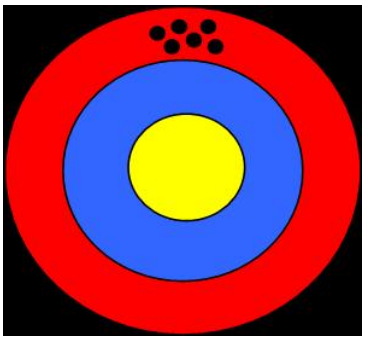
Content validity
This is a non-statistical type of validity that involves the systematic examination of the test content to determine whether it covers a representative sample of the behaviour domain to be measured.

A test has content validity built into it by careful selection of which items to include. Items are chosen so that they comply with the test specification which is drawn up through a thorough examination of the subject domain. By using a panel of experts to review the test specifications and the selection of items the content validity of a test can be improved. The experts will be able to review the items and comment on whether the items cover a representative sample of the behaviour domain.

Face validity
Face validity is very closely related to content validity. While content validity depends on a theoretical basis for assuming if a test is assessing all domains of a certain criterion (e.g. does assessing addition skills yield in a good measure for mathematical skills? - To answer this you have to know, what different kinds of arithmetic skills mathematical skills include).
Face validity relates to whether a test appears to be a good measure or not. This judgment is made on the "face" of the test, thus it can also be judged by the amateur.

Criterion validity
Criterion-related validity reflects the success of measures used for prediction or estimation. There are two types of criterion-related validity: Concurrent and predictive validity. A good example of criterion-related validity is in the validation of employee selection tests; in this case scores on a test or battery of tests is correlated with employee performance scores.

Reliability



Reliability is the consistency of a set of measurements or measuring instrument, often used to describe a test. This can either be whether the measurements of the same instrument give or are likely to give the same measurement (test-retest), or in the case of more subjective instruments, such as oral exams or the assessment of practical skills, whether two independent assessors give similar scores (inter-rater reliability). Reliability is inversely related to random error.
Reliability does not imply validity. That is, a reliable measure is measuring something consistently, but not necessarily what it is supposed to be measuring (e.g a scale can be valid for weighs but does not measure temperatures; a MCQ test does measure knowledge but not practical skills). In terms of accuracy and precision, reliability is precision, while validity is accuracy. An often-used example used to elucidate the difference between reliability and validity in the experimental sciences is a common bathroom scale. If someone that weighs 200 lbs. steps on the scale 6 times, and it reads "200" each time, then the measurement is reliable and valid. If the scale consistently reads "150", then it is not valid, but it is still reliable because the measurement is very consistent.

A common misconception of reliability is that objective assessment (such as the OSCE or MCQ) is always reliable and subjective assessments are always unreliable. Reliability is dependent on characteristics of the test, the conditions of administration, and the group of examinees. A test or assessment by itself is neither reliable nor unreliable. Factors concerning characteristics of the test are test length, item type, and item quality. Conditions of administration contributing to the reliability are proper instructions, time limits, the person administering the test, or physical conditions under which the test is taken.

Estimation of reliability
Reliability may be estimated through a variety of methods that fall into two types: Single-administration and multiple-administration. Multiple-administration methods require that two assessments are administered. In the test-retest method, reliability is estimated between two administrations of the same measure. In the “alternate forms” method, reliability is estimated by the Pearson product-moment correlation coefficient of two different forms of a measure, usually administered together. Single-administration methods include “split-half” and “internal consistency”. The split-half method treats the two halves of a measure as alternate forms. This "halves reliability" estimate is then stepped up to the

full test length. The most common internal consistency measure is Cronbach's alpha, which is usually interpreted as the mean of all possible split-half coefficients.

Each of these estimation methods is sensitive to different sources of error and so might not be expected to be equal. Also, reliability is a property of the "scores of a measure" rather than the measure itself and are thus said to be "sample dependent". Reliability estimates from one sample might differ from those of a second sample (beyond what might be expected due to sampling variations) if the second sample is drawn from a different population because the true reliability is different in this second population. (This is true of measures of all types--yardsticks might measure houses well yet have poor reliability when used to measure the lengths of insects.)

Reliability may be improved by clarity of expression (for written assessments), lengthening the measure, and other informal means. However, formal psychometric analysis, called "item analysis", is considered the most effective way to increase reliability. This analysis consists of computation of item difficulties and item discrimination indices, the latter index involving computation of correlations between the items and sum of the item scores of the entire test. If items that are too difficult, too easy, and/or have near-zero or negative discrimination are replaced with better items, the reliability of the measure will increase.

Reliability in classical test theory
In classical test theory, reliability is defined mathematically as the ratio of the variation of the *true score* and the variation of the *observed score*. Or, equivalently, one minus the ratio of the variation of the *error score* and the variation of the *observed score*:

$$\rho_{xx'} = \frac{\sigma_T^2}{\sigma_X^2} = 1 - \frac{\sigma_E^2}{\sigma_X^2}$$

where $\rho_{xx'}$ is the symbol for the reliability of the observed score, *X*; σ_X^2, σ_T^2, and σ_E^2 are the variances on the measured, true and error scores respectively. Unfortunately, there is no way to directly observe or calculate the true score, so a variety of methods are used to estimate the reliability of a test.
Some examples of the methods to estimate reliability include test-retest reliability, internal consistency reliability, and parallel-test reliability. Each method comes at the problem of figuring out the source of error in the test somewhat differently.

Reliability in item response theory
It was well-known to classical test theorists that measurement precision is not uniform across the scale of measurement. Tests tend to distinguish better for test-takers with moderate trait levels and worse among high- and low-scoring test-takers. Item response theory extends the concept of reliability from a single index to a function called the "information function". The IRT information function is the inverse of the conditional observed score standard error at any given test score. Higher levels of IRT information indicate higher precision and thus greater reliability.

Further information
A good summary of reliability can be found in the National Council on Measurement in Education's (NCME) series "Instrucional Topics in Educational Measurement (ITEMS)". There Traub R.E. and Rowley G.L. have published "Understanding reliability". The series is available online on the NCME website. In the same series, Frisbie D.A. published "Reliability of Test Scores From Teacher-Made Tests".

Summary reliability and validity
A valid assessment is one which measures what it is intended to measure. For example, it would not be valid to assess driving skills through a written test alone. A more valid way of assessing driving skills would be through a combination of tests that help determine what a driver knows, such as through a written test of driving knowledge, and what a driver is able to do, such as through a performance assessment of actual driving. Teachers frequently complain that some examinations do not properly

assess the syllabus upon which the examination is based; they are, effectively, questioning the validity of the exam.

Reliability relates to the consistency of an assessment. A reliable assessment is one which consistently achieves the same results with the same (or similar) cohort of students. Various factors affect reliability – including ambiguous questions, too many options within a question paper, vague marking instructions and poorly trained markers.

A good assessment has both validity and reliability, plus other quality for a specific context and purpose. In practice, an assessment is rarely totally valid or totally reliable. A ruler which is marked wrong will always give the same (wrong) measurements. It is very reliable, but not very valid. Asking random individuals to tell the time without looking at a clock or watch is sometimes used as an example of an assessment which is valid, but not reliable. The answers will vary between individuals, but the average answer is probably close to the actual time. In many fields, such as medical research, educational testing, and psychology, there will often be a trade-off between reliability and validity. A history test written for high validity will have many essay and fill-in-the-blank questions. It will be a good measure of mastery of the subject, but difficult to score completely accurately. A history test written for high reliability will be entirely multiple choice. It isn't as good at measuring knowledge of history, but can easily be scored with great precision. We may generalise from this. The more reliable is our estimate of what we purport to measure, the less certain we are that we are actually measuring that aspect of attainment. It is also important to note that there are at least thirteen sources of invalidity, which can be estimated for individual students in test situations. They never are. Perhaps this is because their social purpose demands the absence of any error, and validity errors are usually so high that they would destabilise the whole assessment industry.

Purpose of assessment

When students were asked about the purpose of assessment (Duffield et al 2002, Med Educ 36:879-886) in an ideal world, they ranked the following from high to low:

1. ensuring competence
2. providing feedback
3. evaluating the curriculum
4. guiding student learning
5. predicting performance as a doctor

Of the five suggested purposes, only 1 and 5 are summative and 2, 3 and 4 are formative. Students in general would prefer assessment with a formative component, as this allows them to grow and learn especially where complex behaviours are measured.

The National Board of Medical Examiners in the US has defined the following purposes of assessment:

- To communicate to students what material is important
- To motivate students to study
- To identify areas of deficiency in need of remediation or further learning
- Determine final grades or make promotion decisions
- To identify areas where the course/curriculum is weak

The content of the exam should match course/clerkship objectives. Important topics should be weighted more heavily than less important ones. The testing time devoted to each topic should reflect the relative importance of the topic and the sample of items should be representative of the instructional goals.

Further information
The National Council on Measurement in Education's (NCME) series "Instrucional Topics in Educational Measurement (ITEMS)" includes articles about this topic:

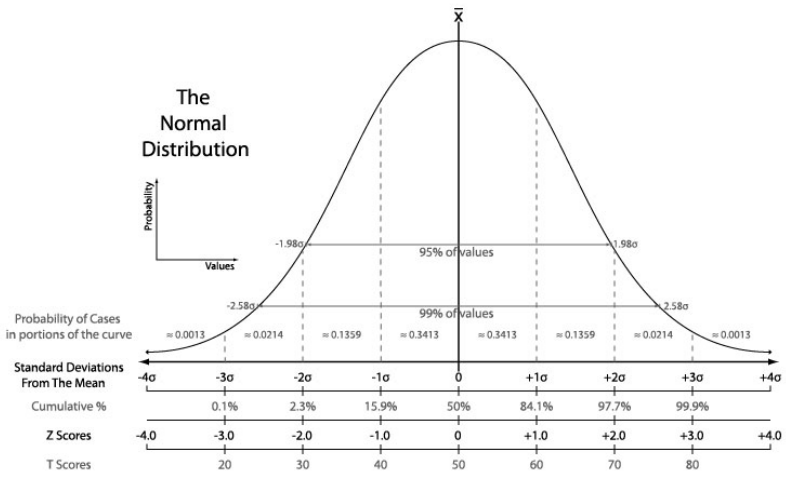
- Stiggins R.J. "High Quality Classroom Assessment: What Does It Really Mean?"
- Stiggins R.J. "Design and Development of Performance Assessment"

Criterion-based assessment

Typically using a criterion-referenced test, criterion-based assessment occurs when candidates are measured against defined (and objective) criteria. Criterion-referenced assessment is often, but not always, used to establish a person's competence (whether s/he can do something). The best known example of criterion-referenced assessment is the driving test, when learner drivers are measured against a range of explicit criteria (such as "Not endangering other road users").

Norm-referenced assessment

Norm-referenced assessment, typically using a norm-referenced test, is not measured against defined criteria. This type of assessment is relative to the student body undertaking the assessment. It is effectively a way of comparing students. The IQ test is the best known example of norm-referenced assessment. Many entrance tests (to prestigious schools or universities) are norm-referenced, permitting a fixed proportion of students to pass ("passing" in this context means being accepted into the school or university rather than an explicit level of ability). This means that standards may vary from year to year, depending on the quality of the cohort; criterion-referenced assessment does not vary from year to year (unless the criteria change).



Grading on a bell curve (as norm-referenced assessment usually does) is a method of assigning grades designed to yield a desired distribution of grades among the students in a class.
Strictly speaking, grading "on a bell curve" refers to the assigning of grades according to the frequency distribution known as the Normal distribution (also called the Gaussian distribution), whose graphical representation is referred to as the Normal curve or the bell curve. Because bell curve grading assigns grades to students based on their relative performance in comparison to classmates' performance, the term "bell curve grading" came, by extension, to be more loosely applied to any method of assigning grades that makes use of comparison between students' performances, though this type of grading does not necessarily actually make use of any frequency distribution such as the bell-shaped Normal distribution.

In true use of bell curve grading, students' scores are scaled according to the frequency distribution represented by the Normal curve. The instructor can decide what grade occupies the centre of the distribution. This is the grade an average score will earn, and will be the most common. Traditionally, in the ABCDF system this is the 'C' grade. The instructor can also decide what portion of the frequency distribution each grade occupies and whether or not high and low grades are symmetrically assigned area under the curve (i.e. if the top 15% of students earn an 'A,' do the bottom 15% fail or might only the bottom 5% fail?). In a system of pure curve grading, the number of students who will receive each grade is already determined at the beginning of a course.

Other forms of "curved" grading vary, but one of the most common is to add to all students' absolute scores the difference between the top student's score and the maximum possible score. For example, if the top score on an exam is 55 out of 60, all students' absolute scores (meaning they have not been adjusted relative to other students' scores in any way) will be increased by 5 before being compared to a pre-determined set of grading benchmarks (for example the common A>90%>B>80% etc. system). This method prevents unusually hard assignments (usually exams) from unfairly reducing students' grades but relies on the assumption that the top student's performance is a good measure of an assignment's difficulty. In the U.S., strict bell-curve grading is unusual at the elementary and secondary school levels (both in age-based grade placement and in standardized testing), but common at the university level.

Benefits and shortcomings
Viewed practically, curved grading is beneficial (to test-givers, not test-takers) because it automatically factors in the difficulty a group of test-takers had with a test. If the majority of students have high (or low) scores then the middling grade will be adjusted there and higher or lower grades awarded based on this performance. In addition, the curve ameliorates the problem of deciding grades that fall very near a grade margin. Clustering of marks establish where the margin should be placed.

However, grading in this way is essentially normative; scores are referenced to the performance of group member. There must always be at least one student who has a lower score than all others, even if that score is quite high when evaluated against specific performance criteria or standards. Conversely, if all students perform poorly relative to a larger population, even the highest graded students may be failing to meet standards. Thus, curved grading makes it difficult to compare groups of students to one another.
An additional shortcoming is that many students can easily become confused between their relative and absolute grades

Further information
A good summary on standard making can be found in the National Council on Measurement in Education's (NCME) series "Instrucional Topics in Educational Measurement (ITEMS)":

- Cizek G.J. "Standard-Setting Guidelines"
- Cizek G.J., Bunch M.B., Koons H. "Setting Performance Standards: Contemporary Methods"

Key Considerations for Selecting Assessment Instruments and Implementing Assessment Systems

The “Accreditation Council for Graduate Medical Education (ACGME)” of the United States has published these guidelines for selecting assessment instruments and implementing assessment systems.

The ability to demonstrate educational outcomes as the achievement of competency-based learning objectives provides evidence of preparing competent physicians who can meet the health care needs of the public. Educational assessment is, therefore, a key component of the Outcome Project and is intended to:

1. Assess residents' attainment of competency-based objectives
2. Facilitate continuous improvement of the educational experience
3. Facilitate continuous improvement of resident performance
4. Facilitate continuous improvement of residency program performance

Assessment is defined as the "process of collecting, synthesizing, and interpreting information to aid decision-making". The results of an assessment should allow sound inferences about what learners know, believe, and can do in defined contexts. Assessment, therefore, integrates several concepts, which are described below.

Assessment Instrument or Approach

1. The assessment approach provides valid data.
   Valid data provide accurate information about what is being assessed. Different types of evidence may be used to infer validity. It may be inferred when assessment results help to predict performance in actual practice. Validity may be inferred also when it is possible to detect change (responsiveness). This occurs, for example, when residents perform poorly on a cardiology assessment prior to completing a cardiology rotation, but perform well on the same assessment following the rotation. In addition, validity may be inferred when there is a strong relationship between data obtained and external indicators (discriminative validity). An example of the latter occurs when medical students perform poorly and cardiologists perform well on the same cardiology quiz. As knowledge about complex assessment advances, however, it is possible that perspectives on validity also will evolve.
2. The assessment approach yields reliable data.
   An assessment approach may be considered reliable when it yields consistent results regardless of when it is used, who uses it, and which item or case is assessed. The importance of a specific type of reliability depends upon what is being assessed and the method by which it is being assessed. Generally speaking, reliability or generalizability coefficients of 0.8 and higher are desired. Inter-observer or inter-rater reliability is an indicator that different assessors have provided similar ratings for the same performance. Inter-case or inter-item reliability is the degree of consistency in an individual's performance across different cases, situations, or items. Test-retest reliability is an indicator of consistency over time. Generalizability theory offers an alternative approach to assessing the individual reliabilities listed above by allowing examination of specific sources of unreliability and providing an overall reliability index termed a G coefficient.
3. The assessment approach is feasible.
   Feasibility depends on several issues that include the following: time and training required

implementing the assessment, equipment or technology required, number of assessments required per examinee, financial cost, and the extent to which an assessment has been used.

4. The assessment approach is likely to apply to my assessment circumstances (external validity).
   When choosing an assessment approach, the conditions in which an assessment has been previously conducted should be considered. These conditions include the purpose for which the assessment was used, the characteristics of those assessed and the assessors, and the setting in which the assessment was conducted. Assessments that have been used in testing centres, for instance, may require modification for use in clinics or wards where the pace may vary and interruptions may occur.
5. The assessment provides valuable information.
   In terms of value, assessment should provide new and useful information that facilitates teaching and learning. For instance, the assessment should allow the collection of enough detailed information that it is possible to know what performance improvements or curricular modifications are needed.

Assessment System

1. Assessment is consistent with curriculum/program objectives.
   Consistency between objectives and assessment occurs when there are clear parallels between what is taught and what is assessed. If , for example, a course is designed to improve knowledge and procedural skills required to conduct upper endoscopies, then both knowledge and skills in this area should be assessed. Consistency between objectives and assessment also increases the likelihood that learners will attend to a broader scope of course objectives and not just content that will be assessed.
2. The educational objectives are representative of the educational domains of interest.
   It is not feasible to assess attainment of all educational objectives in all contexts; therefore, it is necessary to select a sample of what will be assessed. Representative behaviours for each competency in defined contexts should be identified. For the medical knowledge competency, identification may be guided by considering, for instance, common acute and chronic problems that occur in ambulatory settings of specific specialties. For the professionalism competency, development of educational objectives might be guided by considering common ethical dilemmas, relevant cultural contexts of patient care, and key professional courtesies intrinsic to patient care and teamwork for specific specialties in defined settings.
3. Multiple assessment approaches/instruments are employed.
   Because competence is multi-dimensional and individual assessment approaches have limitations, it is unlikely that a single approach to assessment will be adequate. This problem is addressed by using a few different assessment approaches.
4. Multiple observations are conducted.
   Multiple observations improve the reliability or precision of assessment and allow identification of patterns of behaviour over time.
5. Multiple observers/raters provide assessments.
   Using multiple observers improves the reliability or precision of assessment and enhances the scope of assessment.
6. Performance is assessed according to pre-specified standards or criteria.
   Pre-specified standards indicate objective criteria for "good enough" or "borderline" performance and help to reduce subjective assessment.
7. Assessment is fair.
   Fairness pertains to giving all learners the same or equal opportunity to perform. While

fairness may be enhanced by valid and reliable assessment, an assessment may still be unfair if the results are influenced by something other than ability. For example, it would be unfair to compare the assessment results of a learner who was on call the night before an assessment with the results of peers who were not on call. With the exception of baseline or needs assessments, fairness pertains also to providing learners opportunities to learn the material on which they will be assessed. Learners should be informed about what will and will not be assessed. In addition, there should be clarity about the assessment format and how performance will be rated.

Assessment methods

UNESCO-CEPES defines "Assessment" as

1. The process of the systematic gathering, quantifying, and using of information in view of judging the instructional effectiveness and the curricular adequacy of a higher education institution as a whole (institutional assessment) or of its educational programmes (*programme assessment*). It implies the evaluation of the core activities of the higher education institution (quantitative and qualitative evidence of educational activities and research outcomes). Assessment is necessary in order to validate a formal accreditation decision, but it does not necessarily lead to an accreditation outcome.
2. A technically designed process for evaluating student learning outcomes and for improving student learning and development as well as teaching effectiveness (*students assessment*).

The „Accreditation Council for Graduate Medical Education (ACGME)" of the United States has compiled a toolbox of assessment methods which describes some methods used by the council in brief.

Out of the following explanations those from this list were taken out of the „Toolbox of Assessment Methods" of the SCGME Outcomes Project *(©2000 ACGME and ABMS. A product of the joint initiative of the ACGME Outcome Project of the Accreditation Council for Graduate Medical Education (ACGME), and the American Board of Medical Specialities (ABMS). Version 1.1, September 2000)*:

- 360-Degree Evaluation Instrument
- Chart Stimulated Recall Oral Examination (CSR)
- Checklist Evaluation
- Global Rating of Live or Recorded Performance
- Objective Structured Clinical Examination (OSCE)
- Procedure, Operative, or Case Logs
- Patient Surveys
- Portfolios
- Record Review
- Simulations and Models
- Standardized Oral Examination
- Standardized Patient Examination (SP)
- Written Examination (MCQ)

This list does not cover all assessment methods known and available for medical education. There are many more assessment methods like „Key feature-Problems" that might be added in future manuals.

360-Degree Evaluation Instrument

Description

360-degree evaluations consist of measurement tools completed by multiple people in a person's sphere of influence. Evaluators completing rating forms in a 360-degree evaluation usually are superiors, peers, subordinates, and patients and families. Most 360-degree evaluation processes use a survey or questionnaire to gather information about an individual's performance on several topics (e.g., teamwork, communication, management skills, decision-making). Most 360-degree evaluations use rating scales to assess how frequently a behavior is performed (e.g., a scale of 1 to 5, with 5 meaning "all the time" and 1 meaning "never"). The ratings are summarized for all evaluators by topic and overall to provide feedback.

Use

Evaluators provide more accurate and less lenient ratings when the evaluation is intended to give formative feedback rather than summative evaluations. A 360-degree evaluation can be used to assess interpersonal and communication skills, professional behaviors, and some aspects of patient care and systems-based practice.

Psychometric qualities

No published reports of the use of 360-degree evaluation instruments in graduate medical education were found in the literature; however, there are reports of the use of various categories of people evaluating residents at the same time, although with different instruments. Generally the evaluators were nurses, allied health professionals, other residents, faculty/supervisors, and patients. Moderate correlations were found to exist among the scores produced by these evaluators using slightly different assessment tools. Reproducible results were most easily obtainable when five to ten nurses rated residents, while a greater number of faculty and patients were needed for the same degree of reliability. In business, military and education settings, reliability estimates have been reported as great as 0.90 for 360-degree evaluation instruments.

Feasibility / Practicality

In most clinical settings conducting 360-degree-evaluations will pose a significant challenge. The two practical challenges are: constructing surveys that are appropriate for use by all evaluators in the circle of influence, and orchestrating data collection from a potentially large number of individuals that can be compiled and reported confidentially to the resident. Implementing an electronic system should make the 360-degree-evaluation feasible.

Suggested reference

Center for Creative Leadership, Greensboro, North Carolina (http://www.ccl.org).

Chart Stimulated Recall Oral Examination (CSR)

Description

In a chart stimulated recall (CSR) examination patient cases of the examinee (resident) are assessed in a standardized oral examination. A trained and experienced physician examiner questions the examinee about the care provided probing for reasons behind the work-up, diagnoses, interpretation of clinical findings, and treatment plans. The examiners rate the examinee using a well-established protocol and scoring procedure. In efficiently designed CSR oral exams each patient case (test item) takes 5 to 10 minutes. A typical CSR exam is two hours with one or two physicians as examiners per separate 30 or 60-minute session.

Use

These exams assess clinical decision-making and the application or use of medical knowledge with actual patients. Multiple-choice questions are better than CSR at assessing recall or understanding of medical knowledge. Five of the 24 ABMS Member Boards use CSR as part of their standardized oral examinations for initial certification.

Psychometric qualities

Patient cases are selected to be a sample of patients the examinee should be able to manage successfully, for example, as a board certified specialist. One or more scores are derived for each case based upon pre-defined scoring rules. The examinee's performance is determined by combining scores from all cases for a pass/fail decision overall or by each session. If the CSR is used for certification, test scores are analyzed using sophisticated statistical methods (e.g., Item Response Theory (IRT) or generalizability theory) to obtain a better estimate of the examinee's ability. Exam score reliabilities have been reported between 0.65 and 0.88 (1.00 is considered perfect reliability). The physician examiners need to be trained in how to question the examinee and evaluate and score the examinee's responses.

Feasibility / Practicality

"Mock orals," that use resident's cases but with much less standardization compared to board oral exams, often are used in residency training programs to help familiarize residents with the oral exams conducted for board certification. CSR oral exams can be implemented easily to determine if residents can apply knowledge appropriately in managing patients, but for the exams to be used for high stakes decisions about the resident's abilities such as board certification extensive resources and expertise are required to standardize the exam.

Suggested reference

Munger, BS. Oral examinations. In Mancall EL, Bashook PG. (editors) Recertification: new evaluation methods and strategies. Evanston, Illinois: American Board of Medical Specialties, 1995: 39-42.

Checklist Evaluation

Description

Checklists consist of essential or desired specific behaviors, activities, or steps that make up a more complex competency or competency component. Typical response options on these forms are a check (☐) or "yes" to indicate that the behavior occurred or options to indicate the completeness (complete, partial, or absent) or correctness (total, partial, or incorrect) of the action. The forms provide information about behaviors but for the purpose of making a judgment about the adequacy of the overall performance, standards need to be set that indicate, for example, pass/fail or excellent, good, fair, or poor performance.

Use

Checklists are useful for evaluating any competency and competency component that can be broken down into specific behaviors or actions. Documented evidence for the usefulness of checklists exists for the evaluation of patient care skills (history and physical examination, procedural skills) and for interpersonal and communication skills. Checklists have also been used for self-assessment of practice-based learning skills (evidence-based medicine). Checklists are most useful to provide feedback on performance because checklists can be tailored to assess detailed actions in performing a task.

Psychometric qualities

When observers are trained to use checklists, consistent scores can be obtained and reliability in the range of 0.7 to 0.8 is reported (1.0 is perfect reliability). Performance scores derived from checklists can discriminate between residents in different years of training. Scoring practitioners' behavior using checklists is more difficult when checklists assume a fixed sequence of actions because experienced physicians use various valid sequences and are usually parsimonious in their patient care behaviors.

Feasibility / Practicality
To ensure validity of content and scoring rules, checklist development requires consensus by several experts with agreement on essential behaviors/actions, sequencing, and criteria for evaluating performance. Checklists require trained evaluators to observe performance and time to complete a checklist will vary depending on the observation period.
Suggested references
Noel G, Herbers JE, Caplow M et al. How well do Internal Medicine faculty members evaluate the clinical skills of residents? Ann Int Med. 1992; 117: 757-65.
Winckel CP, Reznick RK, Cohen R, Taylor B. Reliability and construct validity of a structured technical skills assessment form. Am J Surg. 1994; 167: 423-27.

Global rating of live or recorded performance

Description
Global rating forms are distinguished from other rating forms in that (a) a rater judges general categories of ability (e.g. patient care skills, medical knowledge, interpersonal and communication skills) instead of specific skills, tasks or behaviors; and (b) the ratings are completed retrospectively based on general impressions collected over a period of time (e.g., end of a clinical rotation) derived from multiple sources of information (e.g., direct observations or interactions; input from other faculty, residents, or patients; review of work products or written materials). All rating forms contain scales that the evaluator uses to judge knowledge, skills, and behaviors listed on the form. Typical rating scales consist of qualitative indicators and often include numeric values for each indicator, for example, (a) very good = 1, good =2, fair = 3, poor =4; or (b) superior =1, satisfactory =2, unsatisfactory =3. Written comments are important to allow evaluators to explain the ratings.
Use
Global rating forms are most often used for making end of rotation and summary assessments about performance observed over days or weeks. Scoring rating forms entails combining numeric ratings with comments to obtain a useful judgment about performance based upon more than one rater.
Psychometric qualities
A number of problems with global ratings have been documented: scores can be highly subjective when raters are not well trained; sometimes all competencies are rated the same regardless of performance; and scores may be biased when raters inappropriately make severe or lenient judgments or avoid using the extreme ends of a rating scale. Research reports are mixed about: discriminating between competence levels of different individuals; rating more skilled/experienced physicians better than less experienced physicians; and reproducibility (reliability) of ratings by the same physician/faculty raters, across different physicians/faculty, and variability across physicians/faculty, residents, nurses, and patients ratings of the same resident. Reproducibility appears easier to achieve for ratings of knowledge and more difficult to achieve for patient care and interpersonal and communication skills. A few studies have reported that faculty give more lenient ratings than residents, especially when the residents believe that the ratings will not be used for pass/fail decisions.
Feasibility / Practicality
Basic global rating forms can be constructed and completed quickly and easily. However, ratings do require time to directly observe performance or interact with the physician being evaluated. Training of raters is important to improve reproducibility of the findings.
Suggested reference
Gray, J. Global rating scales in residency education. Acad Med. 1996; 71: S55-63.

Mini Clinical Evaluation Exercise (mini-CEX or mini-ClinEX)

The mini-CEX is a method to simultaneously assess clinical skills and offer feedback to the trainee.

In the early 70s, the American Board of Internal Medicine introduced the "Clinical Evaluation Exercise" (CEX) to assess young doctors in their specialist training. The CEX and the mini-CEX are methods of formative assessment (even though they could be used for summative assessment as well) aiming to enhance future performance rather than to judge faults and mistakes.

Compared to the formerly used bedside oral examination this new method has had many advantages. A single faculty member is assessing the performance of a trainee in examining a pre-selected patient. The trainee performs a complete history and physical examination, then comes to a diagnostic and therapeutic conclusion and presents his finding in a written report. This takes approximately 2 hours. Compared to other methods of assessment some advantages of this kind of assessment were

- Performance on a real patient
- Provision of educational feedback
- Constructive criticism
- And a complete and realistic clinical challenge for the trainee

Anyway research from the 80s and 90s revealed some problems of the CEX:

The results were not likely to generalise and performance in CEX turned out not to be a good predictor for other patient cases. Since the whole assessment took more than 2 – 3 hours, only few CEXs were carried out in a trainees' career. Also the perception of the performance differed a lot between different teachers. So the reliability and validity of the CEX were poor. Also the trainee was uninfluenced by time constraints so that the examination was not as realistic as thought.

So the mini-CEX was developed. The time the trainee spends with the patient now was limited to 15 minutes. The faculty member now stays in the room and observes the trainee's performance using a standardised marking form. Afterwards feedback is given.

Compared to CEX, the mini-CEX has many advantages. Short assessments on ward are far more feasible than 2 hour long ones. More encounters are possible. Trainees see more cases, more patients and are assessed by different faculty members. They face a broader range of challenges and identification of areas of weaknesses and strengths is easier. All this adds up to increased validity and reliability.

Further information on the mini-CEX can be found in JJ Norcini's article "The mini-CEX: A method for assessing clinical skills" (Ann Intern Med 2003;138:476-481).

Objective Structured Clinical Examination (OSCE)

Description

In an objective structured clinical examination (OSCE) one or more assessment tools are administered at 12 to 20 separate standardized patient encounter stations, each station lasting 10-15 minutes. Between stations candidates may complete patient notes or a brief written examination about the previous patient encounter. All candidates move from station to station in sequence on the same schedule. Standardized patients are the primary assessment tool used in OSCEs, but OSCEs have included other assessment tools such as data interpretation exercises using clinical cases, and clinical scenarios with mannequins, to assess technical skills.

Use

OSCEs have been administered in most US medical schools, many residency programs, and by the licensure boards in Canada for more than five years. The OSCE format provides a standardized means to assess: physical examination and history taking skills; communication

skills with patients and family members, breadth and depth of knowledge; ability to summarize and document findings; ability to make a differential diagnosis, or plan treatment; and clinical judgment based upon patient notes.

Psychometric qualities

OSCEs can provide means to obtain direct measures in a standardized manner of a patient-doctor encounter. OSCEs are not useful to measure skills or abilities in continuity of care with repeated patient encounters or invasive procedures. Because OSCEs often use standardized patients the same advantages and limitations apply (See toolbox description of standardized patient examination). A separate performance score is derived for each task performed at a station and scores are combined across stations or tasks to determine a pass/fail score. Statistical weighting of scores on individual tasks is controversial and not recommended. An OSCE with 14 to 18 stations is recommended to obtain reliable measurements of performance.

Feasibility / Practicality

OSCEs are very useful to measure specific clinical skills and abilities, but are difficult to create and administer. OSCEs are only cost-effective when many candidates are to be examined at one administration. Most OSCEs are administered in medical center outpatient facilities or specially designed patient examining rooms with closed circuit television. A separate room or cubical is needed for each station. For most residency programs developing and administering an OSCE will require the resources and expertise of a consortium of residency programs in an academic institution or metropolitan area.

Suggested reference

Norman, Geoffrey. Evaluation Methods: A resource handbook. Hamilton, Ontario, Canada: Program for Educational Development, McMaster University, 1995: 71-77.

Procedure, operative, or case logs

Description

Procedure, operative, or case logs document each patient encounter by medical conditions seen, surgical operation or procedures performed. The logs may or may not include counts of cases, operations, or procedures. Patient case logs currently in use involve recording of some number of consecutive cases in a designated time frame. Operative logs in current use vary; some entail comprehensive recording of operative data by CPT code while others require recording of operations or procedures for a small number of defined categories.

Use

Logs of types of cases seen or procedures performed are useful for determining the scope of patient care experience. Regular review of logs can be used to help the resident track what cases or procedures must be sought out in order to meet residency requirements or specific learning objectives. Patient logs documenting clinical experience for the entire residency can serve as a summative report of that experience; as noted below, the numbers reported do not necessarily indicate competence.

Psychometric qualities

There are no known studies of case or procedure logs for the purpose of determining accuracy of residents' recording. Unless defined by CPT or other codes, cases or procedures counted for a given category may vary across residents and programs. Minimum numbers of procedures required for accreditation and certification have not been validated against the actual quality of performance of an operation or patient outcomes.

Feasibility / Practicality

Electronic recording devices and systems facilitate the collection and summarization of patient cases or procedures performed. Although there is considerable cost associated with development, testing, and maintenance of electronic systems, these costs generally are not paid by individual programs and institutions, since systems are available commercially for a

relatively small amount (e.g., $2500 annually) or provided free of charge by accrediting or certification bodies. Manual recording is required followed later by data entry unless automated data entry devices are located at or near the point of service. Data entry of manual records typically can be performed by a clerk, but is time consuming depending on the number of residents in the program and log reporting requirements.

Suggested reference

Watts J, Feldman WB. Assessment of technical skills. In: Neufeld V and Norman G (ed). Assessing clinical competence. New York: Springer Publishing Company, 1985: 259-74.

Patient surveys

Description

Surveys of patients to assess satisfaction with hospital, clinic, or office visits typically include questions about the physician's care. The questions often assess satisfaction with general aspects of the physician's care, (e.g., amount of time spent with the patient, overall quality of care, physician competency (skills and knowledge), courtesy, and interest or empathy). More specific aspects of care can be assessed including: the physician's explanations, listening skills and provision of information about examination findings, treatment steps, and drug side effects. A typical patient survey asks patients to rate their satisfaction with care using rating categories (e.g., poor, fair, good, very good, excellent) or agreement with statements describing the care (e.g., "the doctor kept me waiting," --Yes, always; Yes, sometimes; or No, never or hardly ever). Each rating is given a value and a satisfaction score calculated by averaging across responses to generate a single score overall or separate scores for different clinical care activities or settings.

Use

Patient feedback accumulated from single encounter questionnaires can assess satisfaction with patient care competencies (aspects of data gathering, treatment, and management; counseling, and education; preventive care); interpersonal and communication skills; professional behavior; and aspects of systems-based practice (patient advocacy; coordination of care). If survey items about specific physician behaviors are included, the results can be used for formative evaluation and performance improvement. Patient survey results also can be used for summative evaluation, but this use is contingent on whether the measurement process meets standards of reliability and validity.

Psychometric qualities

Reliability estimates of 0.90 or greater have been achieved for most patient satisfaction survey forms used in hospitals and clinics. Reliability estimates are much lower for ratings of residents in training. The American Board of Internal Medicine reports 20-40 patient responses were needed to obtain a reliability of 0.70 to 0.82 on individual resident ratings using the ABIM Patient Satisfaction Questionnaire. Low per-resident reliability has been associated with surveys that use rating scales; survey questions with response options of "yes, definitely," "yes, somewhat," or "no," may provide more reproducible, and useful results.

Feasibility / Practicality

A variety of patient satisfaction surveys are available from commercial developers and medical organizations. Creation of new surveys often begins with gathering input from patients using interviews, focus groups, or questionnaires. Physician attitudes and behaviors patients find to be satisfying or dissatisfying are then translated into survey items. Most patient satisfaction surveys are completed at the time of service, and require less than 10 minutes. Alternatively, they may be mailed after the patient goes home or conducted with patients over the phone. Difficulties encountered with patient surveys are: (1) language and literacy problems; (2) obtaining enough per-resident surveys to provide reproducible results; (3) the resources required to collect, aggregate, and report survey responses; and

(4) assessment of the resident's contribution to a patient's care separate from that of the health care team. Because of these concerns, patient satisfaction surveys are often conducted by the institution or by one or more clinical sites and reports specific to the residency program may or may not be prepared. It may be possible to improve feasibility by utilizing effective survey design principles and using computers to collect and summarize survey data.

Suggested references

Kaplan SH, Ware JE. The patient's role in health care and quality assessment. In: Goldfield N and Nash D (eds). Providing quality care (2nd ed): Future Challenge. Ann Arbor, MI: Health Administration Press, 1995: 25-52.

Matthews DA, Feinstein AR. A new instrument for patients' ratings of physician performance in the hospital setting. J Gen Intern Med. 1989:4:14-22.

Portfolios

Description

A portfolio is a collection of products prepared by the resident that provides evidence of learning and achievement related to a learning plan. A portfolio typically contains written documents but can include video- or audio-recordings, photographs, and other forms of information. Reflecting upon what has been learned is an important part of constructing a portfolio. In addition to products of learning, the portfolio can include statements about what has been learned, its application, remaining learning needs, and how they can be met. In graduate medical education, a portfolio might include a log of clinical procedures performed; a summary of the research literature reviewed when selecting a treatment option; a quality improvement project plan and report of results; ethical dilemmas faced and how they were handled; a computer program that tracks patient care outcomes; or a recording or transcript of counseling provided to patients.

Use

Portfolios can be used for both formative and summative evaluation of residents. Portfolios are most useful for evaluating mastery of competencies that are difficult to evaluate in other ways such as practice-based improvement, use of scientific evidence in patient care, professional behaviors, and patient advocacy. Teaching experiences, morning report, patient rounds, individualized study or research projects are examples of learning experiences that lend themselves to using portfolios to assess residents. The Royal College of Physicians and Surgeons of Canada in the Maintenance of Competence Program (MOCOMPS) has developed a portfolio system for recertification using Internet-based diaries called PCDiary© that could be adapted to residency evaluations.

Psychometric qualities

Reproducible assessments are feasible when there is agreement on criteria and standards for contents of a portfolio. When portfolio assessments have been used to evaluate an educational program (e.g., statewide elementary or high school program) the portfolio products or documentation have been found to be sufficient for program evaluation but are not always appropriate to use in assessing individual students for decisions about promotion to the next grade. However, standard criteria are not necessarily desirable and may be counter-productive when the portfolio purpose is to demonstrate individual learning gains relative to individual goals. The validity of portfolio assessment is determined by the extent to which the products or documentation included in a portfolio demonstrates mastery of expected learning.

Feasibility / Practicality

Acceptance of portfolios in graduate medical education varies according to preferred learning style. Some residents and practicing physicians have found that by maintaining portfolios credit was allowed for some activities that otherwise would have gone undone or un-noticed.

Yet, for others, the time and commitment necessary to create and maintain a portfolio is too great relative to the return.

Suggested reference

Challis M. AMEE medical education guide no. 11 (revised): Portfolio-based learning and assessment in medical education. Med Teach. 1999; 21: 370-86.

Arter J.A., Spandel V. NCME Instructional Module on Using Portfolios of Student Work in Instruction and Assessment. Available online at http://www.ncme.org/pubs/items.cfm.

Record review

Description

Trained staff in an institution's medical records department or clinical department perform a review of patients' paper or electronic records. The staff uses a protocol and coding form based upon predefined criteria to abstract information from the records, such as medications, tests ordered, procedures performed, and patient outcomes. The patient record findings are summarized and compared to accepted patient care standards. Standards of care are available for more than 1600 diseases on the Website of the Agency for HealthCare Research and Quality (http://www.ahrq.gov/).

Use

Record review can provide evidence about clinical decision-making, follow-through in patient management and preventive health services, and appropriate use of clinical facilities and resources (e.g., appropriate laboratory tests and consultations). Often residents will confer with other clinical team members before documenting patient decisions and therefore, the documented care may not be directly attributed to a single resident but to the clinical team.

Psychometric qualities

A sample of approximately eight to 10 patient records is sufficient for a reliable assessment of care for a diagnosis or procedure. One study in office practice demonstrated that six to eight office records selected randomly are adequate to evaluate care. Missing or incomplete documentation of care is interpreted as not meeting the accepted standard.

Feasibility / Practicality

Record reviews by trained staff take approximately 20 to 30 minutes per record on average for records of hospitalized patients. The major limitations are: (1) as a retrospective assessment of care the review may not be completed until sufficient patients have been treated which could delay reports about residents' performance for months after a typical one or two month clinical rotation; (2) criteria of care must be agreed-up and translated into coding forms for staff to review records; (3) staff must be trained in how to identify and code clinical data to assure reasonably reliable findings.

Suggested reference

Tugwell P, Dok, C. Medical record review. In: Neufeld V and Norman G (ed). Assessing clinical competence. New York: Springer Publishing Company, 1985: 142-82.

Simulations and models

Description

Simulations used for assessment of clinical performance closely resemble reality and attempt to imitate but not duplicate real clinical problems. Key attributes of simulations are that: they incorporate a wide array of options resembling reality, allow examinees to reason through a clinical problem with little or no cueing, permit examinees to make life-threatening errors without hurting a real patient, provide instant feedback so examinees can correct a mistaken action, and rate examinees' performance on clinical problems that are difficult or impossible to evaluate effectively in other circumstances. Simulation formats have been developed as paper-

and-pencil branching problems (patient management problems or PMPs), computerized versions of PMPs called clinical case simulations (CCX®), role-playing situations (e.g., standardized patients (SPs), clinical team simulations), anatomical models or mannequins, and combinations of all three formats. Mannequins are imitations of body organs or anatomical body regions frequently using pathological findings to simulate patient disease. The models are constructed of vinyl or plastic sculpted to resemble human tissue with imbedded electronic circuitry to allow the mannequin to respond realistically to actions by the examinee. Virtual reality simulations or environments (VR) use computers sometimes combined with anatomical models to mimic as much as feasible realistic organ and surface images and the touch sensations (computer generated haptic responses) a physician would expect in a real patient. The VR environments allow assessment of procedural skills and other complex clinical tasks that are difficult to assess consistently by other assessment methods.

Use

Simulations using VR environments have been developed to train and assess surgeons performing arthroscopy of the knee and other large joints, anesthesiologists managing life-threatening critical incidents during surgery, surgeons performing wound debridement and minor surgery, and medical students and residents responding to cardio-pulmonary incidents on a full-size human mannequin. Written and computerized simulations have been used to assess clinical reasoning, diagnostic plans and treatment for a variety of clinical disciplines as part of licensure and certification examinations. Standardized patients as simulations are described elsewhere.

Psychometric qualities

Studies of high-quality simulations have demonstrated their content validity when the simulation is designed to resemble a real patient. One or more scores are derived for each simulation based upon pre-defined scoring rules set by the experts in the discipline. The examinee's performance is determined by combining scores from all simulations to derive an overall performance score. When included in Objective Structured Clinical Examinations (OSCEs) the case reliabilities are similar to those reported for OSCEs (See OSCEs).

Feasibility / Practicality

Experts in a specialty carefully craft simulations as clinical scenarios from real patient cases to focus the assessments on specific skills, abilities and "key features" of the case. Technical experts in assessment and simulations then convert the scenarios into simulations as standardized patients, mannequins, computer-based simulations, and other simulations adding when feasible computer-automated scoring rules to record the examinees' actions. Simulations are expensive to create and often require producing many variations of the pathological conditions or clinical problems to make them economical. Grants and contracts from commercial vendors, foundations, governmental agencies and medical schools continue to be the principle source of funding to develop simulations.

Suggested reference

Tekian A, McGuire CH, et al (eds.) Innovative simulations for assessing professional competence. Chicago, Illinois: University of Illinois at Chicago, Dept. Med. Educ. 1999

Standardized Oral Examination

Description

The standardized oral examination is a type of performance assessment using realistic patient cases with a trained physician examiner questioning the examinee. The examiner begins by presenting to the examinee a clinical problem in the form of a patient case scenario and asks

the examinee to manage the case. Questions probe the reasoning for requesting clinical findings, interpretation of findings, and treatment plans. In efficiently designed exams each case scenario takes three to five minutes. Exams last approximately 90 minutes to two and one-half hours with two to four separate 30 or 60-minute sessions. One or two physicians serve as examiners per session. An examinee can be tested on 18 to 60 different clinical cases.

Use

These exams assess clinical decision-making and the application or use of medical knowledge with realistic patients. Multiple-choice questions are better at assessing recall or understanding of medical knowledge. Fifteen of the 24 ABMS Member Boards use standardized oral examinations as the final examination for initial certification.

Psychometric qualities

A committee of experts in the specialty carefully crafts the clinical scenarios from real patient cases to focus the assessment on the "key features" of the case. Cases are selected to be a sample of patients the examinee should be able to manage successfully, for example, as a board certified specialist. One or more scores are derived for each case based upon pre-defined scoring rules. The examinee's performance is determined by combining scores from all cases for a pass/fail decision overall or by each session. Test scores are analyzed using sophisticated statistical methods (e.g., Item Response Theory (IRT) or generalizability theory) to obtain a better estimate of the examinee's ability. Exam score reliabilities have been reported between 0.65 and 0.88 (1.00 is considered perfect reliability). The physician examiners need to be trained in how to provide patient data for each scenario, question the examinee, and evaluate and score the examinee's responses.

Feasibility / Practicality

A committee of physician specialists develops the examination cases and trains the examiners, often with assistance from psychometric experts. "Mock orals," that use cases but with much less standardization compared to board oral exams, are often used in residency training programs to help familiarize residents with the oral exams conducted for board certification. Extensive resources and expertise are required, however, to develop and administer a standardized oral examination.

Suggested reference

Mancall EL, Bashook PG. (eds.) Assessing clinical reasoning: the oral examination and alternative methods. Evanston, Illinois: American Board of Medical Specialties, 1995.

Standardized Patient Examination (SP)

Description

Standardized patients (SPs) are well persons trained to simulate a medical condition in a standardized way or actual patients who are trained to present their condition in a standardized way. A standardized patient exam consists of multiple SPs each presenting a different condition in a 10-12 minute patient encounter. The resident being evaluated examines the SP as if (s)he were a real patient, (i.e., the resident might perform a history and physical exam, order tests, provide a diagnosis, develop a treatment plan, or counsel the patient). Using a checklist or a rating form, a physician observer or the SPs evaluate the resident's performance on appropriateness, correctness, and completeness of specific patient care tasks and expected behaviors (See description of Checklist Evaluation…). Performance criteria are set in advance. Alternatively or in addition to evaluation using a multiple SP exam, individual SPs can be used to assess specific patient care skills. SPs are also included as stations in Objective Structured Clinical Examinations (See description of OSCE).

Use
SPs have been used to assess history-taking skills, physical examination skills, communication skills, differential diagnosis, laboratory utilization, and treatment. Reproducible scores are more readily obtained for history-taking, physical examination, and communication skills. Standardized patient exams are most frequently used as summative performance exams for clinical skills. A single SP can assess targeted skills and knowledge.
Psychometric qualities
Standardized patient examinations can generate reliable scores for individual stations and total performance useful for pass-fail decisions. Training of raters whether physicians, patients or other types of observers is critical to obtain reliable scores. At least one-half day of testing time (four hours) is needed to obtain reliable scores for assessment of hands-on clinical skills. Research on the validity of some SP exams has found better performance by senior residents than junior residents (construct validity) and modest correlations between SP exam scores and clinical ratings or written exams (concurrent validity).
Feasibility / Practicality
Development of an examination using standardized patients involves identification of the specific competencies to be tested, training of standardized patients, development of checklists or rating forms and criteria setting. Development time can be considerable, but can be made more time efficient by sharing of SPs in a collaboration of multiple residency programs or in a single academic medical center. A new SP can learn to stimulate a new clinical problem in 8 to 10 hours; and an experienced SP can learn a new problem in 6 to 8 hours. About twice the training time is needed for SPs to learn to use checklists to evaluate resident performance. Facilities needed for the examination include an examining room for each SP station and space for residents to record medical notes between stations.

Suggested reference
Van der Vleuten, CPM and Swanson, D. Assessment of clinical skills with standardized patients: State of the art. Teach Learn Med. 1990; 2: 58-76.

Written Examination

Description

A written or computer-based MCQ examination is composed of multiple-choice questions (MCQ) selected to sample medical knowledge and understanding of a defined body of knowledge, not just factual or easily recalled information. Each question or test item contains an introductory statement followed by four or five options in outline format. The examinee selects one of the options as the presumed correct answer by marking the option on a coded answer sheet. Only one option is keyed as the correct response. The introductory statement often presents a patient case, clinical findings, or displays data graphically. A separate booklet can be used to display pictures, and other relevant clinical information. The in-training examinations prepared by specialty societies and boards use MCQ type test items. A typical half-day examination has 175 to 250 test questions.
In computer-based examinations the test items are displayed on a computer monitor one at a time with pictures and graphical images also displayed directly on the monitor. In a computer-adaptive test fewer test questions are needed because test items are selected based upon statistical rules programmed into the computer to quickly measure the examinee's ability.
Use

Medical knowledge and understanding can be measured by MCQ examinations. Comparing the test scores on in-training examinations with national statistics can serve to identify strengths and limitations of individual residents to help them improve. Comparing test results

aggregated for residents in each year of a program can be helpful to identify residency training experiences that might be improved.

Psychometric qualities

For test questions to be useful in evaluating a resident's knowledge each test item and the overall exam should be designed to rigorous psychometric standards. Psychometric qualities must be high for pass/fail decisions, but tests used to help residents identify strengths and weaknesses such as in-training examinations need not comply with the same rigorous standards. A committee of experts designing the test defines the knowledge to be assessed and creates a test blueprint that specifies the number of test questions to be selected for each topic. When test questions are used to make pass/fail decisions the test should be pilot tested and statistically analyzed. A higher reliability/reproducibility can be achieved with more test questions per topic. If pass/fail decisions will be made based on test scores a sufficient number of test questions should be included to obtain a test reliability greater than r = 0.85 (1.00 is perfect reliability). Standards for passing scores should be set by a committee of experts prior to administering the examination (criterion referenced exams). If performance of residents is to be compared from year to year at least 25 to 30 percent of the same test questions should be repeated each year.

Feasibility / Practicality

A committee of physician specialists develops the examination with assistance from psychometric experts. For in-training examinations each residency program administers an exam purchased from the specialty society or other vendor. Tests are scored by the vendor and scores returned to the residency director for each resident, for each topic, and by year of residency training. Comparable national scores also are provided. All the 24 ABMS Member Boards use MCQ examinations for initial certification.

Suggested references

Haladyna TM. Developing and validating multiple-choice test items. Hillsdale, New Jersey: L. Erlbaum Associates. 1994.

Case SM, Swanson DB. Constructing written test questions for the basic and clinical sciences. Philadelphia, PA: National Board of Medical Examiners, 1996 (www.nbme.org)

Multiple-choice Questions (MCQs)

With Multiple-choice Questions (MCQs), you first need to decide what you want to include on the test. The amount of attention given to evaluating something should reflect its relative importance. You need to sample topics and also sample skills (e.g., determining the diagnosis, deciding on the next step in management); you cannot ask everything. Performance on the sample provides a basis for estimating achievement in the broader domain that is actually of interest. The nature of the sample determines the extent to which the estimate of true ability is reproducible (reliable, generalizable) and accurate (valid). If the sample is not representative of the broader domain of interest (e.g., including only cardiovascular-related content in a test of competence in general medical practice), exam results will be biased and will not provide a good basis for estimating achievement in the domain of interest. If the sample is too small, exam results may not be sufficiently precise (reproducible, reliable) to ensure that they reflect true proficiency.

With a multiple-choice test, there's almost always one grader (usually the computer) and a series of questions or sets of questions; sampling involves selecting a subset of questions to include on the test. With other evaluation methods (e.g., oral exams based on patient cases, standardized patient exams, essay exams), the sampling is much more complicated. Any method that can't be scored mechanically requires sampling on a second dimension: the dimension of grader. In these exams, you are interested in performance across a range of cases and you want the grade to be independent of who the examiner is. You therefore need to sample across two dimensions: one for

the questions or cases and one for the judges or raters. You need to sample across a range of cases, because performance on one case is not a very good predictor of performance on other cases. You also need to sample across different raters to minimize the effects of rater harshness or leniency, and other issues like halo that cause problems in the consistency of scoring across raters. With broad samples, peaks and valleys in performance and peaks and valleys in rater differences tend to average out.

Item formats

The National Board of Medical Examiners in the US gives these examples for different item formats in its publication "Constructing written test questions for the basic and clinical sciences":

In the 40 years since the first MCQ exam, the National Board has broadened the scope of the A-type item to test reasoning and problem-solving skills by including a clinical vignette in most item stems. Today, the A-type remains the most commonly used item format on the Step examinations. Many other item formats that were developed during this period have been discontinued. These formats (named by a letter in order of origination) are described on the following pages.

A-Type:

Of the following, the most effective prophylactic agent for the prevention of recurrences of rheumatic fever is
A. acetylsalicylic acid
B. para-aminobenzoic acid
C. adrenocorticotrophic hormone
D. cortisone
E. sulfadiazine

B-Type:

B-type items were matching items that consisted of a list of lettered headings followed by a list of numbered words or phrases. The examinee was instructed to select the one heading that was most closely associated with each word or phrase.

Because each response could be used more than once or not at all, B-type items could not be solved by elimination. B-type items were believed to widen the scope of an MCQ examination by allowing testing of a number of related subjects in a single series of items. Unlike the matching formats used today, the B-type items did not typically include a lead-in; as a result, the question being asked was sometimes unclear. These items generally performed well, and were only discontinued recently as the extended-matching format became widely used.

DIRECTIONS: Each set of matching questions in this section consists of a list of three to five lettered options (some of which may be in figures) followed by several numbered items. For each numbered item, select the ONE lettered option that is most closely associated with it and fill in the circle containing the corresponding letter on the answer sheet. Each lettered option may be selected once, more than once, or not at all.

A. Coarctation of the aorta
B. Patent ductus arteriosus
C. Tetralogy of Fallot
D. Aortic vascular ring
E. Tricuspid atresia

1. Benefited by systemic-pulmonary artery anastomosis
2. Most common type of congenital cyanotic heart disease
3. Surgically corrected by resection and end-to-end anastomosis
4. Possible cause of dysphagia in infants and children

5. Hypertension in the arms and hypotension in the legs

D-Type:
D-type items were complex matching items in which each item consisted of three functional disturbances (designated by a letter) and five situations (in a numbered list). The examinee was instructed to 1) select the functional disturbance or category that four of the five situations were related to and 2) indicate the one situation that did not belong in that category. It was believed that these items required discriminatory understanding of a number of similar factors. However, D-type items were difficult to write, and the directions were confusing. In addition, they did not discriminate between knowledgeable and unknowledgeable examinees.

DIRECTIONS: There are two responses to be made to *each* of the following questions. In the lefthand list are three lettered categories. Exactly four of the five numbered items in the right-hand list are related in some way to ONE of these categories. (1) on the appropriate line in the answer sheet blacken the space under the letter of the category in which these four items belong. (2) Then blacken the space under the number of the item in the right-hand list that does NOT belong in the same category with the other four.

A. Eosinophilia of diagnostic significance	1. Trichinosis
B. Plasmacytosis of diagnostic significance	2. Multiple myeloma
C. Lymphocytosis of diagnostic significance	3. Loeffler's syndrome
	4. Hodgkin's disease
	5. Schistosomiasis

K-Type:
K-type items were the most commonly used multiple true/false item format at the National Board. They consisted of a stem followed by four options, one or more of which was correct. It was believed that K-type items tested in-depth knowledge or understanding of several aspects of a disease, a process, or a procedure, and required an examinee to be familiar with several different facts about a given topic. However, K-type items were criticized as being too complicated, requiring the examinee to constantly keep the answer code in mind. In addition, the possible response combinations introduced a cueing effect that reduced item discrimination and lowered test reliability. It was difficult to write good, unambiguous true/false items.
Because the items could include only absolutely true or false facts, K-type items could not be used to assess clinical judgement except in comparisons (e.g., "Drug X is better than Drug Y in treating disease K"). K-type items were more difficult and less discriminating than other item types. In addition, they were less efficient than other MCQ formats, and the relative reliability per unit of test time was lower.

Directions Summarized

A	**B**	**C**	**D**	**E**
1, 2, 3 only	1, 3 only	2, 4 only	4 only	All are correct

A child suffering from an acute exacerbation of rheumatic fever usually has
(1) an elevated sedimentation rate
(2) a prolonged PR interval
(3) an elevated antistreptolysin O titer
(4) subcutaneous nodules

C-Type:
C-type items were similar to B-type items in appearance but were multiple true/false in the task required of examinees. A C-type item consisted of a list of lettered headings followed by a list of

numbered words or phrases. For each numbered item, the examinees were required to decide if A was true, B was true, both were true (option C), or neither was true (option D). This item type was used to compare and contrast two diseases, signs and symptoms, laboratory findings, etc. C-type items match K-types in level of difficulty. The primary problem with C-types was in deciding to what extent something had to be "true" to be selected. If, for example, something was associated with both A and B, but was more strongly associated with A, the examinee had to decide whether an appropriate response was A only or Both A and B. With relatively weak associations, the examinee had to decide whether the association was strong enough to note, or whether "neither" was the appropriate response. These judgements were not related to medical knowledge, but rather forced the examinee to think about what the item writers intended.

DIRECTIONS: Each set of matching questions in this section consists of a list of four lettered options followed by several numbered items. For each numbered item, select the ONE lettered option that is most closely associated with it and fill in the circle containing the corresponding letter on the answer sheet. Each lettered option may be selected once, more than once, or not at all.

A. *Plasmodium vivax* malaria
B. *Plasmodium falciparum* malaria
C. Both
D. Neither

1. A combination of primaquine and chloroquine is treatment of choice for acute attack.
2. Clinical attacks suppressed by ingestion of chloroquine once a week while in an endemic area.
3. Permanently cured by treatment with chloroquine.
4. Infection prevented by ingestion of chloroquine once a week.

<u>E-Type:</u>
E-type items were multiple true/false items that are based on the analysis of relationships. Examinees who took E-type items still refer to them as the "True, True and Unrelated" items. The E-type consisted of a sentence with two main parts: an assertion and a reason for that assertion. The examinee was directed to select A if both were true statements and the reason was a correct explanation of the assertion; B if both were true statements but the reason was not a correct explanation of the assertion; C if the assertion was true but the reason was a false statement; D if the assertion was false but the reason was a true statement; E if both assertion and reason were false statements. It was thought that good reasoning skills and an understanding of the basic principles were required to answer this item type correctly. However, E-type items were difficult to construct, and examinees found them to be confusing.

Directions Summarized			
A	True	True	Reason is a correct explanation.
B	True	True	Reason is NOT a correct explanation.
C	True	False	
D	False	True	
E	False	False	

Assertion		**Reason**
Herpes simplex is usually regarded as an autogenous infection	BECAUSE	patients given fever therapy frequently develop herpes..

Cow's milk is preferable to breast milk in infant feeding	BECAUSE	cow's milk has a higher content of calcium.

H-Type:

H-type items were comparison items that consisted of paired statements describing two entities to be compared in a quantitative sense. The examinee was directed to select A if A was greater than B; B if B was greater than A; and C if the two were approximately equal.

Although it was generally agreed that questions that depend on the memorization of absolute quantitative amounts should be limited, the H-type item was believed to be useful for those instances where recall of quantitative information was believed to be important. The problem for the examinees was in deciding how great the difference needed to be in order to be relevant.

DIRECTIONS: The following paired statements describe two entities that are to be compared in a quantitative sense. On the appropriate line of the answer sheet blacken the space under
A if (A) is greater than (B),
B if (B) is greater than (A),
C if the two are equal or very nearly equal.

1. (A) The usual therapeutic dose of epinephrine
   (B) The usual therapeutic dose of ephedrine

2. (A) Life expectancy with glioblastoma of the occipital lobe
   (B) Life expectancy with glioblastoma of the frontal lobe

I-Type:

The I-type item was similar to the H-type. It consisted of pairs of phrases that describe conditions or quantities that might vary in relation to each other. The examinee was directed to select A if the two phrases were related directly (i.e., an increase in the first was accompanied by an increase in the second or a decrease in the first was accompanied by a decrease in the second); B if the phrases were related inversely (i.e., an increase in the first was accompanied by a decrease in the second or a decrease in the first was accompanied by an increase in the second); or C if the changes were independent of one another.

DIRECTIONS: Each of the following pairs of phrases describe conditions or quantities that may or may not be related. On the appropriate line of the answer sheet blacken the space under
A if increase in the first is accompanied by increase in the second or if decrease in the first is accompanied by decrease in the second
B if increase in the first is accompanied by decrease in the second or if decrease in the first is accompanied by increase in the second
C if changes in the first are not necessarily accompanied by changes in the second.

1. (A) Urine volume
   (B) Urine specific gravity

2. (A) Plasma protein concentration
   (B) Colloid osmotic pressure of plasma

Neither the H- nor I-type formats were particularly popular. Because there were fewer options than in other item types, there was an increased chance of guessing the correct answer. In addition, the items tended to focus on minor details rather than scientific concepts.

In his series in the Federation Bulletin, Morton (1985-86) implied that different item types were included on medical licensure examinations simply to add variety to a lengthy examination. But, 25 years after the National Board converted from an essay exam to MCQ exams, the NBME reviewed the research on the various types of MCQs used, and the variety of item types was then reduced to include A-, B-, C-, G-, K-, X-, and M-type items. Staff again reviewed item types in the mid-1980s. The general consensus, at that time, was that four basic item types provided sufficient variety to test the knowledge specified as important for the awarding of a National Board certificate. These four basic types included A-, B-, C- and K-type items. G-types (sets of A-type items), N-types (sets of K-type items), and M-type items were no longer considered as separate formats.
More recently, the variety of item types has again been reviewed. The current Step examinations include A- and R-type items only. Some of the steps taken to improve the examinations include: concentrating on item types that are psychometrically sound, educating item writers on various item-writing techniques, focusing on clinical decision-making rather than recall items, and pretesting newly written items.

Technical item flaws

The National Board of Medical Examiners in the US gives examples for technical item flaws in its publication "Constructing written test questions for the basic and clinical sciences":

This section describes two types of technical item flaws: testwiseness and irrelevant difficulty. Flaws related to testwiseness make it easier for some students to answer the question correctly, based on their test-taking skills alone. These flaws commonly occur in items that are unfocused and do not satisfy the "cover-the-options" rule. Flaws related to irrelevant difficulty make the question difficult for reasons unrelated to the trait that is the focus of assessment.
The purpose of this section is to outline common flaws and to encourage you to eliminate these flaws from your questions to provide a level playing field for the testwise and not-so-testwise students. The probability of answering a question correctly should relate to the examinee's amount of expertise on the topic being assessed and should not relate to their expertise on test-taking strategies.

Issues related to testwiseness

Grammatical cues: one or more distractors don't follow grammatically from the stem Because an item writer tends to pay more attention to the correct answer than to the distractors, grammatical errors are more likely to occur in the distractors. In this example, testwise students would eliminate A and C as options because they do not follow grammatically or logically from the stem.
Testwise students then have to choose only between B, D, and E.

A 60-year-old man is brought to the emergency department by the police, who found him lying unconscious on the sidewalk. After ascertaining that the airway is open, the first step in management should beintravenous administration of
A. examination of cerebrospinal fluid
B. glucose with vitamin B1 (thiamine)
C. CT scan of the head
D. phenytoin
E. diazepam

Crime is

Logical cues: a subset of the options are collectively exhaustive In this item, Options A, B, and C include all possibilities. The testwise student knows that A, B, or C must be correct, whereas the non-testwise student spends time considering D and E. Often, the item writers add D and E only because they want to list five options. In these situations, the item writer may not have paid much attention to the merits of options D and E; sometimes, they are partially correct and confusing because they cannot be rank-ordered on the same dimension as Options A, B, and C. This flaw is commonly seen in items with options such as "Increases," "Decreases," and "Remains the same."

A. equally distributed among the social classes
B. overrepresented among the poor
C. overrepresented among the middle class and rich
D. primarily an indication of psychosexual Maladjustment
E. reaching a plateau of tolerability for the nation

Absolute terms: terms such as "always" or "never" are used in options In this item, Options A, B, and E contain terms that are less absolute than those in Options C and D. The testwise student will eliminate Options C and D as possibilities because they are less likely to be true than something stated less absolutely. Note that this flaw would not arise if the stem was focused and the options were short; it arises only when verbs are included in the options rather than in the lead-in.

In patients with advanced dementia, Alzheimer's type,the memory defect
A. can be treated adequately with phosphatidylcholine(lecithin)
B. could be a sequela of early parkinsonism
C. is never seen in patients with neurofibrillary tangles at autopsy
D. is never severe
E. possibly involves the cholinergic system

Long correct answer: correct answer is longer, more specific, or more complete than other options

In this item, Option C is longer than the other options; it is also the only double option. Item writers tend to pay more attention to the correct answer than to the distractors. Because you are teachers, you write long correct answers that include additional instructional material, parenthetical information, caveats, etc. Sometimes this can be quite extreme: the correct answer is a paragraph in length and the distractors are single words.

Secondary gain is
A. synonymous with malingering
B. a frequent problem in obsessive-compulsive disorder
C. a complication of a variety of illnesses and tends to prolong many of them
D. never seen in organic brain damage

Word repeats: a word or phrase is included in the stem and in the correct answer This item uses the word "unreal" in the stem, and "derealization" is the correct answer. Sometimes, a word is repeated only in a metaphorical sense, e.g., a stem mentioning bone pain, with the correct answer beginning with the prefix "osteo-".

A 58-year-old man with a history of heavy alcohol use and previous psychiatric hospitalization is confused and agitated. He speaks of experiencing the world as unreal. This symptom is called
A. depersonalization
B. derailment
C. derealization
D. focal memory deficit
E. signal anxiety

Local anesthetics are most effective in the

A. anionic form, acting from inside the nerve membrane
B. cationic form, acting from inside the nerve membrane
C. cationic form, acting from outside the nerve membrane
D. uncharged form, acting from inside the nerve membrane
E. uncharged form, acting from outside the nerve membrane

Convergence strategy: the correct answer includes the most elements in common with the other options This item flaw is less obvious than the others, but it occurs frequently and is worth noting. The flaw is seen in several forms. The underlying premise is that the correct answer is the option that has the most in common with the other options; it is not likely to be an outlier. For example, in numeric options, the correct answer is more often the middle number than an extreme value. In double options, the correct answer is more likely to be the option that has the most elements in common with the other distractors. For example, if the options are "Pencil and pen"; "Pencil and highlighter"; "Pencil and crayon"; "Pen and marker," the correct answer is likely to be "Pencil and pen" (ie, by simple count, "Pencil" appeared 3 times in the options; "Pen" appeared twice; other elements each appeared only once). While this might seem ridiculous, this flaw occurs because item writers start with the correct answer and write permutations of the correct answer as the distractors. The correct answer is, therefore, more likely to have elements in common with the rest of the options; the incorrect answers are more likely to be outliers as the item writer has difficulty generating viable distractors. In this example, the testwise student would eliminate "anionic form" as unlikely because "anionic form" appears only once; that student would also exclude "outside the nerve membrane" because "outside" appears less frequently than "inside". The student would then have to decide between Options B and D. Since three of the five options involve a charge, the testwise student would then pick Option B.

Issues related to irrelevant difficulty

Options are long, complicated, or double

This item illustrates a common flaw. The stem contains extraneous reading, but, more importantly, the options are very long and complicated. Trying to decide among these options requires a significant amount of reading because of the number of elements in each option. This can shift what is measured by an item from content knowledge to reading speed. Please note that this flaw relates only to options. There are many well-constructed test questions that include a long stem. Decisions about stem length should be made in accord with the purpose of the item. If the purpose of the item is to assess whether or not the student can interpret and synthesize information to determine, for example, the most likely diagnosis, then it is appropriate for the stem to include a fairly complete description of the situation.

Peer review committees in HMOs may move to take action against a physician's credentials to care for participants of the HMO. There is an associated requirement to assure that the physician receives due process in the course of these activities. Due process must include which of the following?
A. Notice, an impartial forum, council, a chance to hear and confront evidence against him/her.
B. Proper notice, a tribunal empowered to make the decision, a chance to confront witnesses against him/her, and a chance to present evidence in defense.
C. Reasonable and timely notice, impartial panel empowered to make a decision, a chance to hear evidence against himself/herself and to confront witnesses, and the ability to present evidence in defense.

Numeric data are not stated consistently

Following a second episode of infection, what is the likelihood that a woman is infertile?
A. Less than 20%
B. 20 to 30%
C. Greater than 50%
D. 90%

When numeric options are used, the options should be listed in numeric order and the options should be listed in a single format (i.e., as single terms or as ranges). Confusion occurs when formats are mixed and when the options are listed in an illogical order or in an inconsistent format.

E. 75%

In this example, Options A, B, and C are expressed as ranges, whereas Options D and E are specific percentages. All options should be expressed as ranges or as specific percentages; mixing them is ill-advised. In addition, the range for Option C includes Options D and E, which almost certainly rules out Options D and E as correct answers.

Frequency terms in the options are vague (e.g., rarely, usually)

Research has shown that vague frequency terms are not consistently defined or interpreted, even by experts.

A more complete discussion of this research is included later on.

Severe obesity in early adolescence
A. usually responds dramatically to dietary regimens
B. often is related to endocrine disorders
C. has a 75% chance of clearing spontaneously
D. shows a poor prognosis
E. usually responds to pharmacotherapy and intensive psychotherapy

Language in the options is not parallel; options are in an illogical order

This item illustrates a common flaw in which the options are long and the language makes it difficult and time-consuming to determine which is the most correct. Generally, this flaw can be corrected by careful editing. In this particular item, the lead-in can be changed to "For which of the following reasons can no conclusion be drawn from these results?" The options can then be edited (i.e., A. No follow-up was made of non-vaccinated children; B. The number of cases was too small; C. The trial involved only boys, and a new option can be written for D).

In a vaccine trial, 200 2-year-old boys were given a vaccine against a certain disease and then monitored for five years for occurrence of the disease. Of this group, 85% never contracted the disease. Which of the following statements concerning these results is correct?
A. No conclusion can be drawn, since no follow-up was made of non-vaccinated children
B. The number of cases (i.e., 30 cases over five years) is too small for statistically meaningful conclusions
C. No conclusions can be drawn because the trial involved only boys
D. Vaccine efficacy (%) is calculated as 85-15/100

"None of the above" is used as an option

The phrase "None of the above" is problematic in items where judgement is involved and where the options are not absolutely true or false. If the correct response is intended to be one of the other listed options, knowledgeable students can be faced with a dilemma because they have to decide between a very detailed perfect option and the one that you have intended as correct. They can often construct an option that is more correct than the one you intended to be correct. Use of "none of the above" essentially turns the item into a true/false item; each option has to be evaluated as more or less true than the universe of unlisted options. It will often be possible to fix such items by replacing "none of the above" by an option that means roughly the same thing but is more specific. For example, in an item asking an examinee to specify the most appropriate pharmacotherapy, replacing "none of the above" by "no drug should be given at this time" will eliminate the ambiguity of "none of the above."

Which city is closest to New York City?
A. Boston
B. Chicago
C. Dallas
D. Los Angeles
E. none of the above
If students select E, you don't know if they are thinking about Philadelphia or London.

Stems are tricky or unnecessarily complicated

Sometimes, item writers can take a perfectly easy question and turn it into something so convoluted that only the most stalwart will even read it. This item is a sample of that kind of item. The notation in I: through V: is complex; having to rank order Roman numerals after working through that notation is irrelevant and unnecessarily difficult.

Arrange the parents of the following children with Down's syndrome in order of highest to lowest risk of recurrence. Assume that the maternal age in all cases is 22 years and that a subsequent pregnancy occurs within 5 years. The karyotypes of the daughters are:
I: 46, XX, -14, +T (14q21q) pat *II: 46, XX, -14, +T (14q21q) de novo* *III: 46, XX, -14, +T (14q21q) mat* *IV: 46, XX, -21, +T (14q21q) pat* *V: 47, XX, -21, +T (21q21q) (parents not karyotyped)*
A. III, IV, I, V, II
B. IV, III, V, I, II
C. III, I, IV, V, II
D. IV, III, I, V, II
E. III, IV, I, II, V

Summery of technical item flaws

Issues Related to Testwiseness

- Grammatical cues - one or more distractors don't follow grammatically from the stem
- Logical cues - a subset of the options is collectively exhaustive
- Absolute terms - terms such as "always" or "never" are in some options
- Long correct answer - correct answer is longer, more specific, or more complete than other options
- Word repeats - a word or phrase is included in the stem and in the correct answer
- Convergence strategy - the correct answer includes the most elements in common with the other options

Issues Related to Irrelevant Difficulty

- Options are long, complicated, or double
- Numeric data are not stated consistently
- Terms in the options are vague (e.g., "rarely," "usually")
- Language in the options is not parallel
- Options are in a nonlogical order
- "None of the above" is used as an option
- Stems are tricky or unnecessarily complicated
- The answer to an item is "hinged" to the answer of a related item

General Guidelines for Item Construction

- Make sure the item can be answered without looking at the options OR that the options are 100% true or false.
- Include as much of the item as possible in the stem; the stems should be long and the options short.
- Avoid superfluous information.
- Avoid "tricky" and overly complex items.
- Write options that are grammatically consistent and logically compatible with the stem; list them in logical or alphabetical order. Write distractors that are plausible and the same relative length as the answer.
- Avoid using absolutes such as always, never, and all in the options; also avoid using vague terms such as usually and frequently.

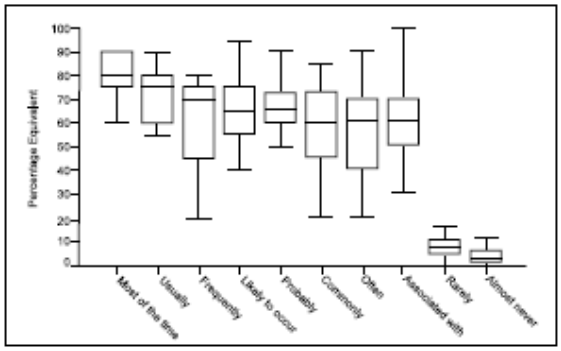
- Avoid negatively phrased items (e.g., those with except or not in the lead-in). If you must use a negative stem, use only short (preferably single word) options.

And most important of all: Focus on important concepts; don't waste time testing trivial facts.

Use of imprecise terms in examination questions

While imprecise terms are used in our everyday speech and in our writing, these terms cause confusion when they are used in the text of examination items. In a study conducted at the NBME, 60 members of eight test committees who wrote questions for various medical specialty examinations reviewed a list of terms used in MCQs to express some concept related to frequency of occurrence and indicated the percentage of time that was reflected by each term.

Results (shown in the box-plot) indicated that the terms do not have an operational definition that is commonly shared, even among the item writers themselves. The mean value plus or minus one standard deviation exceeded 50 percentage points for more than half of the phrases. For example, on average, the item writers believed the term *frequently* indicated 70% of the time; half believed it was between 45% and 75% of the time; actual responses ranged from 20% to 80%. Of particular note is that values for *frequently* overlapped with values for *rarely.*



Box-plot showing distribution of responses for frequency terms.

The implication of these results for the construction of test questions varies by item format. Vague terms create far more severe problems in the various kinds of true/false items (K-, C- and X-type items) than in one-best-answer (A- and R-type) items. For example, imprecise terms cause major problems in true/false items such as this example:

True statements about pseudogout include:
1. It occurs commonly in women.
2. It is often associated with acute pain.
3. It is usually hereditary.
4. Serum calcium levels are frequently increased.

In true/false items, the examinee has to judge whether each option is true or false. When options are not absolutely true or false, examinees rely on their personal definition of the ambiguous terms or their guesses about what these terms meant to the item writer. Alternatively, examinee responses may reflect personal response style (the tendency to respond either true or false when the correct answer is unknown). These response style factors may have more of an effect on whether or not an examinee answers the item correctly than knowledge of the subject matter and may be part of the reason why true/false items tend to perform poorly.

Rewording the options by specifying exact numbers does not correct the problem. For example, the statement, "the incidence among women is 1:2000" would not be an appropriate modification of Option 1 in the example shown. The incidence is not exactly 1:2000, and because a band is not specified, examinees would define their own bands, narrowly or widely, presumably depending on personal response styles. In true/false items, the appropriate treatment of numeric options is either to generate a comparison (e.g., the incidence is greater than that of osteoarthritis) or to specify a range (e.g., the incidence is between 1:1000 and 1:2000).

The issue noted above with true/false items is not as problematic with well-constructed one-best-answer items (i.e., those that pose a clear question and have homogeneous options). For example, the following question includes a vague term in the item stem, yet, because the task is to select the one-best answer, the question is relatively unambiguous.
Which of the following laboratory values is usually increased in patients with pseudogout?
Problems do arise with one-best answer items that have vague terms in the options as in this example:

Patients with pseudogout have pain:
A. frequently
B. usually
C. often
D. commonly

The only way to make such an item more ambiguous would be to use a fifth option "none of the above."

Results are based on responses from 60 members of eight item-writing committees. The horizontal line in each box indicates the median response; the boxes include the ranges for 50% of the responses. The vertical lines extend to the highest and lowest values indicated. For example, the median response for "frequently" indicated 70% of the time; half believed it was between 45% and 75% of the time; actual responses ranged from 20% to 80%, almost overlapping with "rarely."

From: Case SM. (1994) The use of imprecise terms in examination questions: How frequent is frequently? *Academic Medicine*, 69(suppl):S4-S6.

Basic rules for one-best-answer items

The National Board of Medical Examiners in the US has set some basic rules for one-best-answer items (like MCQs) in its publication "Constructing written test questions for the basic and clinical sciences":

- *Each item should focus on an important concept, typically a common or potentially catastrophic clinical problem.*
  Don't waste testing time with questions assessing knowledge of trivial facts. Focus on problems that would be encountered in real life. Avoid trivial, "tricky," or overly complex questions.
- *Each item should assess application of knowledge, not recall of an isolated fact.*
  The item stems may be relatively long; the options should be short. Clinical vignettes provide a good basis for a question. For the clinical sciences, each should begin with the presenting problem of a patient, followed by the history (including duration of signs and symptoms), physical findings, results of diagnostic studies, initial treatment, subsequent findings, etc. Vignettes may include only a subset of this information, but the information should be provided in this specified order. For the basic sciences, patient vignettes may be very brief; "laboratory vignettes" are also appropriate.
- *The stem of the item must pose a clear question, and it should be possible to arrive at an answer with the options covered.*
  To determine if the question is focused, cover up the options and see if the question is clear and if the examinees can pose an answer based only on the stem. Rewrite the stem and/or options if they could not.
- *All distractors (i.e., incorrect options) should be homogeneous.*
  They should fall into the same category as the correct answer (e.g., all diagnoses, tests, treatments, prognoses, disposition alternatives). Rewrite any dissimilar distractors. Avoid

using "double options" (e.g., do W and X; do Y because of Z) unless the correct answer and all distractors are double options. Rewrite double options to focus on a single point. All distractors should be plausible, grammatically consistent, logically compatible, and of the same (relative) length as the correct answer. Order the options in logical order (e.g., numeric), or in alphabetical order.

- *Avoid technical item flaws that provide special benefit to testwise examinees or that pose irrelevant difficulty.*

Do **NOT** write any questions of the form "Which of the following statements is correct?" or "Each of the following statements is correct EXCEPT." These questions are unfocused and have heterogeneous options.

Subject each question to the five "tests" implied by the above rules. If a question passes all five, it is probably well-phrased and focused on an appropriate topic.

Writing one-best-answer items

The National Board of Medical Examiners in the US gives these suggestions to take into consideration when writing one-best-answer items (like MCQs) in its publication "Constructing written test questions for the basic and clinical sciences":

Constructing the stem

The vast majority of questions should be written with a clinical vignette. The stem should begin with the presenting problem of a patient, followed by the history (including duration of signs and symptoms), physical findings, results of diagnostic studies, initial treatment, subsequent findings, etc. Vignettes may include only a subset of this information, but the information should be provided in this specified order. The stem should consist of a single, clearly formulated problem. The lead-in of the stem must pose a clear question so that the examinee can pose an answer without looking at the options. Satisfying the "cover-the-options" rule is an essential component of a good question.

Good stem:

This stem provides sufficient information and can be answered without referring to the options.

A 52-year-old man has had increasing dyspnea and cough productive of purulent sputum for 2 days. He has smoked one pack of cigarettes daily for 30 years. His temperature is 37.2 C (99 F). Breath sounds are distant with a few rhonchi and wheezes. His leukocyte count is 9000/mm3 with a normal differential. Gram's stain of sputum shows numerous neutrophils and gram-negative diplococci. X-ray films of the chest show hyperinflation. Which of the following is the most likely diagnosis?

Stem testing isolated facts:

The following stem contains insufficient information; in order to answer the question, the examinee must use the options as a frame of reference.

Which of the following is true about pseudogout?

Patient vignettes should include some or all of the following components in the order indicated:

- **Age, Gender** (e.g., A 45-year-old man)
- **Site of Care** (e.g., comes to the emergency department)
- **Presenting Complaint** (e.g., because of a headache)
- **Duration** (e.g., that has continued for 2 days).
- **Patient History (with Family History ?)**
- **Physical Findings**
- **+/- Results of Diagnostic Studies**
- **+/- Initial Treatment, Subsequent Findings, etc.**

Make sure that your stem:

- Focuses on important concepts rather than trivial facts
- Can be answered without looking at the options
- Includes all relevant facts; no additional data should be provided in the options
- Is not "tricky" or overly complex
- Is not negatively phrased (i.e., avoid using *except* or *not* in the lead-in)

Fine points on item stem

Use of Real Patients.

We believe it is generally better not to base multiple-choice questions on "real patients," particularly for tests aimed at students. As a general rule, real patients are too complicated, and the elements that are complicated are not necessarily those that are important for assessment. As noted earlier, we do include window dressing (i.e., incidental findings), but do not include "red herrings" (i.e., information that is intended to lead examinees away from the correct answer). Unfortunately, real patients often have "red herrings" among their findings.

Use of Reference Materials.

We believe that it is appropriate to provide information in a test question if, in real life, someone would be likely to refer to a reference source to obtain the information. For example, in many instances, we believe it is appropriate to provide a table of normal laboratory values or a chart showing a recommended schedule of screening tests or immunizations. Of course, you might not just ask questions that require examinees to simply look up information in the chart provided, but you might, for example, ask about immunization of a 6-year-old child who had never been immunized.

Use of Patient's or Physician's Own Words.

We generally do not believe it is useful to include the patient's own words, particularly if the examinee task is to interpret nuances of language that might be affected by tone. On the other hand, it may be useful to ask the examinee to select the most appropriate physician response to a patient by asking the examinee to choose among options phrased as open-ended, closed, or leading questions.

Patients Who Lie.

We believe all multiple-choice patients should tell the truth, or the physician's interpretation of the patient's story should be provided. Physicians use multiple cues to determine how truthful a patient is and many of these cues cannot be translated into written form. Thus, our items would describe a patient's alcohol consumption as "The patient drinks 16 oz of beer with dinner each night" or "The patient's description of his alcohol consumption is contradictory." We would not write something ambiguous, such as "The patient 'claims' to drink only one bottle of beer each night."

Verbosity, Window Dressing, and Red Herrings: Do They Make a Better Test Item?

Most educators stress the importance of writing item stems that are as short as possible, avoiding verbosity (i.e., extra words), "window dressing" (i.e., extraneous material), and "red herrings" (i.e., information designed to mislead the examinee). Somewhat in opposition to this advice, we have emphasized use of clinical vignettes in item writing efforts. For USMLE Step 2, these vignettes consist of paragraph-length descriptions of clinical situations, generally followed by a question related to the diagnosis or next step in patient care. Such items stress application of knowledge by asking examinees to make clinical decisions, rather than to simply recall isolated facts. They are

designed to reflect “real life tasks” by challenging examinees to first identify the findings that are important, then integrate those findings into a diagnosis or clinical action. Such items often require multiple steps in the thinking process. We have found that vignette items tend to have fewer technical item flaws than typical non-vignette items, presumably because vignettes follow a standard structure and pose questions that are clinically natural.

Despite these advantages, some have challenged the use of vignettes, believing that a vignette only makes an item more cumbersome by wrapping window dressing around the real question. Some advocate avoiding vignettes altogether; others advocate using short vignettes and including only relevant positive findings expressed concisely; the rest advocate use of long vignettes that include more complete information that the examinee must review and synthesize.

Several studies were conducted to compare the psychometric characteristics of items developed in three formats: non-vignette, short vignette, and long vignette. The progression was designed to require increasing levels of interpretation, analysis, and synthesis of findings (see sample item in three formats below). As expected, items became more difficult as patient findings were presented in a less interpreted form; however, the differences in discrimination were not statistically significant.

Regardless of the mixed psychometric results, we believe vignette items are generally more appropriate because they test application of knowledge to patient situations and pose appropriate clinical challenges; such items might be viewed as “low fidelity” clinical simulations that improve the content validity of the examinations.
An item written in a non-vignette format typically is written from a “top-down” perspective (ie, given a disease, what are the associated findings). To an expert, items written in this manner may appear identical to items written with a patient vignette.

European Core Curriculum – *the Students' Perspective*

Note: You can find the original document with all the references and a list of the participants of the conference in the SCOME-wikipedia!

At the 5th Bologna follow-up conference in July 2006 in Bristol (UK) more than 40 medical student representatives from 15 European countries met to discuss a common outcome-based European core curriculum from the students' perspective.
In 4 days of hard work they managed to write and reach consensus on the **"European Core Curriculum – *the Students' Perspective*"** which will serve medical students' as a framework for national or local outcome-based core curricula:

Executive summary

From July 6th-10th, 2006, the 5th Bologna process follow-up conference organised by the European Medical Students' Association (EMSA) and the International Federation of Medical Students' Associations (IFMSA) took place in Bristol (UK).

More than 40 student representatives from 15 countries agreed on an outcome-based European core curriculum from the students' perspective. The "European Core Curriculum – *the Students' Perspective*" expresses the medical students' opinion on which abilities, knowledge, and attitudes graduates of medical schools in Europe should have gained and be assessed in accordingly.

Over the last few years, in innovative medical education, focus has shifted from acquisition of knowledge towards the achievement of concrete learning outcomes. Society and stakeholders are now more interested in the final product of the educational programme rather than the processes used to reach them. Therefore the core curriculum does not prescribe neither teaching nor assessment methods to be used but only the final product of the educational process.

The core curriculum is structured in 9 domains with 76 learning outcomes which are listed in alphabetical order:

- Clinical Skills
- Communication
- Critical Thinking
- Health in Society
- Life long learning
- Professionalism – Attitudes, responsibilities and self development
- Teaching
- Teamwork
- Theoretical knowledge

The curriculum will serve medical students and all other stakeholders in medical education as a common framework which can easily be adjusted for specific national or local needs. It serves as a common basis aiming to maintain and even improve the quality of education, healthcare and mobility, therefore furthering the establishment of a European Higher Education Area.

Preamble

Since 2003, the European Medical Students' Association (EMSA) and the International Federation of Medical Students' Association (IFMSA) have collaborated in developing the European Higher Education Area in the field of medicine resulting in widely recognised position papers. The 5th Bologna follow-up conference in Bristol (UK) hosted more than 40 medical student representatives from 15 European countries discussing a **„European Core Curriculum for Medicine – *the Students' Perspective*"**.

Participants of the above mentioned conference agreed on the development of an outcome-based core curriculum designed to fit society's need for optimal patient care and safety. The medical profession differs from other professions in that it is the outcome of education rather than the educational process that has a significant bearing on public health. The focus in medical education has for too long been based on the educational process instead of the product of education. Consequentially, as medical students of Europe we embrace the challenge of working with Medical Schools to take more responsibility for the final product of education instead of focusing on providing knowledge in excess of the core abilities gained by each graduate.

The field of medicine is rapidly expanding; advancing research and technology have extended our core knowledge necessitating a dynamic and modern curriculum to serve new demands. This focused education will empower graduates to serve their population with the most accurate and relevant knowledge and abilities. All stakeholders in medical education should increase communication to develop these curricula, and associated appropriate methods of assessment, optimising the outcome of medical education and the consequential standard of the medical profession.

We aimed neither to reinvent the wheel nor neglect the existence of established and elaborated core curricula. Whilst these documents have paved the way in outcome-based initiatives and we have incorporated some of their key ideas, we wanted to express the opinion of European medical students. As a diverse group currently experiencing a broad base of undergraduate medical education with equally varied educational techniques, we are in an optimal position to propose a curriculum suited to modern healthcare needs. We suggest the use of this curriculum as a framework which could easily be adopted and adjusted for national and local needs.

In developing a core curriculum, harmony and subsequent mobility will be increased throughout the European Higher Education Area. While the core values remain constant throughout Europe, we embrace the individuality and diversity of the countries, regions and individual institutions. This is reflected in the nature of an outcome-based curriculum, not prescribing the educational approach which leads to the end-point, but the overall outcome.

Whilst we believe that the current course of medical studies should lead to a common European medical degree, with specialisation occurring at a post-graduate level, the opportunity to tune individual interests and abilities at the undergraduate level is an important one. Thus, we welcome the opportunity for faculties and their students to foster a unique profile through educational opportunities and programmes. This will facilitate a culture of diversity and increased evolution of the field.

This document is a demonstration of the hard-work and dedication of European medical students to facilitate change and contribute to improved patient care and safety in our future work as medical professionals. We acknowledge our responsibilities and are prepared for the challenges associated with being at the forefront of reform. However, we are only one stakeholder in the field of medical education and therefore present our opinion as a basis for further work and co-operation. This will create a motivational environment for learning leading to further excellence in healthcare.

Note:
The domains of the "European Core Curriculum – *the Students' Perspective*" are listed in alphabetical order.

Clinical Skills

Graduates should have acquired and mastered clinical skills and practical procedures in order to confidently perform them in the professional environment. We appreciate the need for a specific list determining the skills and procedures. Whilst this is beyond the scope of this document, we acknowledge those outlined in previous documents.

Basic diagnostic tools

- Graduates should be able to take a detailed and relevant history.
- Graduates should be able to perform both general and targeted physical examination.
- Graduates should be able to utilise diagnostic procedures, imaging techniques and laboratory (paraclinical) tests where appropriate and interpret results adequately.

Clinical reasoning

- Graduates should be able to demonstrate sufficient clinical reasoning to enable them to use the basic diagnostic tools to arrive at a diagnosis and management plan in light of all the acquired information.

Treatment and care

- Graduates should be able to formulate and carry out an appropriate management plan.
- Graduates should be able to recognize and manage emergency conditions.
  - Graduates should be able to administer advanced life support as defined by international guidelines.
- Graduates should be able to apply appropriate palliative care.

Clinical record keeping

- Graduates should be competent in maintaining clinically and legally valid patient records which are easily readable.

Patient-centred approach

- Graduates should be able to consider the patient as a whole taking into consideration his social and psychological background.

- Graduates should be able to take into account the patient's understanding and experience of their condition and treatment.

- Graduates should be able to adapt treatment to the particular patient, evaluating both effectiveness and evidence.

Communication

Graduates should have the communication skills that facilitate the practise of acquired competencies. This is vital to excellence in patient care.

- Graduates should be able to communicate effectively and efficiently with all relevant parties in the medical environment. This includes:
  - Appropriate communication in every situation using different communication tactics
  - Awareness of their own and others' non-verbal communication.
  - Effective communication with patients, regardless of their backgrounds and/or disabilities.
  - The ability to effectively explain medical issues to a patient
  - Effective communication with other healthcare workers
  - The ability to communicate with all organisations that serve the public

- Graduates should show respect, openness and honesty with patients and aim to communicate with empathy and intuition.

- Graduates should put all their efforts in creating an atmosphere of confidentiality.

- Graduates should find a way to communicate, even when there are barriers to the communication.
  - Graduates should be able to use interpreters and be aware of the difficulties concerning this type of communication.

- Graduates should be able to communicate through all common modalities, including verbal, non-verbal, oral and written communication.

- Graduates should be able to give and receive feedback.

Critical Thinking

Critical thinking is the systematic evaluation of information preceding any professional decision and action. We emphasize that this skill is integral to all aspects of the doctor's role.

- Graduates should be able to question medical procedures and treatment protocols before their application.

- Graduates should be able to find the evidence base for clinical decisions.

- Graduates should stay up-to-date with recent scientific developments and implement evidence based medicine in daily practice. This includes:
  - The ability to evaluate relevant scientific texts and learning resources.

  - An awareness of the limitations of current medical knowledge.

- Graduates should be able to apply quality assurance methods in professional practice.

- Graduates should be able to effectively and critically use resources in professional practice.

Health in Society

As future doctors in a rapidly changing environment we are obliged to adjust our attitudes to the expectations of society. We consider knowledge of the basic principles of public health issues as essential for our work as future physicians at a local, national and international level. Therefore, we stress the importance of including environmental, cultural and international health related issues in our medical curriculum.

Environmental issues

- Graduates should know the impact of social, political and economic factors on the health of individuals and the community.

- Graduates should know the key risk factors, strategies for prevention and screening programmes for the most common conditions.

- Graduates should be able to identify vulnerable populations and respond appropriately.

- Graduates should be able to promote health in individual patients and in society. This includes:
  - Active education of patients.
  - The ability to identify health hazards in the environment and use the existing protocols to notify the responsible authorities accordingly.
  - The ability to formulate their opinion on these issues and participate actively in shaping health policies.

Ethnicity and Cultural issues

- Graduates are able to work with patients from different cultures, religions, social and ethnic backgrounds. This includes:
  - Approaching all patients with equality, regardless of their background.
  - Effective communication with patients, regardless of their background.

- Graduates can identify specific ethnic and social groups susceptible to certain conditions.

- Graduates understand the impact of cultural, religious and social aspects on health, health behaviour and the treatment process.

International Health issues

- Graduates should be familiar with the structure of European and international health politics and all its stakeholders.

- Graduates should be aware of the existence of epidemics and infectious diseases worldwide and know their prevention, treatment and relevant reporting procedures.

- Graduates should be conscious of the limitations of access to healthcare in certain areas of the world and their causes.

Life Long Learning

Life long learning is the refreshment and application of knowledge that physicians should perform with continuity for the rest of their career. A physician should be someone who is constantly up-to-date with their medical knowledge, ensuring that the patient care is evidence based and applied according to the current standards.

- Graduates should be able to identify their own learning needs.

- Graduates should learn strategies to continuously update their relevant medical knowledge and its practice.

- Graduates should assess knowledge and sources of information in terms of their relevance and reliability.

- Graduates should be aware of the benefits of life long learning and realise the consequences of not taking part in learning processes.

Professionalism – Attitudes, responsibilities and self development

Professionalism is an ongoing process, which starts during student-life but continues developing as the student moves into the role of a physician. Undergraduate education leads to a profession and students need to gain the abilities appropriate to a physician's role and identity. Students should play an active part in the development of their role as physicians, and they should be provided with a framework to facilitate this development.

Besides acquiring professional attributes, students should develop an ethical foundation in order to ensure optimal patient care in their future work. In addition, graduates should be aware of society's expectations and should possess sufficient management skills to be able to function within the healthcare sphere.

Professional Attitudes

- Graduates should possess the ability to build a positive professional relationship with the patient. This includes:
  - Showing respect for the patient's autonomy as well as their ability to make informed decisions about their own health and life.
  - Respecting confidentiality as defined by the relevant legal and ethical guidelines.

- Graduates should be willing to constantly refresh and update their knowledge and skills throughout their professional career.

- Graduates should be willing to teach colleagues the knowledge and skills they themselves have mastered.

- Graduates should be prepared to use their knowledge to educate and guide patients and the society in general.

- Graduates should be active in their contribution to the advancement of medicine.

Management, responsibility and decision making

- Graduates should be able to apply basic knowledge concerning leadership and management to professional situations.

- Graduates should employ strategies to cope with crises, conflict, uncertainty, errors and time limits.

- Graduates should have knowledge of the healthcare system in terms of effective patient care and cost effectiveness. They should be able to pay specific attention to rational prescription and use of resources.

- Graduates should have the ability to uphold the STEEEP (Safe, Timely, Efficient, Effective, Equitable, and Patient-centred) principle of patient care.

- Graduates should act responsibly bearing in mind the consequences of their actions, and be able to learn from mistakes.

- Graduates must be able to handle the responsibility needed to work as physicians.

- Graduates should be aware of and able to fulfil their legal responsibilities and obligations as doctors and be able to fulfil those.

- Graduates should know the limits of their knowledge, skills, experience, time, physical capabilities and health. To ensure patient safety, graduates must be able to seek appropriate help and assistance when they are beyond their own capacity.

- Graduates should have the ability to make decisions, both independently and as a part of a team.

- Graduates should be able to make professional decisions knowing that these may have great impact on peoples' lives. Therefore difficult decisions should, where relevant, be taken in conjunction with colleagues, the multidisciplinary team, patients and/or their relatives.

Self awareness

- Graduates must be able to continually evaluate and reflect on their work and role as a practitioner. They should be able to show development in response to both external feedback and self-assessment.

- Graduates need to be confident in their thoughts and actions within their level of competence whilst being aware of their own limits.

- Graduates should be aware of the pressures of a demanding profession and they should be prepared to deal with a stressful environment. Graduates should be familiar with resources available for stress management.

Ethical Principles

- Graduates should apply relevant ethical codes to everyday clinical work and be able to express a well-considered opinion on ethical issues.

- Graduates should be able to ensure appropriate interactions with the healthcare industry.

Teaching

We believe teaching to be an essential component of professional and educational interaction on every level in the medical field and that it plays a key role in maintaining excellence within the healthcare system.

- Graduates should be able to teach colleagues, students, other healthcare providers, patients and their relatives, communities and society at large. This includes:
  - Knowledge of teaching methods.
  - Having the skills to choose the most suitable method and content for the situation and the group or person being taught.
  - The ability to teach the latest up-to-date information in the subject they are teaching.

- Graduates should have knowledge of assessment methods and have the skills to choose the most suitable method for the situation, group or person being assessed

Teamwork

Graduates should aim to ensure optimal patient care by being able to work effectively as part of a team whenever necessary. He should therefore be able to demonstrate the skills and attitudes necessary to fulfil the relevant role.

- Graduates should be able to identify situations where teamwork is necessary and the appropriate composition of the team.

- Graduates should be able to work in a multidisciplinary team.

- Graduates should be able to distinguish the various roles they may be required to play and identify which ones are pertinent to the situation at hand.

- Graduates should demonstrate the attitudes and abilities necessary to work effectively in a team, aiming for excellence in patient centred care. These should include:
  - Leadership where appropriate
  - The ability to share information
  - Showing respect for, and understanding of, other professionals
  - The ability to effectively occupy different roles within a team as required by the situation

- Graduates should be aware of additional diagnostic and therapeutic options available within other healthcare professions.

- Graduates should be familiar with the relevant procedures of collaboration and communication with other bodies within the specific healthcare and legislative framework.

Theoretical knowledge

Graduates must have acquired a scientific foundation for the practice of medicine and be able to translate the knowledge gained into medical practice and professional competence.

They must be aware of the rapid changes and advances in knowledge and recognise the importance of lifelong learning. Newly qualified doctors must make a commitment to exchange of knowledge with peers, be able to recognise the limits of their knowledge, and be able to access appropriate sources of information and evaluate them.

Basic sciences, clinical disciplines and research

- Graduates should have core knowledge relevant to common clinical settings, in basic sciences and clinical disciplines.

- Graduates should understand diseases and biological variation based on knowledge of both the healthy and unhealthy body. They need to apply the principles of basic sciences, including research to clinical practice.

- Graduates should have knowledge of research theory.

Humanities, social and behavioural sciences

- Graduates should have knowledge of medical ethics.

- Graduates should understand the influence of social and behavioural sciences on the practice of medicine.

Community and the environment

- Graduates should understand social, environmental and occupational influences on health in the community.

- Graduates should have knowledge about cultural and religious variation within the population, and understand how people from different cultures or religions present and cope with common illnesses, treatment, death and dying.

Healthcare system

- Graduates should know the structure and functions of the healthcare system, the role of the doctor and other professions in the healthcare system.

- Graduates should know their legal obligations regarding patients' treatment and records.

- Graduates should have sufficient knowledge about the information technology of the healthcare system in which they are working.

- Graduates should know how prevention programmes can improve the health of the community and keep their knowledge up-to-date.

European dimensions

- Graduates should know about other healthcare systems since medical practice cannot be seen only within one country's perspective.

- Graduates should preferably have acquired knowledge (both written and oral) in one or more European foreign language and should have knowledge about European cultures.

Medical Education Systems

There is saying: „Each house with its own customs". This applies perfectly when talking about the medical systems used by different medical schools from all over the world. So it would be yet another „Mission Impossible" to try to put them all in one place.

Instead, it would be more useful to mention the main systems, with their top features, to give the reader a general idea.

"Classical" system

The probably most „classical" system is based on a two cycle structure consisting of a pre-clinical and clinical stage. Basically, pre-clinical are 2 years of basic natural sciences, anatomy, biochemistry and physiology. This is followed by a brief clinical-theoretical stage with basic pharmacology, basic pathology, clinical chemistry and so on. Clinical stage comprises all big clinical subjects like internal medicine, surgery, neurology, etc.
Some systems (e.g. Germany) include a full clinical stage (internship/elective) in the last year.
Studies finish with some kind of licensing exam, probably a state exam.

Integrated system

Some say this is the future. With its large array of teaching methods, interdisciplinary teaching, a focus on longitudinal learning (continuous medical education), contact with patients from the 1st year (compared to 3rd year in „classical" system), the integrated system provides the best flexibility to the medical schools, allowing them to draw up their own academic strategies and accordingly the best tools to achieve it.
Also, the students come in contact to a different way of studying than they are used to from high school, using new „instruments", like PBL.
Currently, some of the most famous med schools are transforming students into young doctors with the aid of an integrated medical system: Berlin, Maastricht, Harvard, Linkoeping (Sweden).

American system

The US medical students spend the most time to become doctors. After 4 years of pre-med college, there are 4 more years of medical school. The cycle ends with the USMLE exam, with 3 „steps", which gives the graduate student the license to practice medicine.

Any non-American graduate medical student who wants to practice medicine in the US and be integrated in the system must take this exam. Depending on their country, students have the option of taking the first „step" at home, but the last exam is in the US.
The US covers an entire continent, from East to West, so a great diversity in medical education systems would be quite expectable. The US medical schools have very different systems, varying from „classical" to integrated (Harvard).

BA/MA (Europe)

One of the most important parts of the Bologna Process, the Bachelor/Master in Medicine is still in its cradle. Although the BP covers all areas of higher education, both Medicine and Architecture have a special, specific-based status.
There is still a great debate concerning the precise definition of the BA/MA system in medicine, with a special focus on its actual implementation.
Currently, the medical schools in Switzerland, Denmark, the Netherlands and some other schools in Europe are implementing the BA/MA system. The issue has been topic of the 2007 IFMSA/EMSA Bologna Process Follow-up conference in Amsterdam and a policy has been written on it (see page 40).

Examples

Note: It differs a lot between the countries, which subjects are included in the pre-clinical phase. In some countries, Pathology, Pharmacology or Microbiology are part of the pre-clinical phase.

The following examples were collected by the participants of the SCOME-session at the third European Regional Meeting (EuRegMe) in April 2006 in Leicester (UK) and of the 55th IFMSA August Meeting in August 2006 in Zlatibor (Serbia). The systems of all countries present at the SCOME-sessions of these meetings can be found at the SCOME-wikipedia (wiki.ifmsa.org/scome).

Austria

3 public faculties (Innsbruck, Salzburg, Vienna), 1 private faculty
To enter medical school: EMS-Test

I. Preclinical:
   - 2 semesters (1 year)
   - Multiple choice test about all the content (SIP 1)

II. Between clinical:
   - 5 semesters (2 ½ years)
   - 3rd semester: 14 weeks of dissection course
   - Organised in ~12 modules such as infectious disease, cardiac & respiratory system, skin & mucosa
   - After the 4th semester: SIP 2
   - After the 7th semester: SIP 3

III. Clinical:
   - 5 semesters (2 ½ years)
   - After the 10th semester: SIP 4
   - 6th year: Internship
   - Final oral exam

In phase II. and III. 20 weeks of internship

After final exam 3 years of intern and residency, afterwards 4 – 6 years of specialisation.

Bosnia and Herzegovina
5 public faculties (Sarajevo, Banja Luka, Tuzla, Mostar, Foca)
I. Preclinical: 3 years
   - In the third year introduction to clinical courses
II. Clinical: 3 years
   - Last year: Only clinic, no lectures
III. Graduation exam: Research presentation
IV. 1 year residency, afterwards state exam

Bulgaria
3 years pre-clinical subjects / 3 years of clinical cycles

Czech Republic
3 years pre-clinical subjects / 3 years of clinical blocks / State exam
Practical training during the summer holidays

Denmark
3 years of pre-clinical subjects / 9 weeks of clinical stay / 3 years of clinical blocks / 6 weeks of elective assignment. Differs a little from university to university

Finland
Entrance exam / 2 years preclinical subjects / Exam / ½ year Pathology, Microbiology,
Immunology / 3 years clinical / ½ year internship / Final Exam
After 4th year you can work as a doctor. After 6th year 2 more years of „EURO-Doctor" phase follow. There is a progress test after each semester.

France
1 year basic sciences / 2 years pre-clinical subjects / 3 years clinical rotation and lectures / National exam

Germany
2 years basic sciences, 3 months nursing training / 1st state exam (320 MCQ & Oral) / 3 years clinical subjects, 4 months clerkship (at least one not in hospital) / „Practical year" (Full-time work on ward in Internal Medicine, Surgery, Elective, 16 weeks blocks) / 2nd state exam (320 MCQ & Oral)
Due to the new law on medical education in Germany, it is possible for a faculty to override the basic structure and also introduce an integrated educational system, which also finishes with the 2nd state exam.

Greece
2 years of basic sciences / 2 years of pre-clinical courses / 2 years of clinical rotations

Netherlands

3 – 3 ½ years pre-clinical but integrated subjects / 2 – 3 years of clinical clerkship

Romania
2 years pre-clinical / 4 years clinical / state exam / 3 to 7 years of residency (based on a national exam)

Turkey
3 years pre-clinical phase with written exams at the end of each year / 2 years clinical phase with rotations through departments / 1 year of internship without theory / Central exam

United Kingdom
2 – 3 years of pre-clinical basic sciences and basic clinical skills training / 3 years of clinical training based in hospitals / Final exam accredited by General Medical Council (GMC).
In the United Kingdom most universities implemented an outcome-based curriculum due to the General Medical Council's publication „Tomorrow's doctors".

Conclusions

Although it might seem that we have overlooked some things or even forgot them all together, just keep in mind that each of the system mentioned above is moulded and transformed by each medical school according to its academic strategy and its needs, but also by its traditions and geographic area.
None of the systems is perfect if used in a rigid manner. Again, students' mobility is affected by the discrepancies between the systems used by different medical schools. This situation can be improved by implementing a core curriculum, which could maybe transcend geographical borders.
Also, traditions and sometimes conservationism stay in the way of a drastic change in old-fashion and outdated ME systems.
Maybe the most important thing about the ME systems and their transformation is that they should have in mind also the student dynamics, both in number and "quality".

Quality Assurance

Accreditation

Accreditation is a process by which an accreditation body evaluates the quality of a higher education institution as a whole (institutional accreditation) or a specific higher education programme (programme accreditation) in order to formally recognise it as having met certain predetermined minimal criteria or standards.

UNESCO-CEPES defines "Accreditation" as

1. The process by which a (non-)governmental or private body evaluates the quality of a higher education institution as a whole or of a specific educational programme in order to formally recognize it as having met certain predetermined minimal criteria or standards. The result of this process is usually the awarding of a status (a yes/no decision), of recognition, and sometimes of a license to operate within a time-limited validity. The

process can imply initial and periodic self-study and evaluation by external peers. The accreditation process generally involves three steps with specific activities:
   1. a self-evaluation process conducted by the faculty, the administrators, and the staff of the institution or academic programme, resulting in a report that takes as its reference the set of standards and criteria of the accrediting body;
   2. a study visit, conducted by a team of peers, selected by the accrediting organization, which reviews the evidence, visits the premises, and interviews the academic and administrative staff, resulting in an assessment report, including a recommendation to the commission of the accrediting body;
   3. examination by the commission of the evidence and recommendation on the basis of the given set of criteria concerning quality and resulting in a final judgment and the communication of the formal decision to the institution and other constituencies, if appropriate.
2. The instrument by which one institution, without its own degree awarding powers or which chooses not to use its awarding powers, gains wide authority to award, and/or gains recognition of its qualifications from another competent authority, and to exercise powers and responsibility for academic provision. This authority might be the State, a government agency, or another domestic or foreign higher education institution.

As defined in the Bologna Declaration, the study structure of the European Higher Education Area (EHEA) should essentially be characterised by two cycles – undergraduate and graduate. Accreditation is a central instrument to support the necessary processes of changes in European higher education systems. Like evaluation, accreditation serves to assure quality when implementing new degree programmes and also to monitor existing ones. Accreditation, i. e. certification of a degree programme, will take place after review of the minimum standards for content and specialisation, the vocational relevance of the degree to be awarded and the coherence and consistency of the general conception of the degree programme. It will be awarded for a limited period of time within the frame of a transparent, formal and external peer review. Thus, the degree programme has to be reviewed after a certain time. The process of a peer review is steered by agencies, which are also reviewed through regular external evaluation. The instrument of accreditation of certificate degree programmes is relatively new in Europe but is increasingly gaining acceptance in the countries involved in the Bologna process.

Evaluation

Evaluation is a systematic and critical analysis leading to judgements and/or recommendations regarding the quality of a higher education institution or a programme.

UNESCO-CEPES defines evaluation as an ability to perform well or to achieve a result without wasted resources, effort, time, or money (using the smallest quantity of resources possible). Educational efficiency can be measured in physical terms (technical efficiency) or in terms of cost (economic efficiency). Greater educational efficiency is achieved when the same amount and standard of educational services are produced at a lower cost, if a more useful educational activity is substituted for a less useful one at the same cost, or if unnecessary educational activities are eliminated. A programme or a higher education institution may be efficiently managed, but not effective in achieving its mission, goals, or objectives.

Apart from accreditation, evaluation is the central activity to assure quality in higher education. To evaluate means to assess teaching and academic studies in a subject or

department and the related degree programmes. Strengths and weaknesses of education and training should be demonstrated by stocktaking and analysis and proposals should be formulated to promote its quality as well. Evaluation is carried out through internal or external procedures. The process of internal evaluation is comprised of the systematic collection of administrative data, questioning of students and graduates, as well as moderated conversations with lecturers and students. As part of the process of external evaluation a review team visits the department in order to review the quality of the academic studies and teaching. External peers are lecturers or persons from vocational practice who discuss with students and young scientists and present a final report. The evaluation of academic studies and teaching has to be followed by an account of how effective the measures of quality assurance are. Besides academic studies and teaching, the performance of research is evaluated at different levels: with reference to national research systems, individual institutions, research programmes or individual projects. In the field of research evaluation internal and external evaluations are also employed.

What is evaluation all about? Is it worth the work at all?

First of all it is an instrument to assess any outcome of any project. It helps you to improve the quality of the project and to optimise processes when running the project again. It helps you to get feedback from all the participants and to create a follow-up.
"If you don't know where you're going,
it doesn't matter which way you take."
But used in an inadequate way it may not help you at all to follow your outline.
It is important to state your goals and objectives before you start thinking about the questionnaire itself. Then the evaluation should be outcome-based, asking the right questions concerning the outline.

How can I start a proper evaluation?

There are eight steps to take to set up an evaluation questionnaire.

1. Define your goals and objectives!
   To assess an outcome you might prefer a summative evaluation (it can be an exam), to improve a project formative evaluation might be the one that should be used.
   Be aware that the time-point of the evaluation in relation to the project is important. Evaluating every single lecture of a long course several weeks later is difficult if you ask to specific questions. On the other hand having an evaluation after each lecture leads to tiredness. Motivating participants to evaluate regularly may be a problem.
2. Brainstorm on the different items you want to evaluate referring to your goals and objectives. The OECD DAC (Organization for Economic Cooperation and Development – Development Assistance Committee) set up certain standard evaluation criteria, which also might fit into your goals and objectives. These are: Efficiency, Effectiveness, Impact, Relevance and Sustainability.
3. Then you should choose the appropriate instrument for your evaluation. Select the tools for descriptive purpose (leading to a summative evaluation) as well as for diagnostic purpose (leading to a formative evaluation). Different evaluation tools are for example assessment, self-assessment, fast-feedback-questionnaire, evaluation-form-questionnaire, interviews, peer-reviews and feedback-groups, but also tests and exams are a kind of evaluation evaluating each student.
4. This step is an optional one, but taking it can prevent serious problems: you should consider potential sources of error. There might be cultural differences among those who

fill the questionnaire. Having language problems and understanding problems may also lead to failure of the evaluation. So be aware of evaluating a heterogeneous group!

5. The next step is a pre-test of the evaluation among a small collective. This helps you find and adjust some coarse mistakes and to apply improvements on answer or question format.
6. Step six is the evaluation itself. Participants are motivated by personal addressing them. The aims of the evaluation should also be explained.
7. The seventh step is the analysis and interpretation.
   The question forms should be recollected directly after the evaluation. It does not make sense to evaluate to late after a project. The participants won't remember the details then, even if you ask only global questions. Then you collect all the data, interpret it and summarize the outcome with regard to your goals and objectives.
   The outcome must then be published. Publishing it creates transparency and motivates both sides of the project. The evaluation is then more accepted among all the persons involved. Maybe some suggestions and changes can also be published at that time.
8. The last step is also the most important one: draw consequences from the outcome!
   You should design a follow-up and make another plan of action. After all the changes are applied, they also need to be evaluated.

Before going into some examples of different question formulation, rating scales, validity and items, I would like to summarize some „take-home-messages":

- Clear goals and objectives are essential for evaluation
- Evaluation without follow-up is useless.
- Transparency and involvement significantly improve the efficiency of the evaluation
- Evaluation can definitely improve your project!

Question formulation

Question:	„Is the food at the restaurant good?"
Statement:	„The food at the restaurant is good."
Positive formulation:	„The food at the restaurant is good."
Negative formulation:	„The food at the restaurant is not good."
Subjective formulation:	„I like the food at the restaurant"
Objective formulation:	„The food at the restaurant is good."

Descriptive formulation is better understandable.
You should avoid double negation (Isn't the food at the restaurant not good?) as it's difficult to understand.
Using rating scales (i.e. not open answer), also double questions (Do you like breakfast and/or diner at the restaurant?) should be avoided; as one doesn't know then, to which part of the question the answer refers to.
The questions and statements should be kept as short and simple as possible.

Rating scales

The scale size has an influence on the result. You can have more precise results having 10 possibilities. But on the other hand even Yes-No may also be enough. It's important at this point, that you keep an eye on your goals and objectives!
The answer format must fit to the question itself, of course. Agree/Disagree-Scales do not fit to any kind of question. The opportunities to answer the question must differ clearly.

It makes a difference weather you use an even (2, 4, 6…) or an odd (3, 5, 7…) scale. Odd scales always give the opportunity to check the middle, which in an extreme way may lead to an evaluation where all middle items are crossed and for this reason don't lead to any conclusion.
The „left-right-placement" also influences the results. In country's using Latin letters there is a tendency to choose options on the left side of the scale. If you put „negative" on the left side the results of course differ compared to having „positive" on the left side.

Valid evaluation

In order to find out if an evaluation is valid, the ratio of participants vs. filled questionnaires must be taken into consideration. You should always give the possibility NOT to take part in the whole evaluation or not answering certain questions. Then it is improbable that somebody just answers the questions without reading them at all.
You must stick to your original outline and should prepare the questions before the project started. Otherwise mistakes made or successes lead to different kind of questions in the questionnaire.
For example, if a session starts delayed or equipment is missing these things would definitively appear in the questionnaire and then easily a predominance either to the positive or the negative things exists. Therefore it is important to add open questions, so that things like missing equipment can still be mentioned by the participants.
Before the evaluation starts it should be reviewed to find mistakes in question and answer format and to eliminate complicated items.

Items

The answer format must be appropriate to the question and the question format. Open questions are important, because you can figure out all the possible things that might happen during the project in advance when setting up the questionnaire.

Further reading

Ronald A. Berk, Professor of Biostatistics and Measurement at the School of Nursing, The Johns Hopkins University, has published an easy-to-read introduction to evaluation theory: "Thirteen strategies to measure college teaching". All principles described there can easily be adopted for any kind of evaluation, questionnaire or scale. The book has been published by Stylus Publishing (ISBN 1-57922-193-9).

Furthermore various articles have been published on the topic of evaluation.

- Morrison J. Evaluation. BMJ 2003; 326; 385-387
- Snell L, Tallett S, Haist S, Hays R, Norcini J, Prince K, et al. A review of the evaluation of clinical teaching: new perspectives and challenges. Med Educ 2000;34:862-70.
- Wilkes M, Bligh J. Evaluating educational interventions. BMJ 1999;318:1269-72.
- Goldie, J. AMEE Education Guide no. 29: Evaluating educational programmes. Med Teach; 28:3, 210-224

Academic quality of IFMSA Professional Exchanges

During the 55th August Meeting of IFMSA in Zlatibor (Serbia) in 2006, a joint SCOPE/SCOME preGA took place on "Improving the Academic Quality of IFMSA Professional Exchanges". Result of this preGA was the "IFMSA Exchange Tutor Kit". It should help medical students participating in the SCOPE programme to get credits for the clerkship.

IFMSA Exchange Tutor Kit

Student Exchanges organized by IFMSA

The International Federation of Medical Students' Associations (IFMSA) is an independent, nongovernmental and non-political federation of medical students' associations throughout the world. The IFMSA currently has 92 members, National Member Organizations from approximatly 80 countries on six continents and represents more than 1 million medical students worldwide.
The IFMSA was founded in May 1951 and is run for and by medical students on a non-profit basis. It is officially recognized as a Non Governmental Organization (NGO) within the United Nations' and recognized by the World Health Organization as the International Forum for medical students. It exists to serve medical students all over the world and was established in the Netherlands as a charity organization.
Since its foundation, IFMSA has sought to give medical students across the world the opportunity to experience medical training in other countries. Each year, more than 7,000 exchanges take place. By participating in exchanges, students gain invaluable insights into other health systems and cultures, enabling them to view the more familiar environment in their own hospitals and communities with fresh perspectives.
The exchanges are initiated and administered entirely by students, which strengthens the network of co-operation within the Federation. The local students willingly take responsibility for the welfare of the visiting students, significantly strengthening the intercultural and international understanding and solidarity between these young people.
The participating medical students spend an average of four weeks in a hospital, on a clinical "clerkship". During this time, they see both the different disease burdens of the local population and practice within a different health system and culture. While knowledge of the local language is extremely helpful, a common knowledge of English, from the tutor and student, is often sufficient.

Potential pitfalls

Students can encounter difficulties on beginning a clinical clerkship. Many of these can be eliminated by adequate preparation and support on arrival at the host institution. We suggest an initial orientation session whereby students and their host department can discuss their respective aims and expectations. This session could consider the following issues:

- Language barriers
- Level of education and competencies
- Relevant local ethical and legal issues

Note on assessment

Why assess

The main objective of the checklist is to implement a standard continuous formative assessment of international students' clinical skills in order to ensure the academic quality of the exchange.
Why is assessment necessary for achieving a high academic quality in clinical education? The learning outcomes and work performed by the students are related to the way in which they are assessed. Assessment of clinical competence of undergraduate medical students plays a key role in their education. In that way assessment has 3 major functions:

- Take decisions over student promotion – the selective function
- Provide feedback to the student – the formative function
- Monitor the quality of the education programme – the accountability function

Assessment in clinical clerkships is essential for providing feedback to the student (formative function). The problem that students experience in applying theoretical knowledge to practice demonstrates the crucial importance of feedback. The clerkship as a teaching programme also benefits from more emphasis on the formative function of assessment. In this way assessment stimulates improvement of the quality of clerkships from a didactic perspective. When skills are assessed students tend to be more motivated. The criteria that are assessed are guiding points for the individual study of the student.

When & how

As clerkship assessment is regarded as formative testing, information should be carried through longitudinally. This method of evaluation of students' clinical skills will reveal weak points and thereby enable them to improve their skills.

The checklist should be used by the student and the tutor as a guide in the daily practice. It is designed to give both student and tutor the possibility to see if the learning goals of the clerkship are achieved. For this, we have drawn from Miller's Pyramid of competence. In order to function in a practical setting, the names of the different layers have been changed.

- Knows: observes
- Knows how: assists
- Shows how: Does under supervision
- Does: Does independently

The checklist also constitutes a document that states that the student has successfully fulfilled the clerkship in a way they might use it for accreditation at their home university. For this reason, it should be used in a responsible way.

The attached logbook is an additional help for the student to monitor his learning progress. It is not supposed to be used as a grading tool. Blank spaces or unmarked items in the checklist, as well as in the logbook, could be interpreted as impossible to encounter or perform due to variations in medical practice and not necessarily as the lack of initiative to seize opportunities.

Why bedside teaching

"To study the phenomena of disease without books is to sail an uncharted sea.
Whilst to study books without patients is not to go to sea at all"
Sir William Osler (1848-1919).

Bedside teaching is an essential component of medical education. It is one of the most effective ways to learn clinical and communication skills. Basic clinical skills are a quick and cheap way of reaching a diagnosis. For example a comprehensive physical examination can provide the diagnosis of more than 70% of diseases. Communication is the basis for a good doctor-patient relationship, which will ensure appropriate patient care, increase patient satisfaction and reduce malpractice lawsuits.

Important Issues to be aware of

Doctors increasingly have to combine clinical, research, administrative and educational duties. This has lead to a decrease in bedside teaching, resulting in students encountering less patients with less opportunities to practise clinical skills. Students are rarely supervised when performing clinical skills and most of the time they do not receive feedback. A further problem encountered by students is the demand on time looking for patient files and waiting for doctors reducing the time available for valuable learning experiences. Careful planning and adequately prepared bedside teaching can overcome these problems. A common concern is that patients

will be uncomfortable with taking part in bedside teaching. Research, however, has shown that the majority of patients enjoy bedside teaching and gain new information about their disease, especially if they are prepared prior to teaching sessions.

Tips for good bedside teaching

A lot of research exists in this area with many researchers developing models for good bedside teaching. Some basic points are however common to all of these.

There should be a defined outcome for the students' clerkships, such as a checklist or logbook. That will ensure that both teachers and students know what to achieve. In order to achieve the outcomes, the teacher should have a basic knowledge of bedside teaching.

- It is very important to ask the patients permission in advance of teaching sessions. It is ethical to inform them of what to expect and give them the opportunity to refuse. They should understand that the students and the teacher will be discussing their disease, and aspects relating to it in front of them.
- The teacher should introduce all the students to the patient making sure that they have understood the role of the medical student as a student doctor as many patients are unsure what the term medical student covers.
- The teacher should know what skills the students already posses and those which they need to acquire, so the teacher can focus the teaching session towards the students requirements. The same applies to the acquisition of practical knowledge.
- A brief overview of the patient's history should be given,physical examinations skills should be practiced and diagnoses and treatment plans should be discussed. Avoiding medical terminology is another important element.
- At the end of the teaching session, the teacher should ask the patient if they have questions or comments.
- It is important to discuss the session with the group of students afterwards and give relevant and encouraging feedback.

Conclusion

It is impossible to imagine medicine taught without patients. Bedside teaching is the only place where history taking, physical examination, empathy and professionalism can be taught, experienced and learnt by example.

The complete document and the checklists can be found at the SCOME-wikipedia (wiki.ifmsa.org/scome) in the article "IFMSA Exchange Tutor Kit".

Appendix 1:

List of IFMSA-Medical Education Directors

1952 – 1953	???	The Netherlands
1957 – 1958	???	The Netherlands
1958 – 1959	Per Hanson	Sweden
1959 – 1960	Per Hanson	Sweden
1960 – 1961	Werner Flohr	Federal Republic of Germany
1961 – 1962	Friedhelm Katzenmeier	Federal Republic of Germany
1962 – 1963	Peter Kussmauer	Federal Republic of Germany
	Karl W. Ostarhild	Federal Republic of Germany
1963 – 1964	Ezard Bertram	Ghana
1964 – 1965	Hannu Vuori	Finland
1965 – 1966	Harry Frey	Finland
1966 – 1967	Wojcech Leszczynski	Poland
1967 – 1968	Hans Heinz Lobner	Austria
1968 – 1969	Hans Heinz Lobner	Austria
1969 – 1970	???	USA
1970 – 1971	Pedrag Banic	Yugoslavia
1971 – 1972	Pedrag Banic	Yugoslavia
1972 – 1973	Avner Hershlag	Israel
1973 – 1974	Dare Demuren	Nigeria
1974 – 1975	Fritz Conrad	Austria
1975 – 1976	???	???
1976 – 1977	???	???
1977 – 1978	???	???
1978 – 1979	???	???
1979 – 1980	Reuven Rabinovici	Israel
1980 – 1981	David Klein	Israel
1981 – 1982	Osnat Geifman	Israel
1982 – 1983	Yoram S. Sandhaus	Israel
1983 – 1984	George Wolfs	The Netherlands
1984 – 1985	George Wolfs	The Netherlands
1985 – 1986	Rick Steele	Denmark
1986 – 1987	Mohamed Saleh Moustafa Hassan	Egypt
1987 – 1988	Itay Goor Aryeh	Israel
1988 – 1989	???	???
1989 – 1990	Gilad Barnea	Israel
1990 – 1991	???	???
1991 – 1992	Cristina Bonetti	Italy
1992 – 1993	Pedro Santos Canonico	Spain
1993 – 1994	Wolfram Antepohl	Germany
1994 – 1995	Wolfram Antepohl	Germany
1995 – 1996	Wolfram Antepohl	Germany
1996 – 1997	Katja Nevala	Finland
	Eva Schmidtke	Sweden
1997 – 1998	Paola Erba	Italy
1998 – 1999	Thiago Monaco	Brazil

1999 – 2000	Thiago Monaco	Brazil
2000 – 2001	Teele Raiend	Estonia
2001 – 2002	Nikola Borojevic	Croatia
2002 – 2003	Özgür Onur	Germany
2003 – 2004	Hans Jacob Westbye	Norway
2004 – 2005	Katja Kovac	Slovenia
2005 – 2006	Jan Hilgers	Germany
	Carl Savage	Sweden
	Note: Carl Savage was suspended from office at the 55th August Meeting	
2006 – 2007	Maja Basnov	Denmark
2007 – 2008	Daniel Rodriguez Muñoz	Spain
2008 – 2009	Nikolaos Davaris	Greece

Appendix 2:

List of IFMSA Liaison Officers on Medical Education Issues

2000 – 2001	Mats Sundberg	Sweden
2001 – 2002	Mats Sundberg	Sweden
2002 – 2003	Özgür Onur	Germany
2003 – 2004	Özgür Onur	Germany
2004 – 2005	Hans Jacob Westbye	Norway
2005 – 2006	Hans Jacob Westbye	Norway
2006 – 2007	Jan Hilgers	Germany
2007 – 2008	Jan Hilgers	Germany
2008 – 2009	Robbert Duvivier	The Netherlands

Index