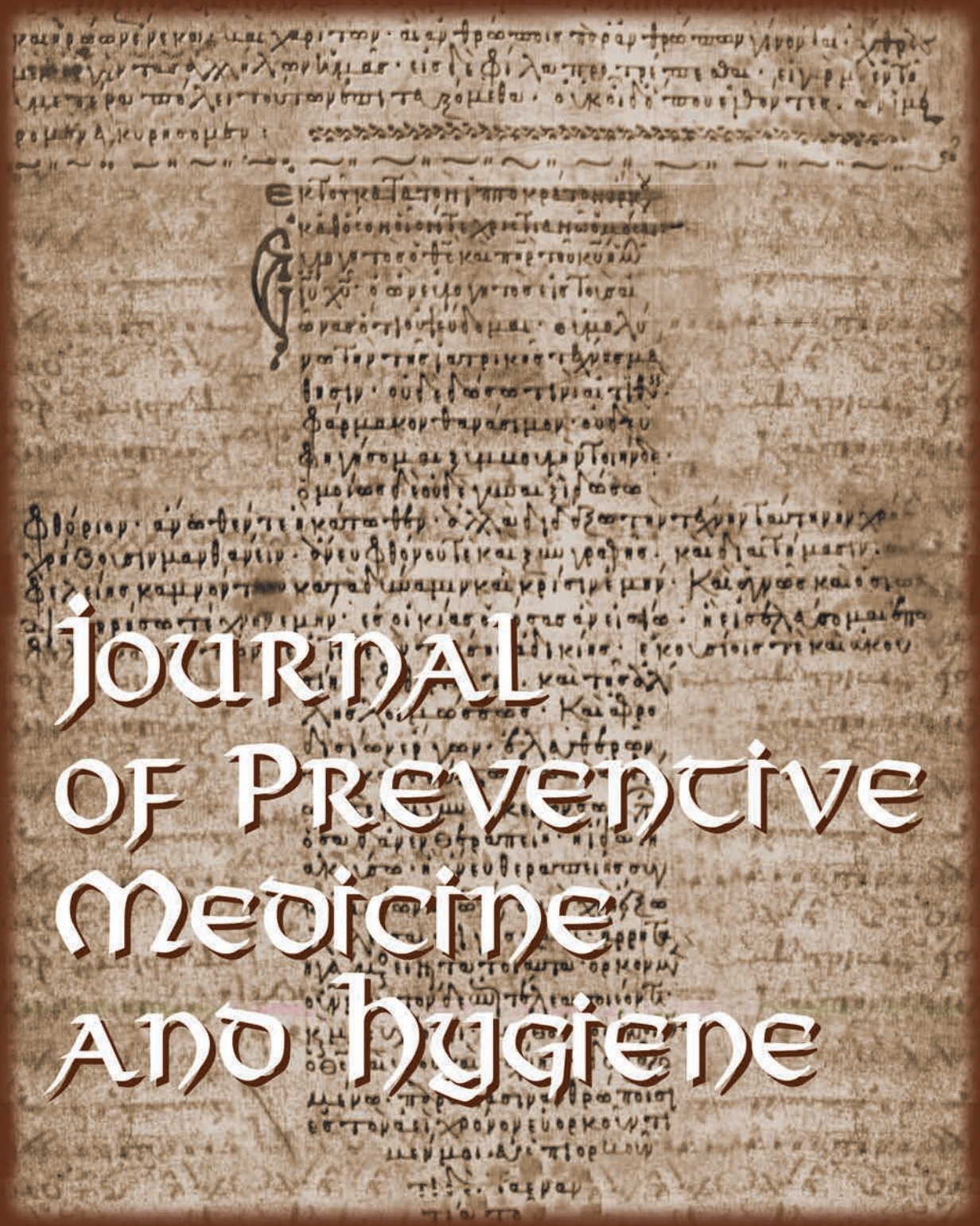
ISSN 2421-4248

vol. 56/4 n.

December 2015

Cited in Index Medicus / Medline
NLM ID 921440 (Pub-Med)



The original document of Hippocrates' oath

# Journal of Preventive Medicine and Hygiene

The Journal has been accreditated, on occasion of the 17th December 2004 Meeting of the Executive and Scientific SItI Councils, by the Italian Society of Hygiene, Preventive Medicine and Public Health

PACINI EDITORE MEDICINA

# JOURNAL OF PREVENTIVE MEDICINE AND HYGIENE

**http://www.jpmh.org**





Managing Editor: Patrizia Alma Pacini
Publisher: Pacini Editore Srl, Via Gherardesca 1, 56121 Pisa, Italy

Published online December 2015.

Authorization Tribunal of Genoa, Italy n. 507 - 10/6/1960
Journal registered at "Registro pubblico degli Operatori della Comunicazione" (Pacini Editore srl registration n. 6269 - 29/8/2001).

CONTENTS

**Original articles**

ORIGINAL ARTICLE

# How can the results of Health Technology Assessment (HTA) evaluations applied to vaccinations be communicated to decision-makers and stakeholders? The ISPOR Rome Chapter Project

R. GASPARINI[1], F.S. MENNINI[2], D. PANATTO[1], P. BONANNI[3], A. BECHINI[3], W. RICCIARDI[4], C. DE WAURE[4], A. MARCELLUSI[1], A. CICCHETTI[5], M. RUGGERI[5], S. BOCCALINI[3]

[1] Department of Health Sciences, University of Genoa, Italy; [2] Faculty of Economics Centre for Economic and International Studies (CEIS) -Economic Evaluation and HTA (EEHTA), University of Rome, Italy; [3] Department of Health Sciences, University of Florence, Italy; [4] Institute of Public Health, Catholic University of Sacred Heart, Rome, Italy; [5] Institute of Health Economics and Management, Catholic University of Sacred Heart, Rome, Italy

Key words

Vaccination • Prevention • HTA

Summary

*HTA is considered the most comprehensive and transparent method of supporting decision-makers in their choices in Public Health.*
*HTA on vaccines is being performed by many experts. However, they often present their studies to colleagues, but not to decision-makers, who should be the main target and current users. It is therefore crucial to improve the transfer of scientific data to decision-makers and all stakeholders.*
*The aims of the present project are: 1) to set up a team of experts to collect economic evaluations and HTA studies on vaccines and assess their actual use in decision-making processes; 2) to constitute regional working groups in order to identify the critical aspects of the communication process and identify the most appropriate method of data transfer.*
*Systematic reviews of economic evaluations and HTA on vaccines and their actual use in decision-making will be used to draw up the basic documents for discussion by the 3 regional working boards. The working groups will discuss the current scientific evidence and communication methods and will try to implement a model of technology assessment with well-defined and objective criteria, in order to better fit pharmaco-economic and HTA methods to the field of vaccinations.*
*Improving the transfer of HTA results to stakeholders, particularly decision-makers, will enable decisions to be taken on the basis of scientific evidence, and appropriate, sustainable actions to be undertaken.*

## Introduction

In accordance with the recommendations of the World Health Organization (WHO) (2005) [1], the current *Italian National Plan for Vaccination Prevention 2012-2014* (PNPV) [2] specifies that the choices made in the sector of prevention, such as those concerning vaccination, should be based on clear, robust, shared criteria. This need is due to the increasing availability of new vaccines, the introduction of new vaccination indications and the limited economic resources of health systems, even in developed countries. Only by prioritizing all possible preventive interventions can we rationalize the use of limited resources and maximize results in terms of health. As *Health Technology Assessment* (HTA) meets the above-mentioned criteria, it may constitute the best approach to this issue.

HTA is regarded as the most comprehensive and transparent method of supporting decision-makers in their choices concerning health technologies. This approach is even more important in the Public Health sector, particularly with regard to vaccine prevention [3]. As vaccination targets healthy individuals, its benefits (such as absence of disease, reduced disease burden etc.) are often unapparent to both the population and decision-makers. Consequently, preventive interventions are viewed only as a considerable cost to the health system, rather than as fundamental investments that generate benefits in the short and long terms. These issues are even more evident today, in the current critical period of spending review in Italy. Indeed, the most debated issue related to the introduction of a new vaccine is that of economic sustainability. In order to verify economic sustainability, pharmaco-economic evaluations are needed. The first step in this process is to evaluate the costs that a disease generates for the National Health Service (point of view of the third-party payer) and the community (point of view of society). Pharmaco-economics offers several approaches. In the field of vaccinations, the most frequently used approach is cost-effectiveness analysis, whereby the cost of vaccination is compared with its clinical benefit (e.g. the number of cases avoided or the years of survival gained). Among cost-effectiveness analyses, cost-utility analysis is preferentially adopted, as benefits are measured in QALYs (quality-adjusted life-years). When comparing two or more alternatives which have different

costs and degrees of efficacy/effectiveness, it is important to obtain information in a single summary measure, i.e. considering both cost and effectiveness. The *Incremental Cost-effectiveness Ratio* (ICER) provides this information and expresses the additional cost of each unit of a health outcome (e.g. one QALY) obtained by one strategy compared with another. The ICER is a fundamental parameter for decision-makers. However, decision-makers must also consider other domains: the epidemiology of the disease in question; vaccine efficacy/effectiveness; the predicted clinical impact on the population, as assessed through mathematical modelling of possible vaccination strategies; ethical, social and legal aspects, and the impact of introducing a new vaccine on health organizations. HTA deals with all these aspects.

In Italy, as in other countries, the application of HTA to vaccines has increased in recent years. Many HTA reports and pharmaco-economic studies, including several in the Italian context, have been published [4-24]. Although the data from these studies are widely available, they are often underused by decision-makers, who may lack the knowledge and ability necessary to exploit them fully. Indeed, decision-makers often evaluate the implementation of new vaccines only on the basis of vaccination costs, without considering the clinical and economic benefits derived from the avoidance of cases of disease.

It is therefore essential to improve the transfer of scientific data to decision-makers and all stakeholders, in order to promote evidence-based decisions in the area of immunization. The introduction of a new vaccine can only be based on Evidence-Based Medicine (EBM), or rather Evidence-Based Prevention (EBP) [25]. To achieve this goal, it is necessary to create a communication network linking all subjects involved in the decisional process: researchers, decision-makers and stakeholders. In this perspective, the project "HTA for public health decision-makers in the area of vaccination" has been planned and endorsed by the ISPOR Rome Chapter.

The *International Society for Pharmacoeconomics and Outcomes Research* (ISPOR), founded in 1995, is a public non-profit educational and scientific association. It promotes pharmaco-economic and outcome research and facilitates the translation of the results of that research into useful information for decision-makers in order to encourage efficiency, effectiveness and equity in health and to improve the health of the population. The ISPOR Italy-Rome Chapter is a working group affiliated to ISPOR and sharing its mission and objectives. It involves healthcare professionals, providers of healthcare technologies, decision-makers and other pertinent subjects in issues of pharmaco-economics and outcome research in the Italian context.

## Objectives

The objectives of the project are:

1) to organize a multidisciplinary working group composed of experts (the Project's founding group) in order to:
   - carry out systematic reviews of economic evaluations and HTA reports on vaccines and vaccination strategies;
   - assess the current use of HTA and economic evaluations in decision-making processes;
   - implement a model of technology assessment with well-defined and objective criteria, in order to better fit pharmaco-economic and HTA methods to the field of vaccinations.
2) to constitute regional working groups composed of experts, decision-makers and stakeholders in order to:
   - identify and define critical points in communication from/to decision-makers and stakeholders and in data transfer processes, in order to identify the most appropriate method of data transfer;
   - increase decisions taken on the basis of scientific evidence in the field of vaccinations;
   - improve the culture of vaccination in the Italian population and counteract the spread of erroneous, confusing and outdated information.

Achievement of these objectives will lead to the identification of the best strategies for creating a "data bridge" between researchers, decision-makers and stakeholders.

## Materials and methods

Two different main activities will be performed: systematic reviews of economic and HTA studies on vaccine/vaccination strategies, and organization of working boards. Table I shows the main activities of the ISPOR Rome Chapter Project.

The following specific activities will be carried out:

- systematic reviews of economic and HTA evaluations of vaccination strategies or of vaccines already included in the Italian vaccination schedule or still under discussion for inclusion in the current regional or national vaccination programs: for example, HPV vaccination in females and males, meningococcal C vaccine, pneumococcal vaccine in children and the elderly, varicella vaccine, quadrivalent conjugate meningococcal vaccine, meningococcal B vaccine, rotavirus vaccine, and herpes zoster vaccine. This systematic review will be performed on PubMed, Scopus and NIHR HTA databases; articles published at any time up to the end of 2015 will be included. PubMed and Scopus will be queried by means of a combination of MESH terms and keywords referring to vaccines/vaccinations and HTA/economic evaluations, while the search on NIHR HTA will be performed through the use of keywords related only to vaccines/vaccinations. Furthermore, a hand-search of the following journals will be carried out: *Global & Regional Health Technology Assessment, Pharmaco-Economics Italian Research Articles, Giornale Italiano di HTA, Politiche Sanitarie, HTA Focus – Pills of Clinical Governance, Pillole di Farmacoeconomia, Giornale Italiano di Farmacoeconomia e Farmacoutilizzazione, Italian Journal of Public*

**Tab. I.** The main activities of ISPOR Rome Chapter project.

| Activities | Working boards |
|---|---|
| Systematic reviews on economic and HTA evaluations of vaccination strategies or vaccines already included in the vaccination schedule in Italy and Europe. | Project's founding group |
| Systematic reviews on economic and HTA evaluations for vaccination strategies or vaccines that are as yet under discussion for inclusion in the current regional or national vaccination programs in Italy and Europe. | Project's founding group |
| Report on the real use of economic and HTA studies in decision-making process in Italy. | Project's founding group |
| Development of mathematical models for compared some vaccination programmes with other health interventions. | Project's founding group |
| Identification of critical points in communication from/to decision-makers and stakeholders and in data transfer processes in order to identify the most appropriate method of data transfer. | Regional working groups |
| Increasing of decisions taken on the basis of scientific evidence in the field of vaccinations. | Regional working groups |
| Improving of the culture of vaccination in the Italian population and counteracting the spread of erroneous, confusing and outdated information. | Regional working groups |
| Implementation of a model of technology assessment with well-defined and objective criteria in order to better fit pharmaco-economic and HTA methods to the field of vaccinations. | Project's founding group |
| National event of working groups | |

*Health and Quaderni dell'Italian Journal of Public Health.* Eligible articles will be identified through a two-step approach: screening of titles/abstracts and reading of full texts. Articles will be considered eligible if they report the results of an economic evaluation or of an HTA performed in Italy or one or more Italian Region;

- evaluation of the real use of economic and HTA studies in the decision-making process in Italy in recent years. This evaluation will be performed by comparing the results of the studies collected with the decisions taken at the national and regional levels in recent years, with regard to the inclusion of new vaccines or vaccination strategies in immunization programmes;
- some vaccination programmes will be compared with other health interventions (therapeutic interventions) by means of mathematical models, in order to increase decision-makers' awareness of the fact that investments in the field of vaccines generate significant benefits and are not only a cost for the health system. For example, HPV vaccination in females will be compared with therapeutic interventions in women with HPV-related lesions/cancer.

These activities will be carried out by the Project's "founding group", made up of researchers working on these subjects at the Universities of Florence, Genoa and Rome (Catholic University and Tor Vergata University). The founding group will also implement a model of technology assessment with well-defined and objective criteria, in order to better fit pharmaco-economic and HTA methods to the field of vaccinations.

Data acquired from these activities will be used to draw up the basic documents for discussion by the working boards that will be organized. The working boards will be composed of all relevant subjects involved in the decision-making process: researchers, decision makers and stakeholders. (Tab. II). The parties that have an interest are: citizens and their associations, health professionals and their associations, scientific societies, those elected by citizens to the various levels of government, businesses and no-profit partners, academics, and volunteers and their associations [26].

The importance of involving as many stakeholders as possible stems from the need to improve the culture of vaccination in the Italian population and to counteract the spread of erroneous, confusing and outdated information. Indeed, the success of a vaccination campaign depends not only on the supply of a vaccine free of charge, but also on the acceptance of vaccination on the part of the target population. The acceptability of vaccination strongly influences vaccine coverage rates, and only high vaccine coverage enables the incidence of disease to be reduced, thereby achieving considerable savings in terms of both healthcare and socioeconomic costs. Opinion leaders and journalists will be involved in the working groups, on account of their significant influence on the choices of the population and decision-makers. In particular, journalists can support the improvement of communication between researchers and decision-makers by identifying the most appropriate method of data transfer. Journalists can also help to spread the culture of vaccination in the population and increase public awareness of the importance of prevention activities.

**Tab. II.** Multidisciplinary working groups of ISPOR Rome Chapter Project.

| | |
|---|---|
| Researchers | University researchers<br>Independent researchers |
| Decision-maker | Minister of Health<br>Institute of Health<br>Regions<br>Minister of Economic<br>Italian Medicine Agency<br>Minister of University |
| Stakeholders | Vaccine manufacturers<br>Scientific Societies<br>Medical Federations<br>Citizens associations<br>Provident Institution<br>Volunteers associations |

## Expected results and discussion

Two systematic reviews will be published: the first will focus on economic and HTA evaluations of vaccination strategies or vaccines already included in vaccination schedules in Italy and Europe; the second will focus on economic and HTA evaluations of vaccination strategies or vaccines that are as yet under discussion for inclusion in current regional or national vaccination programs in Italy and Europe.

A report on the real use of economic and HTA studies in the decision-making process in Italy in recent years will be made.

Three permanent regional working boards will be set up in Northern, Central and Southern Italy, coordinated by the Project's "founding group". These will discuss the current scientific evidence of cost-effectiveness and the cost-benefit profiles of vaccinations and will also investigate other domains relevant to the decisional process (such as ethical, social and legal aspects and the impact of introducing a new vaccine on health organizations). They will also identify and analyse critical points in the different kinds of communication, in order to pick out the most appropriate methods of communication and data transfer. Furthermore, they will discern the information needs of each subject involved and work out how to improve the transfer of knowledge from researchers to decision-makers and stakeholders. The information obtained by these permanent regional boards may be validated in other Italian Regions. Lastly, a national event will be organized in order to extend the indications and information obtained at the regional level to the national level.

The feasibility of the project is guaranteed by the participation of academics and researchers in the permanent regional working boards. However, the achievement of the above-mentioned objectives might be jeopardised by the potentially scant participation of decision-makers and stakeholders in the communication network. To avoid this eventuality, every effort will be made to convince decision-makers and stakeholders of the importance of developing a communication network with scientific experts in order to foster evidence-based decisions that meet the needs highlighted by the decision-makers and stakeholders themselves.

Finally, individuation of a model of technology assessment with well-defined and objective criteria could contribute to the development of new HTA reports according to the requirements and needs evinced by the decision-makers and stakeholders themselves in the future.

## Conclusions

The success of the project will pave the way towards an interactive dialogue among all subjects involved in the decision-making process. In this way, decisions can be taken on the basis of scientific evidence, and appropriate, sustainable actions can be undertaken. Indeed, healthcare priorities can only be established through broadly based, shared evaluations of health interventions. This approach will ensure a more appropriate use of limited resources and enable preventive activities that generate many benefits and not only costs.

### References


[1] World Health Organization (WHO). *Vaccine introduction guidelines. Adding a vaccine to the national immunizaiton programme: Decision and implementation. 2005*. Available at: http://www.who.int/immunization/hpv/plan/vaccine_introduction_guidelines_who_2005.pdf

[2] Ministero della Salute. *Piano Nazionale Prevenzione Vaccinale (PNPV) 2012-14*. Gazzetta Ufficiale n. 47 del 12/03/2012 (Supplemento ordinario). Available at: http://www.salute.gov.it/imgs/C_17_pubblicazioni_1721_allegato.pdf .

[3] La Torre G, de Waure C, Chiaradia G, et al. *The future of best investing in vaccines: the Health Technology Assessment approach.* Vaccine 2008;26:1609-10.

[4] Panatto D, Bragazzi NL, Rizzitelli E, et al. *Evaluation of the economic burden of Herpes Zoster (HZ) infection.* Hum Vaccin Immunother 2015;11:245-62.

[5] Unim B, Saulle R, Boccalini S, et al. *Economic evaluation of Varicella vaccination: results of a systematic review*. Hum Vaccin Immunother 2013;9:1932-42.

[6] Boccalini S, Taddei C, Ceccherini V, et al. *Economic analysis of the first 20 years of universal hepatitis B vaccination program in Italy: an a posteriori evaluation and forecast of future benefits.* Hum Vaccin Immunother 2013;9:1119-28.

[7] Boccalini S, Bechini A, Levi M, et al. *Cost-effectiveness of new adult pneumococcal vaccination strategies in Italy*. Hum Vaccin Immunother 2013;9:699-706.

[8] Boccalini S, Azzari C, Resti M, et al. *Economic and clinical evaluation of a catch-up dose of 13-valent pneumococcal conjugate vaccine in children already immunized with three doses of the 7-valent vaccine in Italy*. Vaccine 2011;29:9521-8.

[9] Bonanni P, Boccalini S, Bechini A, et al. *Economic evaluation of varicella vaccination in Italian children and adolescents according to different intervention strategies: the burden of uncomplicated hospitalised cases.* Vaccine 2008;26:5619-26.

[10] La Torre G, de Waure C, Chiaradia G, et al. *The Health Technology Assessment of bivalent HPV vaccine Cervarix in Italy.* Vaccine 2010;28:3379-84.

[11] La Torre G, de Waure C, Chiaradia G, et al. *Guidance for future HTA applications to vaccines: the HPV lesson.* Hum Vaccin 2011;7:900-4.

[12] Damiani G, Federico B, Visca M, et al. *The impact of socioeconomic level on influenza vaccination among Italian adults and elderly: a cross-sectional study*. Prev Med 2007;45:373-9.

[13] de Waure C, Veneziano MA, Cadeddu C, et al. *Economic value of influenza vaccination*. Hum Vaccin Immunother 2012;8:119-29.

[14] Cadeddu C, De Waure C, Gualano MR, et al. *23-valent pneumococcal polysaccharide vaccine (PPV23) for the prevention of invasive pneumococcal diseases (IPDs) in the elderly: is it really effective*? J Prev Med Hyg 2012;53:101-3.

[15] Panatto D, Amicizia D, Lai PL, et al. *New versus old meningococcal group B vaccines: how the new ones may benefit infants & toddlers.* Indian J Med Res 2013;138:835-46.

[16] Gasparini R, Amicizia D, Lai PL, et al. *Clinical and socioeconomic impact of seasonal and pandemic influenza in adults and the elderly.* Hum Vaccin Immunother 2012;8:21-8.

[17] Sindoni D, La Fauci V, Squeri R, et al. *Comparison between a conventional subunit vaccine and the MF59-adjuvanted subunit influenza vaccine in the elderly: an evaluation of the safety, tolerability and immunogenicity.* J Prev Med Hyg 2009;50:121-6.

[18] Panatto D, Amicizia D, Ansaldi F, et al. *Burden of rotavirus*

*disease and cost-effectiveness of universal vaccination in the Province of Genoa (Northern Italy).* Vaccine 2009;27:3450-3.

[19] Aballéa S, Chancellor J, Martin M, et al. *The cost-effectiveness of influenza vaccination for people aged 50 to 64 years: an international model.* Value Health 2007;10:98-116.

[20] Esposito S, Gasparini R, Bosis S, et al. *Clinical and socio-economic impact of influenza and respiratory syncytial virus infection on healthy children and their households*. Clin Microbiol Infect 2005;11:933-6.

[21] Principi N, Esposito S, Marchisio P, et al. *Socioeconomic impact of influenza on healthy children and their families.* Pediatr Infect Dis J 2003;22(Suppl. 10):S207-10.

[22] Tafuri S, Fortunato F, Cappelli MG, et al. *Effectiveness of vaccination against varicella in children under 5 years in Puglia, Italy 2006-2012.* Hum Vaccin Immunother 2015;11:214-9.

[23] Bechini A, Boccalini S, Baldo V. et al. *Impact of universal vaccination against varicella in Italy.* Hum Vaccin Immunother 2015;11:63-71.

[24] Martinelli D, Pedalino B, Cappelli MG, et al. *Towards the 13-valent pneumococcal conjugate universal vaccination: effectiveness in the transition era between PCV7 and PCV13 in Italy, 2010-2013.* Hum Vaccin Immunother 2014;10:33-9.

[25] Sackett DL, Rosenberg WM, Gray JA, et al. *Evidence based medicine: what it is and what it is not.* BMJ 1996;312:71-2.

[26] Ricciardi W, La Torre G. *Health Technology Assessment. Principles dimensions and tools. SEEd. First Edition*, Turin 2010.

■ Received on September 23, 2015. Accepted on October 20, 2015.

■ Correspondence: Sara Boccalini, Department of Health Sciences, University of Florence, viale Morgagni 48, 50134 Florence, Italy - E-mail: sara.boccalini@unifi.it

ORIGINAL ARTICLE

# A Health Technology Assessment: laparoscopy versus colpoceliotomy

A. DAMONTI[1], L. FERRARIO[1], P. MORELLI[2], M. MUSSI[1], C. PATREGNANI[3], E. GARAGIOLA[1], E. FOGLIA[1], R. PAGANI[1], R. CARMINATI[3], E. PORAZZI[1]

[1] Carlo Cattaneo University, LIUC, Centre for Research on Health Economics, Social and Health Care Management, CREMS, Castellanza, Italy; [2] Public Hospital of Novi Ligure, Italy; [3] Department of Obstetrics and Gynecology of the Civil Legnano Hospital, Italy

Key words

Colpoceliotomy • Laparoscopy • HTA

Summary

**Introduction**. *The objective of this paper is the comparison between two different technologies used for the removal of a uterine myoma, a frequent benign tumor: the standard technology currently used, laparoscopy, and an innovative one, colpoceliotomy. It was considered relevant to evaluate the real and the potential effects of the two technologies implementation and, in addition, the consequences that the introduction or exclusion of the innovative technology would have for both the National Health System (NHS) and the entire community.*

**Methods**. *The comparison between these two different technologies, the standard and the innovative one, was conducted using a Health Technology Assessment (HTA). In particular, in order to analyse their differences, a multi-dimensional approach was considered: effectiveness, costs and budget impact analysis data were collected, applying different instruments, such as the Activity Based Costing methodology (ABC), the Cost-Effectiveness Analysis (CEA) and the Budget Impact Analysis (BIA). Organisational, equity and social impact were also evaluated.*

**Results**. *The results showed that the introduction of colpoceliotomy would provide significant economic savings to the Regional and National Health Service; in particular, a saving of € 453.27 for each surgical procedure.*

**Discussion**. *The introduction of the innovative technology, colpoceliotomy, could be considered a valuable tool; one offering many advantages related to less invasiveness and a shorter surgical procedure than the standard technology currently used (laparoscopy).*

## Introduction

Uterine myoma is the most common benign tumour of the uterus [1], affecting around 20-25% of women over 30 years old [2]. The incidence of this disease increases from age 40 to 50.

This disease is usually asymptomatic; sometimes, however, it may cause disorders such as dysmenorrhea, compromising women's health and quality of life. Symptoms may be severe enough to require treatment.

There are several options for the treatment of myomas, including both medication and surgical procedures, such as myomectomy and hysterectomy. Surgical and other invasive interventions still dominate treatment [3], while the medical therapy is currently limited to the preoperative reduction of symptoms [4].

Myomectomy is the only surgical option for women of childbearing age. With myomectomy, it is necessary to consider both laparoscopy and colpoceliotomy, which, in this paper, is defined as "innovative technology".

Laparoscopy was first described by Semns in 1980 [5]. It is a surgical procedure in which a small incision is made, through which a viewing tube (laparoscope) is inserted. Colpoceliotomy is a surgical procedure in which an opening of the pelvic cavity is made through the fornix of the vagina. The latter procedure is less invasive than the standard one; there is no need of any abdominal incision and the peritoneum is surgically opened via the vaginal fornix; furthermore colpoceliotomy can be useful even in case of large, numerous and intramural fibroids [6]. However, although the first description of this method dating from 1951, thanks to De Bedoya [7], it is an uncommon methodology as it requires a detailed knowledge of the vaginal apparatus and manual transvaginal skills. It is, therefore, a rarely used technique in the removal of uterine myomas. Over the last decades, this methodology has been studied by other authors [8-9], asserting that its surgical access to pelvis and, in general, its surgical time are more rapid if compared with laparoscopy one. However, despite these advantages, gynecologists do not often use vaginal myomectomy [10].

The aim of the present study is to compare laparoscopy with colpoceliotomy, focusing on the advantages and benefits of the innovative technology, based on their use in the Department of Obstetrics and Gynecology of the Health Authority Civil Hospital of Legnano. The final goal of the analysis conducted is the definition of the best procedure to be implemented and used in this specific setting, in accordance with a multidimensional and multi-disciplinary Health Technology Assessment (HTA) approach, considering the Health Authority point of view.

## Materials and methods

In order to achieve the above mentioned objective, HTA was considered the most acknowledged tool for being adopted in the decision-making phase, in professional and knowledge-intensive settings, such as hospitals.
HTA is a multi-disciplinary tool, one that aims at evaluating both the real and the potential effects of technologies, and the consequences that the introduction or the exclusion of a procedure has for the health system, the economy and society.
The present approach is able to analyse different technologies, by examining their economic, social, clinical, ethical and organizational implications [11], thus identifying methodologies that offer a greater benefit to the population. The primary objective of HTA, is not to increase specialists and evaluators' knowledge, but to directly influence the decision making process [12], with an evidence-based, more quantitative and objective approach.
The present paper compares laparoscopy with colpoceliotomy through the implementation of a Hospital-Based HTA, namely IMPAQHTA model (*Implementation of a quick Hospital-based HTA*), assuming the Health Authority Civil Hospital of Legnano perspective.
The proposed IMPAQHTA framework was redesigned, based on exiting models: 1) Core Model [13], because of the completeness of its dimensions; 2) Multi Criteria Decision Analysis [14], because of the quantitative scoring methods evaluation; 3) Lombardy Region Model [15], because of its alignment with policies, laws and country-oriented setting.
The IMPAQHTA framework [16] identified 8 dimensions to be used in the assessment phase: *i)* general relevance; *ii)* safety; *iii)* efficacy; *iv)* effectiveness; *v)* economic and financial impact; *vi)* equity; *vii)* legal, social and ethic impact; and *viii)* organizational impact. The above mentioned dimensions could be evaluated using 13 specific quantitative metrics.
The implemented framework is composed by three distinct logical phases:
*i)* prioritisation of the 8 dimensions of analysis;
*ii)* evaluation of the technologies, thanks to the support of 13 quantitative criteria for the complete quantitative assessment and the production of a final report;
*iii)* determination of a concise result, for the final appraisal of the evaluated innovations.
With regards to the first phase, the 8 dimensions composing the framework are prioritized by specific evaluators, in accordance with the VTS-HTA Lombardy Region model [15]. Thus, the dimensions are prioritised in order to show their relative importance through a numerical value, following a rating scale from 1 to 8 (1 = less important and 8 = more important), using the evaluators' judgment in order to define which is the most important dimension.
A clarification is needed. The chief of the Department of Obstetrics and Gynaecology (who was currently using the innovative technology) needed to evaluate all the impact related to the introduction of Colpoceliotomy in the ordinary clinical practice, thus requiring the judgement of the Strategic Board. The prioritization phase involved 5 members of the Hospital Authority, acting as evaluators.
The second phase consists of a detailed analysis of the dimensions; *i)* identification of different evaluation sub-dimensions; *ii)* attribution to each of them, a three level rating score with 1 (less performant), 2 (equal performant) and 3 (more performant) was applied, in accordance to the Mitton Model [17].
For the complete technology assessment, developing the IMPAQHTA framework, a multi-disciplinary HTA team was required: 6 volunteers, with a certified knowledge and skills on HTA, composed an evaluating team and produced an unabridged IMPAQHTA report, useful for the Strategic Board, to take evidence-based choices during the appraisal phase.
The third phase, according to evidence in the literature, regarding EVIDEM Core Model and Multi-Criteria Decision Analysis, leads to the final synthesis, in order to compare the technologies evaluated. The higher the final rating is, the more preferable the technology is. This final evaluation was conducted by the Strategic Board of the reference Hospital, with the inclusion of the Chief of Department of Obstetrics and Gynaecology.
Moving on from these premises, the evaluation of the above mentioned 8 dimensions was conducted by the multi-disciplinary team, with the development of 13 sub-dimensions using specific operative and quantitative tool (Tab. I).
It is important to clarify that the efficacy dimension was not taken into consideration in the present analysis, because the proposed evaluation was a specific need of the Health Authority of reference, thus grounding the assessment on real data referring to the year 2013.
Safety is a relevant dimension to be investigated, one that leads to the evaluation of adverse events, mortality or morbidity, related to the technologies under assess-

**Tab. I.** Details of the dimensions and the related sub-dimension of the IMPAQHTA model.

| Dimensions | Sub-dimensions |
|---|---|
| General relevance | Quality of scientific evidences |
| | Description of the pathology and the related technologies |
| Safety | Seriousness of Adverse Events (mild, moderate or severe adverse events) |
| Efficacy | Efficacy data |
| Effectiveness | Effectiveness data |
| Economic financial Impact | ABC |
| | Health Economic Evaluation |
| | Budget Impact Analysis |
| Equity | Equity data |
| Legal, social and ethic impact | Legal aspects |
| | Social and ethical impact |
| Organizational Impact | Quantitative impact |
| | Qualitative impact |

ment. In particular, it allows the identification of mild/moderate/severe adverse events, considering the population treated with the technology. Since the lack in literature of this specific information, safety data were collecting through specific interviews to clinicians, in relation to their own clinical experience. In particular, safety information refer to all the possible consequences on the patient, related to the use of the innovative technology, in terms of *a)* seriousness of adverse events; *b)* invasiveness of the procedure; *c)* safety in its implementation.
Equity, legal and organizational impacts were investigated using specific questionnaires, completed by both clinicians and patients, in accordance with the items composing the EUnetHTA Core Model. For instance, the equity impact aimed at investigating *a)* access to care on a local level; *b)* access to care for person of a legally protected status; *c)* hospital waiting list improvement and *d)* the invasiveness of the innovative technology.
The legal and the social impact aimed at evaluating both the patients' satisfaction and the related productivity loss.
Furthermore, the organizational impact had the objective to define the perception of the clinicians involved in the innovative procedure and to quantify all the investments needed if organizational changes occurred: *a)* additional people; *b)* additional room; *c)* training course; *d)* meeting and *e)* learning time of the implementation of the innovation, *f)* investment in equipment; *g)* update of the existing equipment; *h)* impact on the internal and the purchasing processes. According to this, all the items composing the above mentioned dimensions have been evaluated by the multi-disciplinary team (with the support of clinicians), with the rating scale proposed by Mitton and colleagues in 2011, as previously mentioned.
With regard to the economic and financial dimension, the method used for the enhancement of the average cost of the two technologies was the *Activity Based Costing analysis* (ABC), which measured the costs and performances of each activity. The concept of ABC was first defined in 1980 by Cooper and Kaplan [18] and it focused on the activity useful for the final output of the offered service; in particular, it gives the cost of a specific product (in this case, a specific technology), according to the activities through the use of cost drivers [19]. The process consists of the following stages [18, 20-21]: *i)* activities' identification; *ii)* definition of the activities' cost; *iii)* definition of the activities' cost drivers; *iv)* definition of the cost drivers' volume; *v)* definition of a unit cost, per cost driver, for each activity and *vi)* calculation of the unit cost per procedure.
Another methodology applied to compare the two technologies is the *Cost-effectiveness analysis* (CEA), an economic evaluation in which the costs of alternative procedures are compared using outcomes measures expressed in natural units [22]. In order to implement the CEA it is necessary to have suitable measures of effectiveness, because this technique expresses health benefits in simple terms, such as years of life gained [23]. From a methodological point of view, CEA is divided into five stages: *i)* definition of the program; *ii)* computation of net costs; *iii)* computation of net health effects; *iv)* application of decision rules; *v)* sensitivity analysis [23]. This method is often implemented in an HTA report in order to obtain a specific effectiveness with minimum costs.
Whereas a CEA evaluates both costs and outcomes of alternative technologies over a specified time horizon in order to estimate their economic efficiency, a *Budget Impact Analysis* (BIA) is based on their affordability. In fact, its main purpose is to predict the final consequences of the adoption and diffusion of a new technology into a healthcare system with finite resources [24].

## Results

### The sample

The analysis was conducted in accordance with the real data performed by the Department of Obstetrics and Gynaecology of the Health Authority Civil Hospital of Legnano, during the year 2013.
In particular, it emerged that 166 women were treated for the removal of uterine myomas (Tab. II).
The table above shows that the two populations under analysis (118 patients in the arm of colpoceliotomy and 48 patients in the arm of laparascopy) were comparable from both the dimensions and the diameters of the myomas. It could be considered a relevant feature, allowing the comparability of the two population and the related technologies used in this category of procedure.

### Results from the IMPAQHTA model implementation

The first step of the HTA is the prioritisation of the dimensions, by the chief of the Department and the Strategic Board of the Hospital of reference, involving a total of 5 individuals as evaluators. The table below shows that the most important variable for the experts is the patient's safety (Tab. III).
The second most relevant dimension is the economic and financial impact. When the ABC analysis is implemented, the technologies are perfectly super-imposable in terms of the process' description; however, the spe-

**Tab. II.** Description of the sample under assessment.

| | Colpoceliotomy | Laparoscopy | P-value |
|---|---|---|---|
| N. of patients | 118 | 48 | |
| Mean age [years] | 41 | 39 | > 0.005 |
| Length of the procedure [min] | 85 | 80 | > 0.005 |
| Myoma diameters [cm] | 7 | 6 | > 0.005 |
| Myoma dimensions [gr] | 100 | 200 | 0.01 |
| Length of hospitalization | 3 | 5 | 0.000 |
| Adverse events | 1,60% | 6% | 0.000 |

**Tab. III.** Prioritisation.

| Dimensions | Evaluators | | | | | Total | Normalisation |
|---|---|---|---|---|---|---|---|
| | # 1 | # 2 | # 3 | # 4 | # 5 | | |
| Safety | 3 | 1 | 1 | 1 | 2 | 1 | 0.222 |
| Economic and financial impact | 1 | 2 | 3 | 3 | 1 | 2 | 0.194 |
| Effectiveness | 2 | 5 | 4 | 5 | 3 | 3 | 0.167 |
| Organisational impact | 5 | 3 | 5 | 2 | 5 | 4 | 0.139 |
| Efficacy | 7 | 7 | 2 | 4 | 4 | 5 | 0.111 |
| General relevance | 6 | 6 | 6 | 6 | 6 | 6 | 0.083 |
| Equity | 4 | 8 | 7 | 7 | 7 | 7 | 0.056 |
| Social and ethical impact | 8 | 4 | 8 | 8 | 8 | 8 | 0.028 |

cific and related operative phases of the intervention are different (Tab. IV).

The table above shows the differences, between each phase, in terms of costs and impact. It emerges that laparoscopy absorbs economic resources for a total of € 1,789.42, whereas colpoceliotomy a total of €1,336.15, allowing a €453.72 saving for each surgery. This difference may be explained by the expensive instruments used for laparoscopy.

The CEA was calculated as the "ratio" between the unit cost per patient related to the two compared technologies (derived from the previously conducted ABC Analysis) and the effectiveness data. The effectiveness data is defined as the percentage of "non-complications surgery". It emerges that in the colpoceliotomy arm only 1.60% of patients reported problems during surgery.

On the contrary, in the laparoscopy arm 4% of patients reported the onset of mild and moderate adverse events. As previously mentioned, the effectiveness data used for the CEA have been collected from an observational study involving 166 patients in total (see Table II)

In accordance with the real data related to the year 2013, Table V shows that the innovative procedure achieved a higher cost-effectiveness value if compared with the gold standard technology. Colpoceliotomy, on the basis of an increase in the effectiveness data, leads to a decrease of total costs, thus being defined as the dominant strategy.

In the BIA, two scenarios are considered: the first one calculates the annual cost of 118 laparoscopy surgeries and 48 standard procedures, taking into account what actually happened in the year 2013 within the reference Health Authorities; the second one estimates the annual cost of 166 surgeries, if done with the standard technology (it could be considered as the "baseline" scenario).

According to this scenario, the BIA, with the inclusion of both the results deriving from the activity based costing approach, and the organisational investment needed, leads to an overall significant financial and economic saving for the Health Authority of reference, in case of implementation of colpoceliotomy procedure in the clinical practice (with regards to the year 2013).

An economic saving of € 48,955.06 emerged from the evidence, in the first scenario: the data confirmed that colpoceliotomy could be introduced in the clinical practice of the referred hospital. This difference takes into account also the training costs required by the introduction of colpoceliotomy.

This analysis was implemented over a 12-month period due to the fact that the present health technology evaluation was requested by the Strategic Management Board

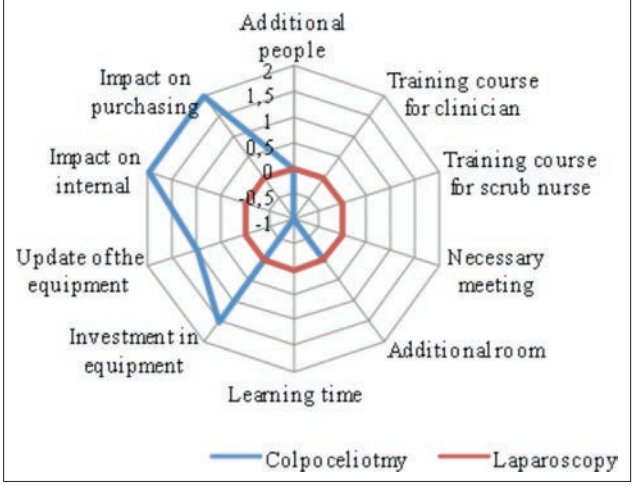
**Tab. IV.** ABC analysis.

| Phases | Laparoscopy | | Colpoceliotomy | |
|---|---|---|---|---|
| Pre-Hospitalisation | 305.08 € | 17% | 305.08 € | 23% |
| Admission | 65.62 € | 4% | 65.62 € | 5% |
| Pre-Surgery Recovery | 38.93 € | 2% | 38.93 € | 3% |
| Surgery | 1,089.64 € | 61% | 636.37 € | 48% |
| Post-Surgery | 201.83 € | 11% | 201.83 € | 15% |
| Discharge | 88.33 € | 5% | 88.83 € | 7% |
| Total | 1,789.42 € | 100% | 1,336.15 € | 100% |

**Tab. V.** Cost-effectiveness and Budget Impact Analysis.

| Cost-Effectiveness Value | CCT | STD |
|---|---|---|
| ABC Analysis | €1,336.15 | €1,789.42 |
| Effectiveness data | 98.40% | 94% |
| CEV | 1,357.87 | 1,903.64 |
| Budget Impact Analysis | CCT | STD |
| ABC Analysis | €1,336.15 | €1,789.42 |
| # surgery per year | 118 | 48 |
| Surgery cost per year | € 243,557.30 | € 297,043.55 |
| Further Training cost | €3,870.00 | |
| | € 660.80 | |
| Total Training cost | €4,530.80 | |
| BIA | €248,088.10 | €297,043.55 |
| Δ BIA | -€48,955.45 | €/year |

Fig. 1. Organizational impact.




of the Hospital Authority, and the financial period of reference for the budget was equal to 12 months.

Additional training courses are relevant from an organizational point of view. In fact, clinicians have to be trained in order to implement the innovative procedure.

The figure shows that colpoceliotomy is preferable from an organizational point of view. A clarification is need: for a proper reading of the picture, it is important to take into consideration the two areas (the wider the area is, the more preferable the technology is).

However, specific training courses are needed for clinicians, scrub nurses and support staff for learning the correct procedure for colpoceliotomy, in terms of *i)* adequate preparation of the women undergoing the intervention, *ii)* proper provision of all the instruments required and *iii)* correct implementation of the innovative technology.

In particular, it emerged that in the specific setting under assessment, 2 surgeons, 3 scrub nurses and 5 health professionals needed to attend a 20-hours training course.

With respect to the last phase of the IMPAQHTA framework implementation, it is important to determine a final score, both for laparoscopy and colpoceliotomy, useful for an evidence-based decision-making appraisal phase.

It is relevant to specify the sub-criteria of each evaluated dimension, starting from the prioritisation shown in Table I. Sub-criteria are detailed in order to appoint a basis score for each of them.

The following table shows the final score of the present comparative study, assigned by all the 5 evaluators that have been involved in the prioritization phase (Tab. VI) In this view, the chief of the Department of Obstetrics and Gynaecology and the Strategic Board assigned (in columns three and four) a rating score for each subcriteria, taking into account all the possible differences related to the two technologies under assessment. In column five and in the final row defining each dimension, the maximum total achievable score is reported.

These data allow the calculation of the effect of each sub-dimension through a ratio between the maximum achievable total score and the established total score.

Finally, it is useful to show the sum of the normalised scores per each technology (normalized score x normalized prioritization), resulting a total score for colpoceliotomy and a total score for the standard procedure. The table shows that colpoceliotomy achieves a higher score than laparoscopy (0.56 *vs* 0.44).

## Discussion

As shown in Table VI, colpoceliotomy achieves a higher score than laparoscopy. The innovative technology presents a lower annual cost and it brings significant economic savings for health care organisations. It is relevant to note the importance of an appropriate and rational implementation of the available technologies, in order to achieve maximum benefits with minimal costs.

However, the introduction of a new technology affects the Health Authority of reference from an organizational point of view, requiring coaching and training courses for the persons involved in the procedure (as mentioned in the previous section). Then, it is necessary to have frequent meetings aimed at communicating the changes introduced by the new technical surgery in the whole organization.

The adoption of colpoceliotomy does not, however, require a purchase of new surgical instruments or equipment. Therefore, the innovative technology could lead to an economic saving due to the lack of need for expensive instruments, such as Trocar. The new procedure also allows a saving of time with some activities, such as the purchase or the maintenance of equipment; this may result in a positive impact on the internal process of gynaecology units and on health workers' safety.

If colpoceliotomy were inserted in the clinical setting, it would have a positive impact also on the access to care, thus positively affecting the equity dimension (average score equal to 2.2). In fact, the innovative procedure would enlarge the treated "target" population, including persons of a legally protected status; thanks to the use of natural orifices, colpoceliotomy is less invasive and specific physical conditions of eligibility are not required.

In addition, colpoceliotomy has a relevant social impact on a patient's life due to a shorter hospitalisation; this means that patients may return to their daily life and work sooner, thus reducing productivity loss.

From the present analysis, it emerged that the average score declared by the expert for the social dimension is equal to 2.5 for colpoceliotomy; the innovative technology could improve the patient's autonomy after the procedure.

However, there is a disadvantage in the implementation of colpoceliotomy: it requires an accurate manual ability, which implies a longer learning time of the new surgical procedure, thus needing coaching and training periods. This has been reflected in the evaluation of the

**Tab. VI.** HTA comparative study.

| | | Basis score | | | | Normalsed score | | | Final score | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Dimensions** | **Sub – dimensions** | **CCT** | **STD** | **TOT** | **Incidence** | **CCT** | **STD** | **Normalized prioritization** | **CCT** | **STD** |
| Safety | | 2 | 2 | 6 | 1.5 | 0.5 | 0.5 | | | |
| | | 4 | | | | 0.5 | 0.5 | 0.22 | 0.11 | 0.11 |
| Economic and financial impact | Activity Bases Costing Analysis | 2 | 1 | 6 | 0.7 | 0.2 | 0.1 | | | |
| | Cost-effectiveness Analysis | 2 | 1 | 6 | 0.7 | 0.2 | 0.1 | | | |
| | Budget Impact Analysis | 2 | 1 | 6 | 0.7 | 0.2 | 0.1 | | | |
| | | 9 | | | | 0.7 | 0.3 | 0.19 | 0.13 | 0.06 |
| Effectiveness | | 2 | 1 | 6 | 2.0 | 0.7 | 0.3 | | | |
| | | 3 | | | | 0.7 | 0.3 | 0.17 | 0.11 | 0.06 |
| Organisational impact | Quantitative Impact | 2 | 1 | 6 | 0.9 | 0.3 | 0.1 | | | |
| | Qualitative Impact | 2 | 2 | 6 | 0.9 | 0.3 | 0.3 | | | |
| | | 7 | | | | 0.6 | 0.4 | 0.14 | 0.08 | 0.06 |
| Efficacy | | 2 | 2 | 6 | 1.5 | 0.5 | 0.5 | | | |
| | | 4 | | | | 0.5 | 0.5 | 0.11 | 0.06 | 0.06 |
| General relevance | Consistency of evidence | 1 | 3 | 6 | 0.3 | 0.0 | 0.1 | | | |
| | Description of technology and comparator | 1 | 2 | 6 | 0.3 | 0.0 | 0.1 | | | |
| | Safety of the new technology and comparator | 2 | 2 | 6 | 0.3 | 0.1 | 0.1 | | | |
| | Target Population | 2 | 1 | 6 | 0.3 | 0.1 | 0.0 | | | |
| | Consistency of the objectives with the adopted strategy | 2 | 2 | 6 | 0.3 | 0.1 | 0.1 | | | |
| | Potential advantaged areas | 2 | 2 | 6 | 0.3 | 0.1 | 0.1 | | | |
| | | 22 | | | | 0.5 | 0.5 | 0.08 | 0.04 | 0.05 |
| Equity | | 2 | 3 | 6 | 1.2 | 0.4 | 0.6 | | | |
| | | 5 | | | | 0.4 | 0.6 | 0.06 | 0.02 | 0.03 |
| Social and ethical impact | Ethical Impact | 3 | 2 | 6 | 0.6 | 0.3 | 0.2 | | | |
| | Social Impact | 3 | 2 | 6 | 0.6 | 0.3 | 0.2 | | | |
| | | 10 | | | | 0.6 | 0.4 | 0.03 | 0.02 | 0.01 |
| | | | | | | | | **Results** | **0.56** | **0.44** |

organizational impact (see Fig. 1): the items related to training course needed for colpoceliotmy achieved a lower score than the standard procedure. The present feature has a substantial impact only on a short term period, because training courses required only 20 hours for person. In general, it emerges that the introduction of colpoceliotmy has a positive organizational impact, achieving an average score equal to 2.50.

Although an HTA study is a useful tool for decision makers, who are involved every day in many different strategic and tactical decisions [25], it is not the only one. In fact, the results and evaluation of an HTA are relevant to decision making only if they are aligned with the mission of the health care organization.

## Conclusions

Colpoceliotomy is an alternative technique to laparoscopy, the latter being the standard method used for the treatment and removal of uterine myomas. At present, colpoceliotomy is still uncommon because it requires a high knowledge of the anatomy of female genitalia (in particular the vaginal canal), high experience and manual ability. Furthermore, the innovative procedure meets institutional constraints that limit its implementation.

In the present study, an HTA evaluation was made, which compared the use of traditional laparoscopy and colpoceliotomy for the removal of uterine myomas. The results showed that the innovative technology is more advanta-

geous in terms of invasiveness and shorter surgery time; thus, this procedure is low-time consuming [26].
The study highlights both the social and the financial impact of colpoceliotomy, considering also the organisational impact on health care companies, represented by the introduction of coaching and training courses both for the medical staff and the support personnel involved in the surgery.
Data show that the new procedure is safe, and that it provides both short and long term benefits in terms of the savings in the purchase and maintenance of machinery.
The innovative technology leads to a significant reduction in adverse events, invasiveness, duration of surgery and post-operative hospitalization.
Considering the economic and financial perspective, the new procedure would lead to a substantial reduction of costs related to the absence of sophisticated equipment, as required by laparoscopy, and to a smaller number of personnel involved.
Therefore, colpoceliotomy may be considered a valuable tool in reducing costs regarding the surgical treatment for uterine myomas.
In conclusion, although the present study highlights its technical superiority; that is, the advantages it ensures over prior or existing technologies [27], a further randomised study is suggested in order to confirm colpoceliotomy's benefits achieved in the present study.

## References


[1] Birsan A, Deval B, Detchev R, et al. *Vaginal and laparoscopic myomectomy for large posterior myomas: results of a pilot study*. Eur J Obstet Gynecol Reprod Biol 2003;110;89-93.

[2] Davies A, Hart R, Magos AL. *The excision of uterine fibroids by vaginal myomectomy: a prospective study*. Fertil Steril 1999;71:961-4.

[3] Stewart EA. *Uterine fibroids and evidence-based medicine. Not an oxymoron*. N Engl J Med 2012;366:471-3.

[4] Lethaby A, Vollenhoven B, Sowter M. *Pre-operative GnRH analogue therapy before hysterectomy or myomectomy for uterine fibroids*. Cochrane Database Syst Rev 2001;2.

[5] Semm K, Mettler L. *Technical progress in pelvic surgery via operative laparoscopy*. Am J Obstet Gynecol 1980;138:121-7.

[6] Carminati R, Ragusa A, Giannica R, et al. *Anterior and posterior vaginal myomectomy: a new surgical technique*. Med Gen Med 2006;8:42.

[7] De Bedoya Y, Hexia G. *Vaginal myomectomy*. Gynecol Prat 1951;2:486.

[8] Bognoni V, Quartuccio A Jr, Quartuccio A. *Therapeutic posterior colpo-celiotomy. Authors' experience*. Minerva Ginecol 1999;51:153-5.

[9] Magos AL, Bournas N, Sinha R, et al. *Vaginal myomectomy*. BJOG 1994;101:1092-4.

[10] Plotti G, Plotti F, Di Giovanni A, et al. *Feasibility and safety of vaginal myomectomy: a prospective pilot study*. J Minim Invasive Gynecol 2008;15;166-71.

[11] Drummond MF, Schwartz JS, Jönsson B, et al. *Key principles for the improved conduct of health technology assessments for resource allocation decisions*. Int J Technol Assess Health Care 2008;24:244-58.

[12] Drummond MF. *Principles of Economic Appraisal in Health Care*. Oxford: Oxford University Press 1980.

[13] EUnetHTA. *HTA Core Model for medical and surgical interventions* 1.0r, 2008. Available online at: http://www.eunethta.eu/outputs/hta-core-model-medical-and-surgical-interventions-10r

[14] Thokala P. *Multicriteria Decision Analysis for Health Technology Assessment*, 2011. Available online at: http://www.nicedsu.org.uk/MCDA%20for%20HTA%20DSU.pdf

[15] Radaelli G, Lettieri E, Masella C, et al. *Implementation of Eunethta core model® in Lombardia: The VTS framework*. Int J Technol Assess Health Care 2014;30:105-12.

[16] Pagani R, Porazzi E, Manzini R, et al. *Assessing Innovative Healthcare Technologies in Hospitals: Lessons Learnt From an Empirical Experiment*. RnD Management Conference 2015.

[17] Mitton C, Dionne F, Damji R, et al. *Difficult decisions in times of constraint: Criteria based Resource Allocation in the Vancouver Coastal Health Authority.* BMC Health Serv Res 2011;11:169.

[18] Cooper R, Kaplan RS. *How Cost Accounting Distorts Product Costs*. Manag Account 1988.

[19] Bogdanoiu, C. *Activity Based Costing from the perspective of competitive advantage*. J Appl Econ Sci 2009;1;5-11.

[20] Bruggeman W, Everaert P. *Time-driven Activity-based Costing: exploring the Underlying Model*. Cost Management 2009;21;16-20.

[21] Krumweide KR. *The Implementation Stages of ABC and the Impact of Contextual and Organizational Factors*. J Manag Account Res 1988;10;239-78.

[22] Robinson R. *Cost Effectiveness analysis*, BMJ 1993;307:793-5.

[23] Shepard S, Thompson M. *First principles of Cost-Effectiveness analysis in Health*, Public Health Rep 1979;94;535-43.

[24] Mauskopf JA, Sullivans SD, Annemans L, et al. *Principles of good practice for budget impact analysis: report of the ISPOR task force on good research practices-budget impact analysis*. Value in Health 2007;10:336-47.

[25] Lazzarotti V, Manzini R, Pizzurno E. *Setting up innovation networks: the case of a small company in the calibration industry*. Asia Pac Manag Rev 2012;17;281-300.

[26] Carminati R. *La chirurgia vaginale oltre la tradizione*. Padova: Piccin 2009.

[27] Chiesa V, Frattini F, Gilardoni E, et al. *Searching for factors influencing technological asset value*. Eur J Innovation Manag 2007;10:467-88.


■ Received on July 31, 2015. Accepted on October 23, 2015.

■ Correspondence: Lucrezia Ferrario, corso Matteotti 22, 21053 Castellanza (VA), Italy - Tel. +39 0331 572504 - Fax +39 0331 572513 - E-mail: lferrario@liuc.it

ORIGINAL ARTICLE

# Impact of different scoring algorithms applied to multiple-mark survey items on outcome assessment: an in-field study on health-related knowledge

A. DOMNICH, D. PANATTO, L. ARATA, I. BEVILACQUA, L. APPRATO, R. GASPARINI, D. AMICIZIA
Department of Health Sciences, University of Genoa, Italy

Key words

Multiple-mark items • Multiple answer items • Pick-N items

## Summary

**Introduction.** *Health-related knowledge is often assessed through multiple-choice tests. Among the different types of formats, researchers may opt to use multiple-mark items, i.e. with more than one correct answer. Although multiple-mark items have long been used in the academic setting – sometimes with scant or inconclusive results – little is known about the implementation of this format in research on in-field health education and promotion.*
**Methods.** *A study population of secondary school students completed a survey on nutrition-related knowledge, followed by a single-lecture intervention. Answers were scored by means of eight different scoring algorithms and analyzed from the perspective of classical test theory. The same survey was re-administered to a sample of the students in order to evaluate the short-term change in their knowledge.*
**Results.** *In all, 286 questionnaires were analyzed. Partial scoring algorithms displayed better psychometric characteristics than the dichotomous rule. In particular, the algorithm proposed by Ripkey and the balanced rule showed greater internal consistency and relative efficiency in scoring multiple-mark items. A penalizing algorithm in which the proportion of marked distracters was subtracted from that of marked correct answers was the only one that highlighted a significant difference in performance between natives and immigrants, probably owing to its slightly better discriminatory ability. This algorithm was also associated with the largest effect size in the pre-/post-intervention score change.*
**Discussion.** *The choice of an appropriate rule for scoring multiple-mark items in research on health education and promotion should consider not only the psychometric properties of single algorithms but also the study aims and outcomes, since scoring rules differ in terms of biasness, reliability, difficulty, sensitivity to guessing and discrimination.*

## Introduction

Knowledge of topics related to lifestyle, health and healthcare may guide people's health-related choices and determine their health status. Assessment of these issues is growing, as any inadequacies identified can be targeted by specifically designed health promotion interventions [1]. Health-related knowledge may be seen as a part of health literacy [2], which is a broader concept defined as "the constellation of skills, including the ability to perform basic reading and numerical tasks required to function in the healthcare environment" [3]. Health literacy is considered a priority public health goal [4], as its low level is a well-known predictor of poor health-related outcomes (reviewed in [5]).

Factual knowledge on health-related topics is usually assessed by means of questionnaires and, in particular, multiple-choice tests or quizzes. Of the different types of question formats, multiple-choice (type-A or one-out-of-N items, i.e. with the single best option) and true-false items are among the most widely used to assess health-related knowledge. The widespread use of these two formats is probably due to such characteristics as their objectivity, rapidity in testing numerous subjects and receiving respondents' feedback, and the possibility of automatic scoring [6]. Alongside these strengths, however, multiple-choice questions also have weaknesses [6], such as reduced validity due to the possibility (or even encouragement) of guessing [7] and their failure to distinguish between partial knowledge and absence of knowledge [8, 9]. Alternative item formats may partially solve the shortcomings of type-A tests. Among these, multiple true-false (MTF, also known also as type-X) and multiple-mark (MM) items with several correct answers (also dubbed multiple choice multiple answer or pick-N items) have been extensively studied [10-19]. MTF tests, in which the respondent classifies each option as a separate true-false statement [14], are somewhat similar to MM items, in which the correct options chosen are regarded as true options while unmarked distracters are classed as false options. Indeed, Cronbach [10] has established that MTF and MM formats are very similar in terms of reliability, validity and respondents' performance. In certain situations, the MM format may be advantageous in terms of item construction, in that it allows more natural wording of both questions and response options and may need fewer distracters [15]. Pomplun and Omar [13] have demonstrated that MM questions are a feasible objective format and display acceptable reliability and validity, while Berk [20] has underlined that the MM

format preserves the main qualities of the type-A format while at the same time quantifying complex cognitive outcomes by assessing respondents' lines of reasoning in selecting answers. Moreover, MM items are useful in evaluating people with average and above-average knowledge of a topic [12].

One of the main issues regarding the MM format is the choice of an appropriate scoring rule. The most computationally simple scoring algorithm (SA) is the so-called dichotomous rule, whereby the respondent gets the full score for all correctly marked options, but nothing otherwise. An important drawback of the dichotomous SA, however, is its inability to give credit for partial knowledge; a respondent who gets all but one answer correct obtains the same score as one who is unable to provide any correct answer or even selects all wrong answers [15, 21, 22]. Indeed, in research on health education and promotion, laypeople's knowledge of health-related topics is often dubbed as partial knowledge.

In recent years, several SAs that are able to award partial credit, with or without penalties for guessing, have been developed and studied [12, 15, 16, 18, 19]. However, the results of these studies have often been inconsistent. Thus, Hsu et al. [12] established that none of the six SAs used in their study was significantly better than the others, while a partial-credit SA developed by Ripkey et al. [16] proved to be superior to the dichotomous SA in terms of item difficulty and discrimination parameters. These latter findings were later replicated by Bauer et al. [19], who documented the superiority of two different partial SAs to the dichotomous SA. More recently, the balanced SA, specifically designed for MM items, has been proposed as an improvement on Ripkey's algorithm [18].

Most of the above-mentioned studies were carried out in the academic setting in order to evaluate students' performances in exams and find an optimal item format. However, little is known about how different scoring rules applied to MM survey items would affect the evaluation of health-promotion outcomes. The present study therefore aimed to evaluate whether the choice of a scoring rule could impact on the evaluation of findings. Specifically, we posed two research questions: (1) do the psychometric properties of different SAs applied to the evaluation of factual health-related knowledge differ? and (2) do different SAs applied to the evaluation of factual health-related knowledge impact on the outcome?

## Methods

### Study design and setting

The nutrition-related knowledge of students from seven secondary schools in the Genoa metropolitan area was assessed in 2012/2013 by means of a self-administered paper-and-pencil survey. Participation was voluntary and anonymity was assured. No time limit was placed on compilation of the questionnaire, though students took less than 20 min.; survey administration was strictly supervised in order to prevent cheating. The study and the test were approved by the boards of each school.

This initial assessment of nutrition-related knowledge was followed by a single interactive lecture given by appropriately trained medical staff accompanied by teachers. The lecture lasted approximately 45 minutes and covered both general food- and nutrition-related topics (e.g. healthy diet, dietary recommendations, notions of macro- and micronutrients) and questions frequently asked by the students during the pre-intervention survey administration.

To evaluate changes in knowledge scores, the same survey was re-administered to a sample of students 2 weeks after the lecture.

### Survey instrument for assessing nutrition-related knowledge

The factual nutrition-related knowledge part of the survey consisted of 14 items. Two knowledge items were excluded from the analysis, as formal flaws (poor specification of questions) were detected after survey administration; a total of 12 items were therefore analyzed. The survey also contained 7 perceived knowledge items (such as, *Do you know what carbohydrates are?*) and 2 open-ended items (such as, *What would you like to know about nutrition?*). These items were introduced after agreement with teachers, in order to plan the content of the upcoming lecture and of future school-based health-promotion interventions, and were not analyzed in the present study. Conceptually, the survey consisted of two nutrition-related topics, namely the understanding of food terms and the main sources of nutrients. Two formats were adopted: 9 items were MM, while the remaining 3 were type-A. The items did not conform to a single pattern; among the MM items, the number of options ranged from 4 to 8, the number of correct options from 2 to 5, and the number of distracters from 1 to 5. The type-A items had 2 or 3 distracters. To discourage guessing [23], a "don't know" option was also provided. All questionnaires were checked for quality control and responses were entered into an *ad hoc* database.

### Scoring algorithms

The type-A items were scored by the conventional method: one point if the respondent marked only the keyed correct option and zero otherwise. To score the MM items, a total of eight SAs were implemented (Tab. I). The first was the dichotomous algorithm, which does not allow partial knowledge to be quantified ("all or nothing"); this SA has been widely used as a comparator versus partial SAs [12, 16, 18, 19]. The partial SAs 2-5 were adapted from the paper by Hsu et al. [12]; SAs 2, 4 and 5 involve some penalty for incorrectly chosen options, while SA3 does not. The formula of SA2 is similar to that of SA3, except for the fact that it penalizes incorrect answers; SA2 has been judged rather "severe" regardless of the number of marked distracters and unmarked correct answers provided by a respondent [12]. SA4 and its modifications are among the first methods of partial scoring described in the literature [24, 25]; SA4 consists of subtracting the

Tab. I. Description of scoring algorithms used in the study.

| Scoring algorithm | Definition | Reference |
|---|---|---|
| SA1 | S = 1 if IC = 0, otherwise S = 0 | 12, 16, 18, 19 |
| SA2 | S = (CC – IC)/TO | 12 |
| SA3 | S = CC/TO | 12 |
| SA4 | S = MCO/CO – (MIO/(TO – CO)) | 12 |
| SA5 | S = CC/TO – ((TO!/IC!·(TO – IC)!)/2^TO) | 12 |
| SA6 | p = MCO/CO, if p > 0 ⇒ x = MO/TO – CO/TO, otherwise p = S;<br>if x > 0 ⇒ S = p – (x/(1 – CO/TO)), otherwise p = S | 18 |
| SA7 | S = MCO/CO if MCO ≤ CO, otherwise S = 0 | 16 |
| SA8 | S = 1 if IC = 0, S = 0.5 if 0.5·CO ≤ MCO < CO, otherwise S = 0 | 19 |

S: Respondent's score on a multiple-mark item (max = 1); CO: Number of keyed correct options; CC: Correct choices made by a respondent (both marked correct answers and unmarked distracters); IC: Incorrect choices made by a respondent (both marked distracters and unmarked correct answers); TO: Total number of item options; MCO: Correct options marked by a respondent; MIO: Incorrect options marked by a respondent; MO: Options marked by a respondent; p: Points for MCO; x: Penalty.

proportion of marked distracters from that of marked correct answers. SA5 involves a binomial coefficient and assumes that the incorrect choices made by a respondent are the result of guessing [12]. SAs 2, 3 and 5 treat MM items as MTF ones. SA6, known as balanced SA, has recently been described by Tarasowa and Auer [18]; it includes some logical operators and a penalty is applied only when the number of marked options exceeds that of keyed correct options. The SA7 proposed by Ripkey [16] yields a proportion-of-possible-points score only if the number of marked options does not exceed the number of keyed correct options. SA8, dubbed $PS_{50}$ by Bauer et al. [19], awards the full score if all correct options are marked (no distracters must be marked), half the score if least 50% of correct options are marked, and zero points otherwise.

Items to which no response was given or the "don't know" option was selected were awarded zero points. The "don't know" option was not included in the count of the total number of options used for scoring and data analysis.

Scores of individual items were summed to produce a total score. By agreement between the research team and teachers, for scoring purposes all 12 items were assumed to have the same level of difficulty of 1; the highest possible score was therefore 12.

### Independent variables

Demographic variables of age, sex and immigrant background were recorded from each participant. Body mass index (BMI) was calculated from self-reported height and weight, mapped to the BMI-for-age growth charts and classified in underweight (< $5^{th}$ percentile), normal weight ($5^{th}$-$85^{th}$ percentile), overweight ($85^{th}$-$95^{th}$ percentile) and obese (≥ $95^{th}$ percentile) categories.

### Statistical analysis

For purposes of analysis, the factual nutrition-related knowledge part of the survey was divided (by item type format) into two subsets: the MM subset and the whole survey, which also included 3 type-A items.

Students' scores calculated according to the different SAs were compared by means of repeated-measures analysis of variance (rANOVA); the Greenhouse-Geisser correction for sphericity was applied by applying a significant Mauchly's test statistic. Post-hoc *t* tests for paired data, with *p*-values corrected by means of Bonferroni's method, were subsequently performed. Tarasowa and Auer [18] have suggested that the dichotomous SA1 should be used as a reference rule for scoring MM items (as it virtually excludes the probability of guessing) and that respondents' rankings should then be compared among different SAs; we therefore calculated Spearman's ρ coefficients with 95% confidence intervals (CIs) in order to compare students' rankings yielded by SA1 and the other seven SAs.

The psychometric properties of each SA were evaluated from the perspective of classical test theory. To measure internal consistency, Cronbach's α coefficients with 95% CIs were computed. The eight dependent α coefficients and subsequent pairwise comparisons with adjusted *p*-values were compared by means of Feldt's formulas [26, 27] implemented in the cocron R package [28]. The standard errors (SEs) of students' scores were determined as $SD\sqrt{1-\alpha}$, where SD is the standard deviation of the scores [29]. The efficiency of an SA was evaluated by means of the coefficient of effective length; two SAs with a coefficient of effective length of 1 were considered equally efficient (relative efficiency) [12]. The Spearman-Brown prophecy formula was applied in order to estimate the number of items needed to reach a desirable α of 0.7 and to compare the reliability coefficients of type-A and MM items, considering their different numbers. The item-difficulty index *p*, calculated as the mean score of an item, was categorized as "difficult" ($p < 0.2$), "acceptable" ($0.2 < p < 0.8$) and "easy" ($p \geq 0.8$) [29, 30]. The mean difficulty indexes of the eight SAs were compared by means of rANOVA. The item-discrimination index *D* was computed for each SA; items with $D > 0.2$ were considered acceptable [31].

The differences in the total scores according to the independent variables of interest (gender, immigration background and BMI categories) were quantified by applying standardized mean differences (SMDs) with 95% CIs; SMD was interpreted as large (0.8), medium (0.5) and small (0.2) [32]. Any association between the total score and the independent variables was checked by means of analysis of variance (ANOVA), while that

between the score and the participants' age was checked by means of Pearson's correlation coefficient *r*. These tests were performed separately for each SA.
Assuming an SMD of 0.5 between pre- and post-lecture scores when two-sided α is 0.05 and β is 0.9, we calculated that at least 44 subjects were needed. Cochran's *Q* test was performed to evaluate whether the different SAs had identical effects on the pre- to post-lecture change in individual scores (improved vs. not improved). The pre/post score changes were quantified by means of SMDs. The statistical significance level was conventionally set to two-sided $p < 0.05$. All data were analyzed by means of the R stats package, version 3.1.2 [33] and GPower, version 3.1.9.2 [34].

## Results

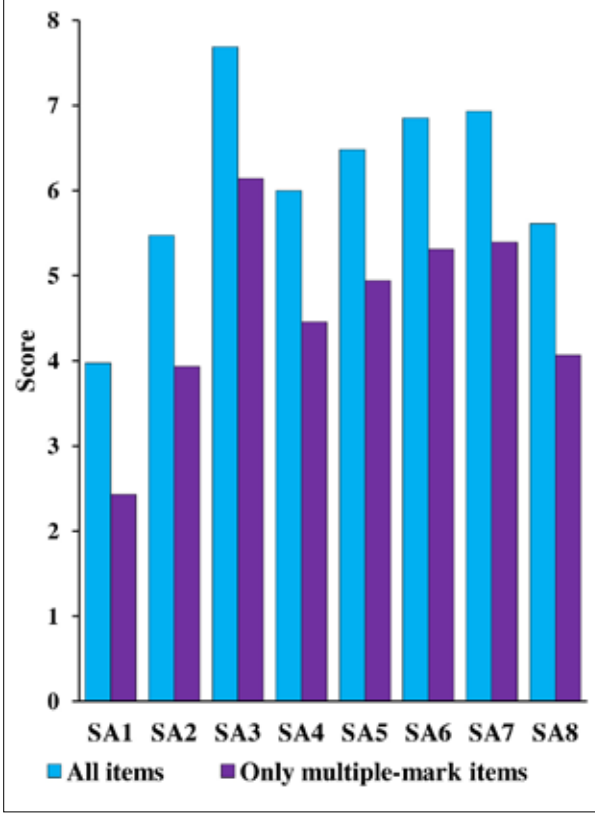
### Sample characteristics

Students took an active part in the survey, completed questionnaires (total 298) being received from all participants. However, 12 questionnaires did not pass the quality check: 9 students had not attempted to answer any question, including demographic ones, while 3 questionnaires contained unlikely answers (such as improbable weight or height). These 12 were discarded and a total of 286 questionnaires were analyzed. Male and female students participated in approximately equal proportions (males: 51.0%) and their mean age was 16.1 (SD 1.1, range 14-20) years. Approximately a quarter of subjects [22.7% (95% CI: 18.0-28.0%)] were from an immigrant background. As calculated from self-reported height and weight, more than four fifths [82.2% (95% CI: 77.2-86.4%)] of students were of normal weight for their age and sex, while 2.1% (95% CI: 0.8-4.5%), 12.2% (95% CI: 8.7-16.6%) and 3.5% (95% CI: 1.7-6.3%) were classified as underweight, overweight and obese, respectively.

### Difference in students' performance, by algorithm

As shown in Figure 1, the summary scores of the seven partial SAs were higher than those yielded by the dichotomous algorithm (Δ means: 1.50, 3.71, 2.03, 2.51, 2.88, 2.96 and 1.64 for SA2-8, respectively); as expected, the partial, non-penalizing SA3 yielded the highest scores. The mean scores of SA6 and SA7 were very close to each other; the mean scores yielded by SA7 were 1.1% and 1.5% higher than those of SA6 in the whole survey and the MM subset, respectively. rANOVA with corrected for sphericity (ε = 0.41) degrees of freedom showed a significant ($p < 0.001$) within-subject effect of SA on students' performances. All pairwise comparisons proved statistically significant. As shown by rank correlation coefficients (Tab. II), students' scores calculated according to SA8 were the most highly correlated with those of SA1; the lower limit of 95% CIs of ρ between SA1 and SA8 did not overlap with the upper limits of most of the other pairwise coefficients.

**Fig. 1.** Students' scores, by scoring algorithm and survey subset.




### Psychometric properties of the scoring algorithms

There was a perceptible difference in the reliability measures of the SA: the dichotomous SA displayed a lower α coefficient (0.48) than any of the partial algorithms (Tab. III). Notably, considering only MM items, SA3, SA6 and SA7 increased their reliability coefficients, but only SA7 reached an $\alpha > 0.7$. The SE of measurement was lowest for SA3 (0.68), while SA1, SA2 and SA4 showed substantially higher SEs (1.08, 1.03 and 1.06, respectively). Analogously, the Spearman-Brown prophecy formula showed that, in order to achieve an α of 0.7, the number of items would need to be more than doubled for the dichotomous SA1, while for the balanced SA6 only three items would need to be added (Tab. III). The reliability coefficient of the three type-A

**Tab. II.** Spearman's ρ correlation coefficients between the dichotomous scoring algorithm 1 (SA1) and the other seven partial scoring rules applied to the multiple-mark survey subset (all $p < .001$).

| Scoring algorithm | ρ | 95% CI |
|---|---|---|
| SA2 | 0.79 | 0.74-0.83 |
| SA3 | 0.73 | 0.67-0.78 |
| SA4 | 0.74 | 0.68-0.79 |
| SA5 | 0.82 | 0.78-0.85 |
| SA6 | 0.81 | 0.77-0.85 |
| SA7 | 0.78 | 0.73-0.82 |
| SA8 | 0.88 | 0.85-0.90 |

**Tab. III.** Reliability measures of the scoring algorithms (SAs), by survey subset.

| Scoring algorithm | α (95% CI) | | N of items needed to reach α = 0.7 | |
|---|---|---|---|---|
| | All (N = 12) | MM (N = 9) | All (N = 12) | MM (N = 9) |
| SA1 | 0.48 (0.38-0.56) | 0.42 (0.32-0.52) | 31 | 29 |
| SA2 | 0.60 (0.53-0.67) | 0.57 (0.49-0.64) | 19 | 16 |
| SA3 | 0.65 (0.59-0.71) | 0.66 (0.60-0.72) | 15 | 11 |
| SA4 | 0.59 (0.51-0.65) | 0.53 (0.45-0.61) | 20 | 19 |
| SA5 | 0.65 (0.58-0.71) | 0.65 (0.59-0.71) | 16 | 12 |
| SA6 | 0.66 (0.60-0.72) | 0.68 (0.63-0.74) | 15 | 10 |
| SA7 | 0.67 (0.62-0.73) | 0.71 (0.65-0.76) | 14 | – |
| SA8 | 0.60 (0.53-0.68) | 0.59 (0.52-0.66) | 19 | 15 |

items was 0.32 (95% CI: 0.17-0.44). The projected coefficient for N = 9 type-A items was estimated to be 0.58, which was lower than the α coefficients of the 9 MM items scored according to SAs 3, 5-8 (Tab. III).

As demonstrated by the coefficients of effective length (Tab. IV), SA1 was the least efficient algorithm, while SA7 was the most efficient. More generally, SAs 3, 5-7 were at least twice as efficient as SA1. The eight reliability coefficients of both survey subsets differed significantly ($p < 0.001$). Several pairwise comparisons of α coefficients proved statistically significant in both survey subsets (Tab. V). In the MM survey subset, the α of SA7 was significantly higher than those of the other seven SAs, while, considering all items, the α of SA7 did not differ significantly from those of SA3 and SA6.

The mean difficulty index (Tab. VI) was the lowest when SA1 was applied, while the quiz was the "easiest" when SA3 was used. The differences among mean difficulty coefficients adjusted for sphericity violations proved to be highly significant in both subsets ($p < 0.001$). All type-A items had difficulty indexes $p$ between 0.2 and 0.8; thus the numbers of easy and difficult items in both survey subsets matched. The highest number (N = 4) of difficult items ($p < 0.2$) was observed when SA1 was used, while according to SAs 3, 5-7, no difficult questions were present in the survey. Conversely, according to SA3, three items were classified as easy ($p > 0.8$), while none were when the dichotomous SA1 was applied.

The item discrimination analysis reported in Table VII did not reveal any negative total item correlation coefficient, while the number of items with $D > 0.2$ varied by SA. SA2 and SA4 showed slightly higher mean discrimination indexes; notably, all MM items scored by SA4 had desirable point-biserial coefficients.

**Tab. IV.** Relative efficiency of the scoring algorithms, as measured by the coefficient of effective length of all items (upper right triangle) and only multiple-mark items (lower left triangle).

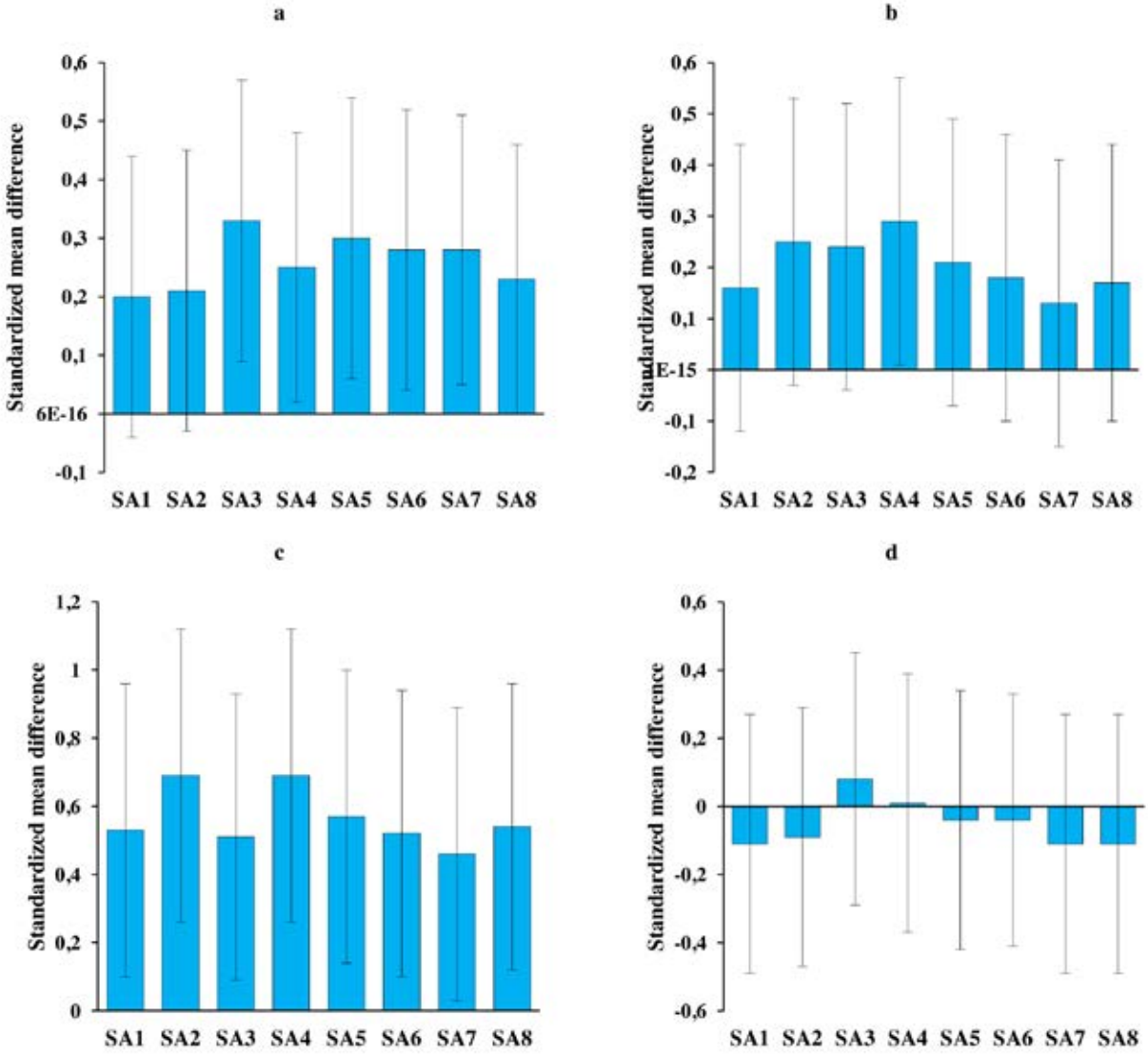
| Scoring algorithm | SA1 | SA2 | SA3 | SA4 | SA5 | SA6 | SA7 | SA8 |
|---|---|---|---|---|---|---|---|---|
| SA1 | – | 1.63 | 2.01 | 1.56 | 2.01 | 2.10 | 2.20 | 1.63 |
| SA2 | 1.83 | – | 1.24 | 0.96 | 1.24 | 1.29 | 1.35 | 1.00 |
| SA3 | 2.68 | 1.46 | – | 0.77 | 1.00 | 1.05 | 1.09 | 0.81 |
| SA4 | 1.56 | 0.85 | 0.58 | – | 1.29 | 1.35 | 1.41 | 1.04 |
| SA5 | 2.56 | 1.40 | 0.96 | 1.65 | – | 1.05 | 1.09 | 0.81 |
| SA6 | 2.93 | 1.60 | 1.09 | 1.88 | 1.14 | – | 1.05 | 0.77 |
| SA7 | 3.38 | 1.85 | 1.26 | 2.17 | 1.32 | 1.15 | – | 0.74 |
| SA8 | 1.99 | 1.09 | 0.74 | 1.28 | 0.77 | 0.68 | 0.59 | – |

**Tab. V.** Pairwise comparisons* of Cronbach's α coefficients of all items (upper right triangle) and only multiple-mark items (lower left triangle)..

| Scoring algorithm | SA1 | SA2 | SA3 | SA4 | SA5 | SA6 | SA7 | SA8 |
|---|---|---|---|---|---|---|---|---|
| SA1 | – | < 0.001 | < 0.001 | 0.011 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| SA2 | 0.003 | – | 0.036 | 0.99 | < 0.001 | < 0.001 | < 0.001 | 0.99 |
| SA3 | < 0.001 | < 0.001 | – | < 0.001 | 0.99 | 0.99 | 0.94 | 0.046 |
| SA4 | 0.31 | 0.23 | < 0.001 | – | < 0.001 | < 0.001 | < 0.001 | 0.99 |
| SA5 | < 0.001 | < 0.001 | 0.99 | < 0.001 | – | 0.99 | 0.048 | 0.001 |
| SA6 | < 0.001 | < 0.001 | 0.99 | < 0.001 | 0.12 | – | 0.51 | < 0.001 |
| SA7 | < 0.001 | < 0.001 | 0.037 | < 0.001 | < 0.001 | 0.004 | – | < 0.001 |
| SA8 | < 0.001 | 0.99 | 0.018 | 0.99 | 0.001 | < 0.001 | < 0.001 | – |

*: Results are reported as $p$-values corrected by means of Bonferroni's method.

Fig. 2. Standardized mean differences in total scores between females and males (a), native and immigrant students (b), native male and immigrant male students (c) and native female and immigrant female students (d), by scoring algorithm.




### In-field assessment: impact of the scoring algorithm on outcome assessment

Neither BMI category nor age was associated with the total score yielded by any SA. Female students scored significantly higher than males on 5 of the 7 partial SAs. The effect size was, however, judged small. By contrast, SA1, SA2 and SA8 were unable to highlight the effect of gender on the respondents' nutrition knowledge (Fig. 2a). Foreign students tended to score lower than Italians, though the difference reached the significance level (low effect size of 0.29) only when SA4 was applied (Fig. 2b). However, the total score yielded by most algorithms was probably determined by a combined effect of gender and immigration background; foreign-born males scored much lower than native male students (Fig. 2c), while no obvious pattern emerged regarding differences in scores between immigrant and Italian females (Fig. 2d). The effect size in scores between foreign and native male students was medium for all but one (SA7) rule. Final ANOVA models of the main effects of gender and nationality and their interaction confirmed the results of univariable statistics, although patterns of main and interaction effects differed by SA (Tab.VIII).

A total of 42 students completed the post-lecture survey. Most students improved their pre-lecture scores, though the proportions differed significantly ($p = 0.006$) by SA (54.8%, 76.2%, 73.8%, 66.7%, 71.4%, 61.9%, 66.7% and 64.3% on using SA1–SA8, respectively). Figure 3 reports effect sizes for pre- and post-lecture scores. The highest effect sizes were observed for SA4 (0.60) and SA2 (0.53), and were judged medium, while the other SAs displayed low effect sizes.

## Discussion

The present study investigated the application of eight different scoring rules for MM items and demonstrated

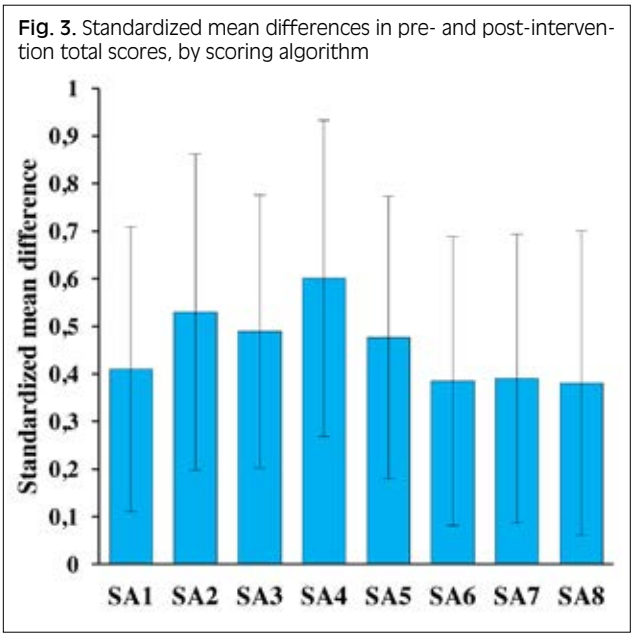



**Fig. 3.** Standardized mean differences in pre- and post-intervention total scores, by scoring algorithm

how they may affect the evaluation of health education outcomes. In line with previous findings [16, 18, 19], we found greater internal consistency and relative efficiency of Ripkey's rule (SA7) and its modifications, such as the balanced algorithm (SA6), in scoring MM items. The SA proposed by Ripkey also performed comparatively well with regard to item difficulty. We found that the choice of SA may have a great influence on student performance; application of the non-penalizing SA3 approximately doubled the total score yielded by the dichotomous SA1. This finding is consistent with previous research [12, 16, 18, 19], which has indicated that MM items scored dichotomously are relatively difficult. In addition, our results support those of Tarasowa and Auer [18], in that SA6 and SA7 penalized respondents more than SA3, which is an MTF-like algorithm, but less than the other rules. This may have important implications for the comparison of knowledge scores obtained from different studies. For instance, previous European studies on nutrition knowledge among adolescents [35-38] have found about 60% of correct responses in multiple choice tests, which roughly corresponds to our estimate (mean percent scores of 54%-64%) obtained by applying partial SAs 3,5-7. By contrast, the dichotomous algorithm produced a substantially lower score of 33%.

Despite the somewhat superior psychometric properties of the Ripkey and the balanced scoring rules, our analysis revealed that SA4 was the only one that identified the negative impact of an immigrant background on the total score. This observation was probably due to a slightly higher discriminatory ability of SA4. The relationship between immigrant status and knowledge scores seems to be plausible; indeed, a large European study [35] conducted in nine countries found a 10% difference in nutrition knowledge scores between native and immigrant adolescents. This coincides with our estimate of a 10.6% mean score difference between Italian and migrant teenagers. On the other hand, the association between immigrant background and knowledge scores probably depends on sex, as shown by the fact that foreign-born male students displayed the poorest performance, regardless of the scoring rule used. Similarly, in the quasi-experimental part of the study, all algorithms were able

**Tab. VI.** Difficulty parameters of survey items, as measured by the different scoring algorithms, by survey subset.

| Scoring algorithm | Mean difficulty, *p* (SD) | | N of easy items | N of difficult items |
|---|---|---|---|---|
| | All (N = 12) | MM (N = 9) | | |
| SA1 | 0.33 (0.27) | 0.27 (0.25) | 0 | 4 |
| SA2 | 0.46 (0.29) | 0.44 (0.32) | 1 | 2 |
| SA3 | 0.64 (0.20) | 0.68 (0.17) | 3 | 0 |
| SA4 | 0.50 (0.26) | 0.50 (0.28) | 1 | 2 |
| SA5 | 0.54 (0.22) | 0.55 (0.22) | 1 | 0 |
| SA6 | 0.57 (0.20) | 0.59 (0.20) | 1 | 0 |
| SA7 | 0.58 (0.21) | 0.60 (0.20) | 1 | 0 |
| SA8 | 0.47 (0.25) | 0.45 (0.26) | 1 | 2 |

**Tab. VII.** Discrimination parameters of survey items, as measured by the different scoring algorithms, by survey subset.

| Scoring algorithm | Mean discrimination index, *D* (SD) | | N of items with $D > 0.2$ | |
|---|---|---|---|---|
| | All (N = 12) | MM (N = 9) | All (N = 12) | MM (N = 9) |
| SA1 | 0.33 (0.21) | 0.35 (0.28) | 8 | 5 |
| SA2 | 0.36 (0.12) | 0.38 (0.16) | 11 | 8 |
| SA3 | 0.30 (0.16) | 0.30 (0.10) | 8 | 6 |
| SA4 | 0.36 (0.12) | 0.37 (0.15) | 11 | 9 |
| SA5 | 0.32 (0.15) | 0.30 (0.14) | 9 | 7 |
| SA6 | 0.32 (0.14) | 0.30 (0.11) | 9 | 8 |
| SA7 | 0.32 (0.13) | 0.31 (0.12) | 9 | 8 |
| SA8 | 0.32 (0.16) | 0.32 (0.16) | 10 | 6 |

Tab. VIII. Results of ANOVA models for sex and nationality, by scoring algorithm.

| Scoring algorithm | Sex | | Nation | | Sex : Nation | |
|---|---|---|---|---|---|---|
| | F | p | F | p | F | p |
| SA1 | 2.88 | 0.091 | 1.87 | 0.17 | 5.43 | 0.021 |
| SA2 | 3.42 | 0.066 | 3.90 | 0.049 | 8.05 | 0.005 |
| SA3 | 8.07 | 0.005 | 4.13 | 0.043 | 3.07 | 0.081 |
| SA4 | 4.66 | 0.032 | 5.53 | 0.019 | 6.47 | 0.012 |
| SA5 | 6.64 | 0.011 | 3.28 | 0.071 | 5.35 | 0.021 |
| SA6 | 5.79 | 0.017 | 2.53 | 0.11 | 4.41 | 0.037 |
| SA7 | 5.62 | 0.018 | 1.44 | 0.23 | 4.40 | 0.037 |
| SA8 | 3.84 | 0.051 | 1.98 | 0.16 | 5.49 | 0.020 |

to highlight the efficacy of the single-lecture intervention in improving students' scores, although SA4 yielded the highest effect size for the pre-/post-intervention difference. Given the above-mentioned patterns, we believe that the choice of an SA for MM items should take into account not only the psychometric properties of single SAs but also the study aims, study population and research topic. This supports the principal conclusions of Muijtjens et al. [39], who suggested that the choice between less biased number-right (e.g. one point for each correct response) and more reliable formula-based scoring rules should be balanced by considering several education factors. For instance, it is acknowledged that females know more about nutrition than males do [35, 40, 41]; it has also been established that females are less likely to guess in multiple-choice tests [42, 43]. It could therefore be speculated that SAs with a correction for guessing would, to some extent, adjust for gender difference in scores. Furthermore, the choice of scoring rule for MM items may affect the statistical power of the analysis, and thus somehow alter outcome assessment. An appropriate SA should therefore be chosen during the design and planning (e.g. sample size calculation) of surveys on health-related knowledge containing this type of item format.

More generally, our results support the principal findings and conclusions of earlier studies [10, 13, 18, 19] on the feasibility of the MM format, since the MM items scored by most of the partial algorithms displayed at least equal internal consistency of the type-A items. MTF and MM items are not rare in health-related knowledge surveys [44, 45], including food- and nutrition-related ones [46], and these items have usually been scored by means of the conventional number-right method. Nevertheless, the guidelines for assessing nutrition-related knowledge, attitudes and practices issued by the *Food and Agriculture Organization of the United Nations* [47] discourage the use of multiple-choice and true-false formats because of the probability of "lucky guessing", and thus overestimation of knowledge scores. However, a correctly guessed answer may be the result of either a blind guess (i.e. a random response given by a fully uninformed subject) or an educated guess (i.e. a response given by a partially informed subject) [48]. Despite its main disadvantage of giving no credit for partial knowledge, use of the dichotomous rule in scoring MM items almost excludes the measurement error due to blind guesses [18]. SA8 showed the highest rank correlation with the dichotomous reference rule (SA1); this confirms the findings of Bauer et al. [19], which indicated that partially scored MM items with a 50% threshold of correct answers may separate the two types of guessing. Scoring MM items as MTF items did not yield any advantage; SAs 2, 3 and 5 neither displayed better psychometric characteristics nor were superior to the others in the on-field outcome evaluation. Despite some similarities between MM-item and MTF-item structures, Cronbach [10, 49] noted a significant difference in questions marked as true, and dubbed this an "acquiescence bias"; poor respondents tend to perform better on items to which the correct answer is "true" rather than "false". In turn, this bias contributes to the skewness of responses [18]. An added advantage of algorithms, especially the balanced SA6, that do not treat MM items as MTF items, is that they allow both MM and type-A questions to be scored. In other words, MM items scored in accordance with SA6 and similar rules make these items a "subspecies" of the type-A items widely used and recognized in health education/promotion research [18].

Overall, our sample may be considered as representative of the adolescent population of Genoa. Furthermore, the distribution of BMI was very close to the estimates obtained by the *Health Behavior in School-Aged Children* (HBSC) study [50] in the Liguria region (underweight: 2.3%, normal weight: 83.1%, overweight: 13.2%, and obese: 1.5%). A very high participation rate enabled us to minimize the response bias. Alongside its strengths, the present study had some limitations. First of all, we used a survey instrument that had not been fully validated, although it was highly comprehensible (as shown by a Gulpease readability index of 78.4, i.e. easy for subjects with a middle-school education) and sensitive to changes. Secondly, relatively low reliability coefficients of the knowledge part of the questionnaire were observed; this was probably due to the small number of survey items. However, Cronbach's $\alpha$ of $> 0.6$ is still acceptable [51] and the coefficients yielded by some partials SAs were comparable to those of well-established literacy instruments (e.g. the Spanish version of the New Vital Sign has an $\alpha$ of 0.69 [52]).

In conclusion, the past few years have seen a revival of the use of MM items to assess factual knowledge [22],

including health-related knowledge [44-46]. In research on health education and promotion, the choice between number-right and formula-based scoring rules, and between formulas that penalize guessing and those that do not, should balance the psychometric properties of single scoring rules and the outcomes of interest. The dichotomous "all or nothing" algorithm should be applied with caution to MM items, especially in cross-sectional study designs, owing to its poorer reliability, item difficulty and discrimination properties. Considering its high sensitivity to blind guessing, we believe that implementation of the dichotomous scoring rule should be limited to highly standardized survey instruments with excellent content validity. However, since school-based health-promotion interventions often require close collaboration with teachers in preparing knowledge-evaluation surveys, the validity of these questionnaires may be far from optimal. In the present study, the scoring rule proposed by Ripkey [16] and the balanced algorithm described by Tarasowa and Auer [18] showed greater internal consistency and relative efficiency in scoring MM items, while the penalizing SA4 was associated with largest effect sizes in the in-field evaluation.

## References


[1] Domnich A, Panatto D, Signori A, et al. *Uncontrolled web-based administration of surveys on factual health-related knowledge: a randomized study of untimed versus timed quizzing*. J Med Internet Res 2015;13:17:e94.

[2] Baker DW. *The meaning and the measure of health literacy*. J Gen Intern Med 2006;21:878-83.

[3] Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, American Medical Association. *Health literacy: report of the Council on Scientific Affairs*. JAMA 1999;281:552-7.

[4] Nutbeam D. *Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century*. Health Promot Int 2000;15:259-67.

[5] Dewalt DA, Berkman ND, Sheridan S, et al. *Literacy and health outcomes: a systematic review of the literature*. J Gen Intern Med 2004;19:1228-39.

[6] Roberts TS. *The use of multiple choice tests for formative and summative assessment*. In: *Proceedings of the 8th Australasian Conference on Computing Education*. Australian Computer Society 2006;52:175-80.

[7] Newble DI, Baxter A, Elmslie RG. *A comparison of multiple-choice tests and free-response tests in examinations of clinical competence*. Med Educ 1979;13:263-8.

[8] Chang SH, Lin PC, Lin ZC. *Measures of partial knowledge and unexpected responses in multiple-choice tests*. Educ Technol Soc 2007;10:95-109.

[9] Lau PNK, Lau SH, Hong KS, et al. *Guessing, partial knowledge, and misconceptions in multiple-choice tests*. Educ Technol Soc 2011;14:99-110.

[10] Cronbach LJ. *An experimental comparison of the multiple true-false and multiple multiple-choice tests*. J Educ Psychol 1941;32:533.

[11] Albanese MA, Kent TH, Whitney DR. *Cluing in multiple-choice test items with combinations of correct responses*. Acad Med 1979;54:948-50.

[12] Hsu TC, Moss PA, Khampalikit C. *The merits of multiple-answer items as evaluated by using six scoring formulas*. J Exp Educ 1984;52:152-8.

[13] Pomplun M, Omar MH. *Multiple-mark items: An alternative objective item format?* Educ Psychol Meas1997;57:949-62.

[14] Frisbie DA, Druva CA. *Estimating the reliability of multiple true-false tests.* J Educ Meas 1986;23:99-105.

[15] Duncan GT, Milton EO. *Multiple-answer multiple-choice test items: responding and scoring through Bayes and minimax strategies*. Psychometrika 1978;43:43-57.

[16] Ripkey DR, Case SM, Swanson DB. *A 'new' item format for assessing aspects of clinical competence*. Acad Med 1996;71(Suppl. 10):S34-6.

[17] Bandaranayake R, Payne J, White S. *Using multiple response true-false multiple choice questions*. Aust N Z J Surg 1999;69:311-5.

[18] Tarasowa D, Auer S. *Balanced scoring method for multiple-mark questions.* CSEDU 2013 – Proceedings of the 5th International Conference on Computer Supported Education 2013, pp. 411-6.

[19] Bauer D, Holzer M, Kopp V, et al. *Pick-N multiple choice-exams: a comparison of scoring algorithms.* Adv Health Sci Educ Theory Pract 2011;16:211-21.

[20] Berk RA. *A consumer's guide to multiple-choice item formats that measure complex cognitive outcomes*. Available at: http://images.pearsonassessments.com/images/NES_Publications/1996_12Berk_368_1.pdf

[21] Frary RB. *Partial-credit scoring methods for multiple-choice tests.* Appl Meas Educ 1989;2:79-96.

[22] Verbić S. *Information value of multiple response questions*. Psihologija 2013;45:467-85.

[23] Pennington HR, Pachana NA, Coyle SL. *Use of the facts on aging quiz in New Zealand: validation of questions, performance of a student sample, and effects of a don't know option*. Educ Gerontol 2001;27:409-16.

[24] Dressel PL, Schmid J. *Some modifications of the multiple-choice item*. Educ Psychol Meas1953;13:574-95.

[25] Morgan MRJ. *MCQ: An interactive computer program for multiple-choice self-testing*. Biochem Educ 1979;7:67-9.

[26] Feldt LS, Woodruff DJ, Salih FA. *Statistical inference for coefficient alpha*. Appl Psychol Meas1987;11:93-103.

[27] Feldt LS. *A test of the hypothesis that Cronbach's alpha reliability coefficient is the same for two tests administered to the same sample*. Psychometrika 1980;45:99-105.

[28] Package "cocron". Available at: http://cran.r-project.org/web/packages/cocron/cocron.pdf

[29] Kline P. *The handbook of psychological testing. 2nd edition*. Abingdon: Routledge 1993.

[30] Mackison D, Wrieden WL, Anderson AS. *Validity and reliability testing of a short questionnaire developed to assess consumers' use, understanding and perception of food labels*. Eur J Clin Nutr 2010;64:210-7.

[31] Crocker L, Algina J. *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston 1986.

[32] Cohen J. *Statistical power analysis for the behavioral sciences. 2nd edition*. Hillsdale, NJ: Lawrence Erlbaum Associates 1988.

[33] R Core Team R. *A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing 2014. Available at: http://www.R-project.org/.

[34] Faul F, Erdfelder E, Buchner A, et al. *Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses*. Behav Res Methods 2009;41:1149-60.

[35] Sichert-Hellert W, Beghin L, De Henauw S, et al. *Nutritional knowledge in European adolescents: results from the HELENA (Healthy Lifestyle in Europe by Nutrition in Adolescence) study*. Public Health Nutr 2011;14:2083-91.

[36] Reinehr T, Kersting M, Chahda C, et al. *Nutritional knowledge of obese compared to non obese children*. Nutr Res 2003;23:645-9.

[37] Osler M, Hansen ET. *Dietary knowledge and behaviour among*

*schoolchildren in Copenhagen, Denmark*. Scand J Soc Med 1993;21:135-40.

[38] Cunningham-Sabo LD, Davis SM, Koehler KM, et al. *Food preferences, practices, and cancer-related food and nutrition knowledge of southwestern American Indian youth*. Cancer 1996;78(Suppl. 7):1617-22.

[39] Muijtjens AM, Mameren HV, Hoogenboom RJ, et al. *The effect of a 'don't know' option on test scores: number-right and formula scoring compare*d. Med Educ 1999;33:267-75.

[40] Shepherd R, Towler G. *Nutrition knowledge, attitudes and fat intake: application of the theory of reasoned action*. J Hum Nutr Diet 2007;20:159-69.

[41] Shah P, Misra A, Gupta N, et al. *Improvement in nutrition-related knowledge and behaviour of urban Asian Indian school children: findings from the 'Medical education for children/ Adolescents for Realistic prevention of obesity and diabetes and for healthy aGeing' (MARG) intervention study*. Br J Nutr 2010;104:427-36.

[42] Slakter MJ, Koehler RA, Hampton SH, et al. *Sex, grade level, and risk taking on objective examinations*. J Exp Educ1971;39:65-8.

[43] Ben-Shakhar G, Sinai Y. *Gender differences in multiple-choice tests: the role of differential guessing tendencies*. J Educ Meas 1991;28:23-35.

[44] Suda AL, Jennings F, Bueno VC, et al. *Development and validation of Fibromyalgia Knowledge Questionnaire: FKQ*. Rheumatol Int 2012;32:655-62.

[45] Maciel SC, Jennings F, Jones A, et al. *The development and validation of a Low Back Pain Knowledge Questionnaire – LKQ*. Clinics (Sao Paulo) 2009;64:1167-75.

[46] Byrd-Bredbenner C, Wheatley V, Schaffner D, et al. *Development and implementation of a food safety knowledge instrument*. J Food Sci Education 2007;6:46-55.

[47] Macías YF, Glasauer P. *Guidelines for assessing nutrition-related knowledge, attitudes and practices.* Rome: Food and Agriculture Organization of the United Nations 2014.

[48] Mondak JJ, Davis BC. *Asked and answered: Knowledge levels when we will not take "don't know" for an answer*. Polit Behav 2001;23:199-224.

[49] Cronbach LJ. *Studies of acquiescence as a factor in the true-false test*. J Educ Psychol 1942;33:401-15.

[50] Health Behavior in the School-Aged Children study: rapporto sui dati 2010. Rapporto ISTISAN 13/5, 2013. Available at: http://www.hbsc.unito.it/it/images/pdf/hbsc/report_nazionale_2010.pdf

[51] Nunnally J. *Psychometric theory. Second edition*. New York: McGraw-Hill 1978.

[52] Weiss BD, Mays MZ, Martz W, et al. *Quick assessment of literacy in primary care: the newest vital sign*. Ann Fam Med 2005;3:514-22.

■ Received on June 10, 2015. Accepted on October 30, 2015.

■ Correspondence: Alexander Domnich, Department of Health Sciences, University of Genoa, via Pastore 1, 16132 Genoa, Italy - Tel. +39 010 3538524 - Fax +39 010 3538541 - E-mail: alexander.domnich@gmail.com

ORIGINAL ARTICLE

# The relationship between physical activity and quality of life in prisoners: a pilot study

A. MANNOCCI[1], D. MASALA[2], D. MIPATRINI[1], J. RIZZO[1], S. MEGGIOLARO[1], D. DI THIENE[1], G. LA TORRE[1]
[1] Department of Public Health and Infectious Diseases, "Sapienza" University of Rome, Italy; [2] Department of Human, Social and Health Sciences, University of Cassino and Southern Lazio, Italy

Key words

Prison • Physical activity • Quality of life

Summary

**Introduction**. *Imprisoned people have usually a poor health status in comparison with the general population. The aim is to investigate a possible association between the quality of life and physical activity level in male inmates.*
**Methods**. *A cross-sectional pilot study was carried out between 2010 and 2011. A questionnaire contained SF12 and International Physical Activity Questionnaire was administered. Mental Component Score (MCS) and Physical Component Score (PCS) were computed. The physical activity level was measured using Metabolic Equivalents score (MET).*
**Results**. *121 inmates entered the survey. MCS is directly correlated to MET of physical activity ($\beta = 0.23$; $P = 0.03$) and negatively to BMI variations ($\beta = -0.24$; $P = 0.02$) and smoking status ($\beta = -0.24$; $P = 0.02$).*
**Discussion**. *This pilot study suggests to improve the investigation to support the promotion of physical activity programs in Italian prisons in order to improve inmates QoL and allow a better social integration at the end of detention.*

## Introduction

Imprisoned people have usually a poor health status in comparison with the general population [1].
This health disparity has been attributed to various causes. The behavioral factors such as use of intravenous drug, alcohol misuse, smoking, and inactivity could increase the risk of morbidity, mortality and mental disorders [2-5].
There is clear evidence of a connection between physical exercise and *Health Related Quality of Life* (HRQoL) [6-8], and this association has been proven for detention environments too [9-10]. Studies show a direct relationship between self-reported exercise, calculated in minutes per week, and good health of prisoners [4, 11].
In Italy no studies evaluated the association between quality of life and physical activity in the prison population while two studies suggest an association between physical activity and improvement of mental and physical health among prisoners [3, 12].
The hypothesis of the investigation is the possible association between the HQoL, assessed through the short form 12 questionnaire (SF12), and *Metabolic Equivalents* (MET) of physical activity in male inmates, assessed through the IPAQ questionnaire.

## Methods

### Study design, setting, participants

This cross-sectional pilot study was carried out between 2010 and 2011.
STROBE Statement was applied to plan the study [13].
The male inmates were recruited on a voluntary basis in two prisons situated in the Center Italy, Rebibbia in Rome and the Cassino penitentiary. The prisons were chosen on an opportunistic basis.
We enrolled only the prisoners apparently healthy who showed good conduct during their detention and we excluded the inmates in higher security and mental health units.
The questionnaires were administered twice for prison, during each day of administration around 30 prisoners answered the questionnaire: 10 minutes to explain the project and 50 to support and complete the questionnaire were spent and reduce the missing data. The prisoners enrolled were informed concerning the nature and the aim of the research project.

### Questionnaire and variables

The questionnaire included three parts:
the *International Physical Activity Questionnaire* (IPAQ) [14] investigating the time spent carrying out moderate/intensity activity, walking and sitting in a week and referring adapted examples for inmates [15]. IPAQ responses were converted to MET minutes per week

(MET-min/wk) according to the IPAQ scoring protocol [16] through the analysis of total minutes over last seven days spent on vigorous activity, moderate-intensity activity, and walking;
Short Form-12 (SF-12) questionnaire [17] presenting 12 questions evaluating the HRQoL. The answers provide two outcome measures of the physical and the mental QoL: the *Physical Componet score* (PCS) and *Mental Componet Score* (MCS) and the score ranges from 0 to 100 (bad and good quality respectively). Although the use of SF12 rather than SF36 has been suggested for big samples (more than 500 subjects) the SF-12 was chosen for two reasons. First of all we were interested in the two summary subscales rather than eight health dimensions provided by the SF36 and secondly SF-12 has been already used in community-based cross-sectional studies [18].
Socio-demographic and biometric variables were included because they were considered potential confounding: age, gender, nationality, educational level, period of detention (dichotomized in 0-2 and ≥ 2; in years), civil status, smoking habits, weight in kilograms and height in centimeters auto-referred, subsequently combined to compute the BMI. Weight at the entrance was asked to and differential BMI was calculated.
The questionnaire was translated in three languages (Italian, English and France) allowing foreign prisoners to participate to the study.
A written consent to participate into the survey was collected from each participants. The study was approved by the Local Ethical Committee.

### Statistical analysis

All results were recorded on paper and then double entered into SPSS19.0 database for analysis.
A probability level of 0.05 was set as limit for statistical significance.
Statistical analysis included descriptive statistics (mean, standard deviation and proportion) and inferential statistics.
In order to assess association between PCS and MCS of HRQoL and socio-demographic, biometric and attitudinal variables different statistic tests were performed:
to assess the normal distribution of the continuous variables, including the outcomes, the Kolmogorof-Smirnoff 's test was applied;
Pearson's Correlation for quantitative variables (age, BMI at the imprisonment, Current BMI, BMI variation, MET of physical activity);
Mann Whitney's test for dichotomous qualitative variables (civil status, smoking status, years of detention, nationality);
Kruskal-Wallis test has been performed for qualitative variables with more than two modalities.
Multivariate analysis was performed. According with Hosmer and Lemeshow [19] variables associated to PCS and MCS with probability level of 0,25 or less at the univariate and bivariate analysis were included in the model. Given the exploratory nature of the research project and the little literature available on the issue we chose to adopt backward elimination model. Only covariates still present at the last elimination step were reported in Table III.

## Results

121 inmates entered the survey, 64 in Rome and 57 in Cassino detection centers (response rate = 87%).
The illiterate or the low educational level and no knowledge of English, Italian or French were the most common missing data sources that were referred from researchers that have administered the questionnaires.
The socio-demographic, biometric and attitudinal characteristics of the sample are described in Table I.
The bivariate analysis shows a direct association between PCS of HRQoL and MET of physical activity ($r = 0.22$; $P = 0.04$) and an inverse association with BMI variation ($r = -0.23$; $P = 0.02$). MCS of HRQoL is directly associated with MET of physical activity ($r = 0.25$; $P = 0.02$) and with age of inmates ($r = 0.21$; $P = 0.03$) and an inversely associated with the smoking attitude ($P = 0.02$). Educational level, civil status, nationality and years of detention are not significantly associated with the inmates' quality of life.
Two multivariate analyses were performed, one with each HRQoL score (PCS and MCS) as outcome. The following variables entered the analysis of PCS as covariates: current BMI, BMI variation since the start of detention, MET of physical activity, Educational level, smoking status and years of detention. PCS resulted di-

**Tab. I.** Descriptive characteristics of the inmates studied.

| Continuous variables | | N. of responders | Mean | SD |
|---|---|---|---|---|
| Age (years) | | 121 | 35.0 | 11.1 |
| BMI at the start of the detention | | 114 | 25.9 | 4.7 |
| BMI at the present | | 112 | 26.1 | 3.4 |
| PCS | | 113 | 48.7 | 7.7 |
| MCS | | 113 | 45.6 | 11.0 |
| Total MET per week | | 95 | 9278.9 | 13845.9 |
| Years of detention | | 120 | 5.20 | 6.1 |
| | | | | |
| **Qualitative variables** | | | % | |
| Educational level | None or primary school | 29 | 24.6 | |
| | Secondary school | 68 | 57.6 | |
| | High school / university | 21 | 17.8 | |
| Smoking status | Smokers | 83 | 69.7 | |
| | No smokers | 36 | 30.3 | |
| Civil status | Married | 41 | 33.9 | |
| | Not married | 80 | 66.1 | |
| Years of detention | 0-2 | 54 | 45.0 | |
| | > 2 | 66 | 55.0 | |
| Nationality | Italian | 77 | 66.4 | |
| | Foreign | 39 | 33.6 | |

**Tab. II.** The impact of demographic and anthropometric characteristics and behaviors on the Quality of Life, bivariate analysis.

| | Physical score | | Mental score | |
|---|---|---|---|---|
| | Pearson's correlation | P | Pearson's correlation | P |
| **Age** | 0.023 | 0.807 | 0.209 | **0.026** |
| BMI at the imprisonment | 0.077 | 0.428 | 0.061 | 0.533 |
| Current BMI | -0.135 | **0.172** | -0.108 | 0.274 |
| BMI variation | -0.232 | **0.021** | -0.164 | **0.106** |
| MET total | 0.217 | **0.040** | 0.248 | **0.018** |
| Years of detention | 0.088 | 0.355 | -0.041 | 0.666 |
| | **Median (IQR)** | **P*** | **Median (IQR)** | **P*** |
| **Educational level** | | | | |
| None/primary school | 50.2 (13.3) | | 47.0 (15.7) | |
| Secondary school | 48.7 (11.7) | **0.232** | 49.6 (15.7) | 0.591 |
| High school / university | 51.7 (9.8) | | 43.3 (21.0) | |
| **Civil status** | | | | |
| Not married | 50.2 (12.7) | 0.988 | 48.3 (15.0) | 0.955 |
| Married | 49.5 (11.43) | | 46.4 (19.2) | |
| **Smoking status** | | | | |
| No | 52.0 (8.7) | **0.226** | 52.1 (12.3) | **0.022** |
| Yes | 49.5 (13.2) | | 45.0 (18.2) | |
| **Years of detention** | | | | |
| 0-2 | 48.6 (11.4) | **0.194** | 50.7 (19.2) | 0.898 |
| > 2 | 50.8 (12.5) | | 47.1 (15.2) | |
| **Nationality** | | | | |
| Italian | 49.4 (13.0) | 0.400 | 47.6 (18.3) | 0.614 |
| Foreign | 50.4 (10.0) | | 47.5 (17.6) | |

IQR: interquartile range

* P: p-value of Mann Whitney's test for dichotomous qualitative variables and Kruskal-Wallis test for qualitative variables with more than two modalities.

Bold: $p < 0.05$.

**Tab. III.** The impact of demographic and anthropometric characteristics and behaviors on the Quality of Life, multivariate analysis.

| Covariates* | Physical component score | | Mental component score | |
|---|---|---|---|---|
| | Beta | P | Beta | P |
| BMI variation | **-0.262** | **0.017** | **-0,242** | **0,023** |
| MET total | **0.250** | **0.023** | **0,232** | **0,030** |
| Current BMI | -0.135 | 0.217 | | |
| Smoke | -0.119 | 0.270 | **-0,242** | **0,023** |
| Educational level | 0.018 | 0.872 | | |
| Age | | | 0,153 | 0,140 |
| $R^2$ | 0.090 | | 0.143 | |

* Only variables still present at the last step of the backward elimination model are shown in the table.

Bold: $p < 0.05$.

rectly associated to physical activity ($r = 0.25$; $P = 0.02$) and inversely correlated to BMI variations ($r = -0.26$; $P = 0.02$). The following variables were included in the model for MCS: age. BMI variation since the start of detention, MET of physical activity and smoking status. MCS is influenced positively by MET of physical activity ($r = 0.23$; $P = 0.03$) and negatively by BMI variations ($r = -0.24$; $P= 0.02$) and smoking status ($r = -0.24$; $P = 0.02$).

## Discussion

Our results are consistent with the scientific literature on the issue. Even if HRQoL is rarely investigated among inmates population, physical activity has been proven to reduce depression and anxiety [3], increase functional capacity and improving the health status of incarcerated people [12]. Moreover in a US high security prison women participating in a focus group described limited physical activity as a specific prison-based factors that affected their physical health [20]. Our study reveals that physical activity is associated to both the physical and the mental components of the HRQoL in the inmates population. According with Meek et al. [21] we can state that physical activity can contribute to achieving "healthy prison" objectives in practice.

Moreover, there is evidence from our survey that BMI variations are inversely associated to the HRQoL and tobacco smoking is inversely associated with the mental component of HRQoL. Demographic factors and the educational level does not appear to affect the HRQoL. Likely this can be partially explained by the low sample size, it would be appropriate to continue to investigate this aspect.

The strengths of this study are mainly two. First the innovation that brings to the Italian landscape where few experiences have been conducted on the topic. Few studies have been carried out in this field. Even if the focus of this work is specific it can contribute to the evidence building on the prisoners health. Secondly, the possibility to actively intervening to bettering the HRQoL of inmates through the realization of physical training in Italian prisons. Physical activity is a proven cost effective tool to challenge chronic diseases in general population and this research seems to support albeit weakly to a number of limitations to the idea that physical activity can be beneficial to the quality of life. Larger studies are needed to find out [22].

However, we have to acknowledge the study presents some limitations. Above all the sample size (120 participants) and composition (only males) could not infer the associations we found to the general inmate population. Moreover, only few behavioral factors have been evaluated and not pathologic or environmental factors have been assessed due to the nature of pilot study.

The goodness of fit of the model could have been increased adding other factors to the multivariate analysis. Anyway, the results are promising and a multicenter study is ongoing to expand the sample and include the

female population in the study. Further research approaches could include the realization of a randomized trial with the organization of physical training in prisons and qualitative analysis to identify other determinants of inmates HRQoL.
If physical activity will be confirmed to highly affect the inmates HRQoL practical implications have to be considered. The realization of training programs for detention centers could be a simple, effective, and cost effective way to improve the Quality of Life of prisoners and to contribute to the re-educational aim of the prison detention. According with the Italian Constitution Document "punishments must aim at rehabilitating the condemned" for the restitution of healthy and renewed persons to the community [23].
Physical activity programs in prisons could contribute to accomplish this target.

## Conclusions

Both physical and mental components of quality of life are affected by physical activity that seems to be the main determinant, joint with the BMI variation and smoking attitude, of inmates' quality of life [9].
A multicenter study is ongoing in Italy to confirm preliminary results.
Programs promoting physical activity should be planned and implemented in Italian prisons in order to improve inmates QoL and allow a better social integration at the end of detention.

### References


[1] Fazel S, Baillargeon J. *The health of prisoners*. Lancet 2011;377:956-65.

[2] Fischer J, Butt C, Dawes H, et al. *Fitness levels and physical activity among class A drug users entering prison.* Br J Sports Med 2012;46:1142-4.

[3] Battaglia C, di Cagno A, Fiorilli G, et al. *Participation in a 9-month selected physical exercise programme enhances psychological well-being in a prison population*. Crim Behav Ment Health 2014;n/a – n/a.

[4] Cashin A, Potter E, Butler T. *The relationship between exercise and hopelessness in prison.* J Psychiatr Ment Health Nurs 2008;15:66-71.

[5] Brinded PM, Simpson AI, Laidlaw TM, et al. *Prevalence of psychiatric disorders in New Zealand prisons: a national study.* Aust N Z J Psychiatry 2001;35:166-73.

[6] Zucker DM, Villemaire L, Rigali C, et al. *The evolution of a labyrinth walking program in corrections.* J Forensic Nurs 2013;9:101-4.

[7] Brown DW, Brown DR, Heath GW, et al. *Associations between physical activity dose and health-related quality of life.* Med Sci Sports Exerc 2004;36:890-6.

[8] Anokye NK, Trueman P, Green C, et al. *Physical activity and health related quality of life*. BMC Public Health 2012;12:624.

[9] Zucker DM, Villemaire L, Rigali C, et al. *The evolution of a labyrinth walking program in corrections*. J Forensic Nurs 2013;9:101-4.

[10] U.S. Department of Health and Human Services. *Physical Activity and Health: A Report of the Surgeon General.* Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion [Internet]. 1996: http://www.cdc.gov/nccdphp/sgr/index.htm

[11] Pérez-Moreno F, Cámara-Sánchez M, Tremblay JF, et al. *Benefits of exercise training in Spanish prison inmates.* Int J Sports Med 2007;28:1046-52.

[12] Battaglia C, di Cagno A, Fiorilli G, et al. *Benefits of selected physical exercise programs in detention: a randomized controlled study.* Int J Environ Res Public Health 2013;10:5683-96.

[13] Von Elm E, Altman DG, Egger M, et al. *The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for reporting observational studies.* Int J Surg 2014;12:1495-9.

[14] International Physical Activity Questionnaire [Internet]: https://sites.google.com/site/theipaq/

[15] Mannocci A, Di Thiene D, Del Cimmuto A, et al. *International Physical Activity Questionnaire: validation and assessment in an Italian sample*. Ital J Publ Health 2010;7:369-76.

[16] Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ) – Short and Long Forms Contents [Internet]: http://www.academia.edu/5346814/Guidelines_for_Data_Processing_and_Analysis_of_the_International_Physical_Activity_Questionnaire_IPAQ_Short_and_Long_Forms_Contents

[17] Apolone G, Mosconi P, Ware JJ. *Questionario sullo stato di salute SF-36* [Internet]. Guerini e Associati, 1997.

[18] Domnich A, Amicizia D, Panatto D, et al. *Use of different subjective health indicators to assess health inequalities in an urban immigrant population in north-western Italy: a cross-sectional study*. BMC Public Health 2013;13:1006.

[19] Hosmer DW, Hosmer T, Le Cessie S, et al. *A comparison of goodness-of-fit tests for the logistic regression model*. Stat Med 1997;16:965-80.

[20] Harner HM, Riley S. *Factors contributing to poor physical health in incarcerated women*. J Health Care Poor Underserved 2013;24:788-801.

[21] Meek R, Lewis G. *The role of sport in promoting prisoner health*. Int J Prison Health 2012;8:117-30.

[22] Roux L, Pratt M, Tengs TO, et al. *Cost effectiveness of community-based physical activity interventions.* Am J Prev Med 2008;35:578-88.

[23] Art. 27 of the Italian Constitution Document. 1947.


■ Received on May 29, 2015. Accepted on September 7, 2015.

■ Correspondence: Giuseppe La Torre, Dipartimento di Sanità Pubblica e Malattie infettive, "Sapienza" Università di Roma, Piazzale Aldo Moro 5, 00187 Roma, Italy - Tel. +39 06 49914308 - Fax +39 06 4453845 - E-mail:giuseppe.latorre@uniroma1.it

ORIGINAL ARTICLE

# Susceptibility of multidrug resistant clinical pathogens to a chlorhexidine formulation

F. GÜNTHER, S.J. KAISER, T. FRIES, U. FRANK, N.T. MUTTERS
Heidelberg University Hospital, Heidelberg, Germany

Key words

Chlorhexidine • Drug resistant • Gram-negative • Hospital-acquired infection • MRSA • VRE

Summary

Multidrug resistant pathogens are a widespread problem in the hospital setting especially on intensive care units (ICU). This study evaluated the susceptibility of clinical isolates of gram-negative extensively drug resistant organisms (XDR), methicillin-resistant *Staphylococcus aureus* (MRSA), and vancomycin-resistant *Enterococcus* (VRE) to a proprietary chlorhexidine digluconate (CHG) formulation used in one brand of CHG-impregnated cloths. Ten isolates each of XDR *Pseudomonas aeruginosa*, XDR *Acinetobacter baumannii*, XDR *Klebsiella pneumoniae*, XDR *Escherichia coli,* MRSA, and vancomycin-resistant *Enterococcus faecium* from our hospital were tested. All isolates were susceptible to the proprietary CHG formulation (0.5%, 1%, 2%), with 99% to 100% suppression of growth at the earliest time point in time kill assays (1 minute for gram-positive and 15 seconds for gram-negative organisms). Minimum inhibitory concentrations ranged from 1 : 4096 to 1 : 65536 for MRSA, 1 : 1024 to 1 : 2048 for VRE, 1 : 2048 to 1 : 4096 for XDR *E. coli*, 1 : 512 to 1 : 2048 for XDR *A. baumannii*, 1 : 512 to 1 : 1024 for XDR *P. aeruginosa*, and 1 : 512 to 1 : 1024 for XDR *K. pneumoniae*. Cloths impregnated with this CHG formulation provide effective protection against colonization and infection by many pathogens. This study provides *in vitro* evidence that the proprietary CHG formulation used in one brand of CHG-impregnated cloths is effective against XDR gram-negative organisms, MRSA, and VRE.

## Introduction

The rapid emergence and spread of multidrug resistant organisms (MDROs) in hospitals is a growing problem worldwide [1, 2]. Hospital-acquired infections, particularly those caused by MDROs, are associated with excess mortality and morbidity as well as increased hospital costs [3-7]. Methicillin-resistant *Staphylococcus aureus* (MRSA), vancomycin-resistant *Enterococcus* (VRE), and multidrug resistant (MDR) gram-negative pathogens commonly cause hospital-acquired infections [8]. Managing hospital-acquired infections and MDROs is a daily challenge in hospitals, especially from the perspective of critical care.

Universal decolonization using topical antiseptic agents that reduce the population of microorganisms on patients' skin represents a simple, cost-effective way to prevent healthcare-associated infections [9, 10]. Universal decolonization by daily bathing of ICU patients with chlorhexidine digluconate (CHG)-impregnated cloths resulted in a substantial reduction in bloodstream infections and MRSA acquisition [9, 11-13]. CHG-impregnated cloths have also proven effective in reducing the skin burden of MRSA and VRE in ICU patients [9, 14, 15].

Evidence suggests that skin colonization with gram-negative bacterial pathogens may be a root cause of hospital-acquired infections. ICU patients who have diarrhea can be particularly at risk of gram-negative bacteria dissemination from feces to skin areas on distant parts of the body [16]. There is less information concerning the utility of CHG on antibiotic-resistant gram-negative bacteria [15], although a recent study showed that a proprietary CHG formulation reduced hospital-acquired infections caused by gram-negative bacteria [17]. The objective of this study was to quantify *in vitro* the antimicrobial effectiveness of a proprietary CHG formulation used in one brand of CHG-impregnated cloths against gram-negative MDR and extensively drug resistant (XDR) clinical isolates (e.g., *Pseudomonas aeruginosa*, *Klebsiella pneumoniae, Escherichia coli*, *Acinetobacter baumannii*), as well as against clinical isolates of MRSA and VRE.

## Materials and methods

### Bacterial Isolates

Clinical isolates of *P. aeruginosa*, *K. pneumoniae, E. coli*, and *A. baumannii* selected for testing were classified as MDR or XDR as described elsewhere [18]. Clinical isolates of MRSA and vancomycin-resistant *E. faecium* were also tested. Ten consecutive isolates per species were collected from patients who were diagnosed with an infection cause by a MDRO during their stay in at the Heidelberg University Hospital, Germany, in 2014.

### Formulations and Neutralizer

A proprietary 2% CHG formulation (Sage Products LLC, Cary, Illinois, US) was tested. The solution used

to neutralize the antimicrobial properties of the CHG was composed of Caso-bouillon and LTHTh (Heipha, Eppelheim, Germany).

### Time Kill Assays

The antimicrobial properties of the proprietary CHG formulation at concentrations of 2% (20 mg/ml, the original proprietary formulation concentration), 1% (10 mg/mL), and 0.5% (5 mg/mL) were determined using the quantitative suspension methods described by Gebel et al. [19]. Briefly, bactericidal efficacy was determined without organic load. The CHG formulation was diluted in water of standardized hardness. One milliliter of the test organism suspension and 1 mL of sterile water of standardized hardness were mixed and incubated for 2 minutes, after which the test substance was added. The resulting solutions were incubated for 1, 3, and 9 minutes, respectively; gram-negative organisms were also incubated for 15 seconds, based on the test procedures outlined in *the FDA Tentative Final Monograph for Topical Antimicrobial Drug Products for Over-The-Counter Human Use* (59 FR, 31444, June 17, 1994). At the end of the incubation time, 1 mL of the test solution was transferred to 10 mL of Caso-bouillon and LTHTh and neutralized for 5 minutes. Thereafter, 100 µl and 500 µl of the neutralized test solution was spread onto two agar plates. After incubation for 24 hours at 37 °C the colony forming units (CFU) were counted. The log 10 reduction was determined as the logarithm to the base 10 of the difference between the number of cells in the test solution at the beginning of the contact time and at the end.

### Determination of the Minimum Inhibitory Concentration (MIC)

MIC testing procedures were adapted based on those outlined in the *FDA Tentative Final Monograph for Topical Antimicrobial Drug Products for Over-The-Counter Human Use* (59 FR, 31444, June 17, 1994). Briefly, a 96-well microtiter plate (Sarstedt, Nümbrecht, Germany) containing doubling dilutions of the proprietary CHG in RPMI medium (Gibco, Darmstadt, Germany) was set up. CHG concentrations started at 2% and were then diluted by half. Concentrations for MIC testing were diluted down to 1 : 65,536. Wells containing only RPMI were used as growth controls. An overnight broth culture of each isolate was standardized to 1 x $10^8$ CFU/mL and 50 µL volumes of this were added to the microtiter plate. Serial dilutions were done and CFUs counted on the plates to achieve a concentration of 1 x $10^5$. Plates were incubated for 24 hours at 37 °C. The MIC was defined as the lowest concentration of CHG at which no bacterial growth was observed visually on the microtiter plate. Conversion of resazurin to resarafin (Sigma-Aldrich, Hamburg, Germany) was used as a visual indicator.

## Results

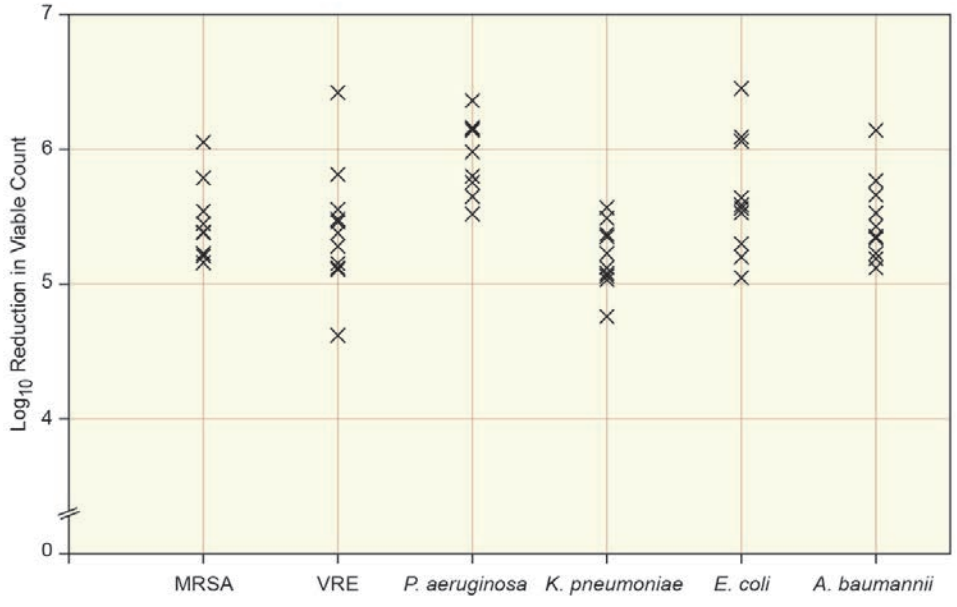
### Isolates

Ten isolates each of VRE, MRSA, XDR *P. aeruginosa*, XDR *K. pneumoniae,* and *E. coli* were collected and tested. Nine isolates of MDR *A. baumannii* and one of XDR *A. baumannii* were collected and tested.

### In Vitro Time Kill Studies

All isolates were highly susceptible to the proprietary CHG formulation (Fig. 1). Suppression rates were 99%

**Fig. 1.** Time kill at 1 minute with a 2% CHG proprietary solution of clinical isolates of MRSA, vancomycin-resistant *E. faecium*, and the *XDR gram-negative pathogens P. aeruginosa, K. pneumoniae, E. coli,* and *A. baumannii.*



to 100% for all isolates, at all concentrations including the lower concentrations of 1% (10 mg/mL) and 0.5% (5 mg/mL). The 0.5% CHG formulation provided a 99.9% reduction of XDR *P. aeruginosa*, XDR *K. pneumoniae*, XDR *E. coli*, and XDR and MDR *A. baumannii* from the earliest 15-second time point, and for MRSA and VRE at the earliest 1 minute time point.

### MIC Determinations

MICs for the proprietary CHG solution against MRSA ranged from 1 : 4096 for a single isolate to 1 : 65536 (Tab. I). Nine out of ten XDR *E. coli* isolates had an MIC of 1 : 4096 with the CHG formulations. XDR *A. baumannii* showed relatively low MICs, with an MIC of 1 : 1024. Vancomycin-resistant *E. faecium* most commonly demonstrated MICs of 1 : 2048. XDR *K. pneumonia* and XDR *P. aeruginosa* had slightly higher MICs. The MICs for *P. aeruginosa* MICs ranged from 1 : 512 to 1 : 1024. MICs for XDR *K. pneumoniae* ranged from 1 : 512 to 1 : 4096.

## Discussion

Hospital-acquired infections are leading causes of preventable morbidity and mortality [2, 20]. Skin decolonization with CHG-impregnated cloths has been shown to reduce the risk of some types of these infections [11, 9, 12, 13, 16, 21]. This method was shown to be an effective and cost saving way to reduce the risk of transmission of MDROs such as MRSA and VRE in the hospital, a setting where rapid emergence and spread of MDROs is well known [9, 12, 15, 10, 1]. The *in vitro* time kill studies confirmed that all of the German clinical isolates tested, including MDR and XDR gram-negative bacteria, MRSA, and vancomycin-resistant *E. faecium*, were highly susceptible to all concentrations of the proprietary CHG formulation used. This was true at the earliest time point tested (15 seconds). These results corroborate and quantify the effectiveness of the proprietary CHG formulation against clinical isolates of MDR and XDR gram-negative bacteria *in vitro*.

MIC data were more variable than time kill data. We observed relatively low MICs for MRSA, *E. coli*, *A. baumannii* and vancomycin-resistant *E. faecium*, but higher MICs for *P. aeruginosa and K. pneumoniae.* Lack of a consistent relationship between the CHG MIC and increased CHG killing has been described previously with MRSA [22]. The authors noted that the MIC did not appear to affect *in vitro* rate of kill or *in vivo* skin test results, and did not represent CHG resistance. The discordance between MIC and kill rate did not affect the utility of CHG decolonization in that study, and the same may well be true for the pathogens we studied. For example, even the MICs we observed against *P. aeruginosa* (e.g., 19.53 µg/mL or 1 : 512) would be at least 10-fold lower than the lowest concentrations of CHG deposited on skin when the 2% proprietary CHG formulation was used [11]. We conducted our *in vitro* evaluations using the proprietary CHG formulation specifically because the CHG impregnated cloths have been shown to provide consistently high concentrations of antiseptic coverage when applied to skin [11].

Tab. I. MICs of a proprietary CHG formulation against ten clinical isolates of MRSA, vancomycin-resistant *E. faecium,* and the XDR gram-negative pathogens *P. aeruginosa, K. pneumoniae, E. coli,* and *A. baumannii* demonstrate the susceptibility of MDR isolates.

| Clinical isolates | Number of isolates | MIC |
|---|---|---|
| *S. aureus (MRSA)* | 1 | 1 : 4096 |
| | 3 | 1 : 8192 |
| | 5 | 1 : 16384 |
| | 1 | 1 : 65536 |
| *E. faecium (VRE)* | 2 | 1 : 1024 |
| | 8 | 1 : 2048 |
| *P. aeruginosa (XDR)* | 7 | 1 : 512 |
| | 3 | 1 : 1024 |
| *K. pneumoniae (XDR)* | 7 | 1 : 512 |
| | 2 | 1 : 1024 |
| | 1 | 1 : 4096 |
| *E.coli (XDR)* | 1 | 1 : 2048 |
| | 9 | 1 : 4096 |
| *A. baumannii (XDR)* | 1 | 1 : 512 |
| | 6 | 1 : 1024 |
| | 2 | 1 : 2048 |
| *A. baumannii (MDR)* | 1 | 1 : 1024 |

## Conclusions

The present study provides *in vitro* evidence that the proprietary CHG formulation is effective against MDR gram-negative organisms, MRSA, and VRE. The solution we studied is available only in CHG-impregnated cloths which are known to provide effective protection against colonization and infection by drug resistant pathogens. Of course, the 60 clinical isolates tested would not be representative of all strains a patient might encounter in a German hospital. In addition, higher concentrations of the CHG product than were tested *in vitro* may also be more representative of the amount deposited by cloths in real-life use. Future research should evaluate the potential of the cloths to prevent MDROs from colonizing the skin and leading to hospital-acquired infections.

### Acknowledgments

Writing assistance was provided by Wendy Horn, Insight Communications Group LLC, and funded by Sage Products LLC. This research was funded by Sage Products LLC, Cary, Illinois, USA.

Partial results were presented at the 3rd International Conference on Prevention & Infection Control (ICPIC), 2015."

### References

[1] Luyt C-E, Bréchot N, Trouillet J-L, et al. *Antibiotic stewardship in the intensive care unit*. Critical Care 2014;18:480-91.

[2] Sievert DM, Ricks P, Edwards JR, et al. *Antimicrobial-re-*

*sistant pathogens associated with healthcare-associated infections: summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2009-2010.* Infect Control Hosp Epidemiol 2013;34:1-14.

[3] Mutters NT, Mersch-Sundermann V, Mutters R, et al. *Control of the spread of vancomycin-resistant Enterococci in hospitals: epidemiology and clinical relevance.* Dtsch Arztebl Int 2013;110:725-31.

[4] Nelson MU, Bizzarro MJ, Dembry LM, et al. *One size does not fit all: Why universal decolonization strategies to prevent methicillin-resistant* Staphylococcus aureus *colonization and infection in adult intensive care units may be inappropriate for neonatal intensive care units.* J Perinatol 2014;34:653-5.

[5] Schwab F, Meyer E, Geffers C, et al. *Understaffing, overcrowding, inappropriate nurse: ventilated patient ratio and nosocomial infections: which parameter is the best reflection of deficits?* J Hosp Infect 2012;80:133-9.

[6] Cheah AL, Spelman T, Liew D, et al. *Enterococcal bacteraemia: factors influencing mortality, length of stay and costs of hospitalization*. Clin Microbiol Infect 2013;19:E181-9. doi:10.1111/1469-0691.12132.

[7] Lambert ML, Suetens C, Savey A, et al. *Clinical outcomes of health-care-associated infections and antimicrobial resistance in patients admitted to European intensive-care units: a cohort study.* Lancet Infect Dis 2011;11:30-8.

[8] Mattner F, Bange F-C, Meyer E, et al. *Preventing the spread of multidrug-resistant gram-negative pathogens.* Dtsch Arztebl Int 2012;108:39-45.

[9] Climo MW, Yokoe DS, Warren DK, et al. *Effect of daily chlorhexidine bathing on hospital-acquired infection.* N Engl J Med 2013;368:533-42.

[10] Huang SS, Septimus E, Avery TR, et al. *Cost savings of universal decolonization to prevent intensive care unit infection: Implications of the REDUCE MRSA Trial*. Infect Control Hosp Epidemiol 2014;35:S23-S31.

[11] Edmiston Jr CE, Krepel CJ, Seabrook GR, et al. *Preoperative shower revisited: can high topical antiseptic levels be achieved on the skin surface before surgical admission?* J Am Coll Surg 2008;207:233-9.

[12] Huang SS, Septimus E, Kleinman K, et al. *Targeted versus universal decolonization to prevent ICU infection.* N Engl J Med 2013;368:2255-65.

[13] Milstone AM, Elward A, Song X, et al. *Daily chlorhexidine bathing to reduce bacteraemia in critically ill children: a multicentre, cluster-randomised, crossover trial*. Lancet Infect Dis 2013;381:1099-106.

[14] Vernon MO, Hayden MK, Trick WE, et al. *Chlorhexidine gluconate to cleanse patients in a medical intensive care unit: the effectiveness of source control to reduce the bioburden of vancomycin-resistant enterococci*. Arch Intern Med 2006;166:306-12.

[15] Derde LPG, Dautzenberg MJD, Bonten MJM. *Chlorhexidine body washing to control antimicrobial-resistant bacteria in intensive care units: a systematic review.* Intensive Care Med 2012;38:931-9.

[16] Cassir N, Papazian L, Fournier PE, et al. *Insights into bacterial colonization of intensive care patients' skin: the effect of chlorhexidine daily bathing*. Eur J Clin Microbiol Infect Dis 2015;34:999-1004.

[17] Cassir N, Thomas G, Hraiech S, et al. *Chlorhexidine daily bathing: Impact on health care-associated infections caused by gram-negative bacteria.* Am J Infect Control 2015. doi: 101016/jajic201502010 (Epub ahead of print)

[18] Magiorakos AP, Srinivasan A, Carey RB, et al. *Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance.* Clin Microbiol Infect 2012;18:268-81.

[19] Gebel J, Werner HP, Kirsch-Altena A. et al. *Standardmethoden der DGHM zur Prüfung chemischer Desinfektionsverfahren.* Berlin: MHP Verlag 2001.

[20] Umscheid CA, Mitchell MD, Doshi JA, et al. *Estimating the proportion of healthcare-associated infections that are reasonably preventable and the related mortality and costs.* Infect Control Hosp Epidemiol 2011;32:101-14.

[21] Quach C, Milstone AM, Perpête C, et al. *Chlorhexidine bathing in a tertiary care neonatal intensive care unit: impact on central line-associated bloodstream infections.* Infect Control Hosp Epidemiol 2014;35:158-63.

[22] Cookson BD, Bolton MC, Platt JH. *Chlorhexidine resistance in methicillin-resistant* Staphylococcus aureus *or just an elevated MIC? An in vitro and in vivo assessment*. Antimicrobial Agents Chemother 1991;35:1997-2002.

■ Received on July 29, 2015. Accepted on October 20, 2015.

■ Correspondence: Nico T. Mutters, Heidelberg University Hospital, Department of Infectious Diseases, Im Neuenheimer Feld 324, D-69120 Heidelberg, Germany - Tel. +49 0 6221 56-8208 - Fax +49 0 6221 56-5627 - E-mail: nico.mutters@med.uni-heidelberg.de

ORIGINAL ARTICLE

# Determinants of failure to thrive (FTT) among infants aged 6-24 months: a case-control study

H. HABIBZADEH[1], H. JAFARIZADEH[2], A. DIDARLOO[3]
[1] School of Nursing and Midwifery, Urmia University of Medical Sciences, Serow Highway, Nazlou, Urmia, Iran; [2] Department of Nursing, School of Nursing and Midwifery, Urmia University of Medical Sciences, Serow Highway, Nazlou, Urmia, Iran; [3] Department of Health and Preventive Medicine, School of Medicine, Urmia University of Medical Sciences, Serow Highway, Nazlou, Urmia, Iran

Key words

Risk factors • Failure to thrive • Infants

Summary

**Introduction.** *Failure to thrive (FTT) in children is one of the most important health issues around the world, especially in developing countries. Lack of success in identifying and controlling this health problem may lead to dangerous health consequences for children. The aim of this research was to explore the risk factors for this health problem in infants under two years of age in Urmia, Northwest of Iran.*

**Methods.** *This case-control study was carried out on 445 infants of 6 to 24 months (180 as cases, and 265 as controls) in Urmia, Northwest of Iran, during 2013. The study samples were selected from six health centers, using the purposeful sampling method. To collect data, a questionnaire including items regarding socio-demographics of the children's families, and demographic and nutrition-related variables of infants was utilized. To analysis data and determine the real effect of the aforementioned factors on growth status of infants, a chi-square test and logistic regression analysis were applied.*

**Results.** *The regression analysis revealed that education level of infants' mothers [AOR = 1.421, 95% CI (1.172, 1.724)], duration of breastfeeding [AOR = 1.859, 95% CI (1.212, 2.852)], birth weight of infants [AOR = 2.777, 95% CI (1.276, 7.166)], family's monthly income [AOR = 1.492, 95% CI (1.117, 2.230)] were correlated with FTT as significant risk factors (P < 0.05). Birth order of infants [AOR = .741, 95% CI ( .573- .958)], however, appeared to be a protective factor for child growth (P < 0.05).*

**Discussion.** *The findings of the study may help health care providers in designing and implementing appropriate interventions for improving children's health. In addition, taking into account the importance of healthy growth of children, educating mothers/caretakers would seem beneficial in preventing dangerous diseases in children.*

## Introduction

Children, as a large group in any community, are the most vulnerable population to diverse risk factors [1]. In the world, especially in developing countries, one of the most common problems threatening younger children is failure to thrive (FTT) [2]. Its rapid diagnosis and control can have positive health outcomes on infants and help the state of health and progress in societies [3]. FTT is seen in all socioeconomic strata, but the incidence is especially high among urban and rural families living in poverty. Infants and children who develop FTT are at a higher risk of long term growth, development, and behavior problems [4]. It is estimated that FTT affects 5-10% of young children and approximately 3-5% of children admitted to teaching hospitals [5]. The peak incidence of FTT occurs in children aged 9-24 months, with no significant gender difference [6]. *Iranian National Committee for Child Nutrition* (INCCN) revealed that FTT is still very common in urban and rural children under five years (9-15% suffer from stunting, 10.4% suffer from moderate and severe underweight, and 4.9% suffer from thinness) [7].

Infants 6 to 24 months, due to the rapid transition from breastfeeding to complementary feeding, are more likely to be at risk of FTT. Considering that the brain growth and development occur during the first six months of life, an incidence of growth failure will lead to harmful consequences such as childhood mental retardation [8].

FTT is variously defined, but it is a term used to describe inadequate growth or the inability to maintain growth, usually in early childhood. FTT is not a diagnosis, but a syndrome that results from many different medical, social, or environmental processes [9]. Kholdi et al. defines FTT as a decrease in a child's weight (a minimum of 50 grams) at each attendance, compared to the previous evaluation [10]. Various health indices such as weight for age, height for age, and weight for height have been utilized for diagnosing FTT. Recent research, however, has validated the weight-for-age approach is the simplest and most reasonable marker for FTT [11].

By measuring the weight and monitoring the growth of every child, health care providers can identify and prevent growth failure of infants and subsequent complications using suitable interventions [12]. Compared to height and arm circumference, weight is a more sensitive indicator of health status, dietary adequacy, and

growth delay or recent nutritional problems in younger children [13].

Diverse medical and non-medical causes in this age group lead to increased growth failure, which in turn results in brain impairments, learning difficulties, and an elevated risk of other life-threating problems [14].

As previously mentioned, considering deleterious consequences of infants' growth failure on their health, it is vital to identify causes of this health issue in every society. Therefore, the aim of the current study is to determine factors influencing growth failure among infants aged 6-24 months living in Urmia, northwest of Iran. It seems that the current research results would be helpful to health care providers in designing and implementing appropriate interventions for improving growth of children.

## Materials and methods

This case-control study was carried out on infants aged 6-24 months in Urmia, Northwest of Iran, during 2013. Inclusion criteria were as follows: 1) a history of weight decrease (a minimum of 50 grams) in at least two consecutive months, based on growth monitoring charts, 2) willingness of the child's family to participate in the study, 3) not consuming drugs or boosters with an effect on physical growth of the infant, 4) lack of congenital disorders in the infant, 5) lack of gestational diabetes in the child's mother. The mothers not interested in taking part in the study were excluded from the investigation. To diagnose and select infants with growth failure, investigators applied weight-for-age approach. The reason for choosing this approach was that recent research has validated weight-for-age as the simplest and most reasonable marker for FTT [15]. First Six health centers were randomly selected from all health centers located in Urmia, Iran. Then, from among those health centers, based on the definition of FTT by kholdi et al. [10] and using health care files of children, infants with a weight decrease (a minimum of 50 grams) at each attendance compared to the previous evaluation were selected and entered into the study as subjects with FTT (180 subjects as the case group). Against the case group, 360 infants without growth disorder were selected and entered into the study as the control group (two controls against one case). The study groups were matched in terms of sex and age adopting the group matching approach during the sampling process. However, due to different factors (absence of some subjects, defects in filling the questionnaire, lack of response to the instrument etc.), 95 controls were excluded from the study and 265 remained.

To collect data on risk factors related to growth failure of infants, a self-report questionnaire was developed by the authors based on the available scientific literature [10, 16-18]. The questionnaire included two subscales: 1) socio-demographic information related to the children's families (13 items), 2) demographic and nutrition-related variables of the infants (11 items). To determine validity of the questionnaire, quantitative methods of content validity such as content validity ratio (CVR) and content validity index (CVI) were utilized. The CVR is an item statistic useful in the rejection or retention of specific items. In CVR, experts are requested to specify whether an item is necessary or not. Moreover, they are asked to score each item from 1 to 3 with a three-degree range of "not necessary, useful but not essential, essential". Greater levels of content validity exist as larger numbers of panelists agree that a particular item is essential. The numeric value of content validity ratio is determined by the Lawshe table [19].

The CVI was another approach taken for measuring content validity of the research questionnaire applied by researchers. In this approach, the panel of experts are asked to rate each item of the instrument in terms of relevancy and clarity and score each item from 1 to 4 with a four-point scale of 1 = not relevant, 2 = somewhat relevant, 3 = quite relevant, 4 = highly relevant or 1 = not clear, 2 = needs some revision, 3 = clear but needs minor revision, 4 = very clear. The study tool was sent via email to 10 academics to act as a panel of experts and examine the items in terms of necessity, relevance, and clarity. Then, feedback received from the specialists was analyzed and suggested changes were made to the study tool. Overall, five items did not receive the minimum scores of CVR (0.62) and CVI (> 0.79) and were discarded from the questionnaire. Content validity of the finalized questionnaire was confirmed based on results of CVR and CVI.

To determine the reliability of the questionnaire, the study instrument was completed by interviewing the mothers of all participants from the two groups. Next, Cronbach's alpha coefficient was computed and its value was found to be 0.89. Therefore, reliability of the questionnaire was also confirmed. After collecting, organizing, and classifying the data, statistical analyses were performed using descriptive and inferential statistical methods (chi-square test for independence, also called Pearson's chi-square test, is used to discover if there is a relationship between two categorical variables. In addition, logistic regression test predicts the probability that an observation falls into one of the two categories of a dichotomous dependent variable based on one or more independent variables that can be either continuous or categorical). In this study, a P value of less than 0.05 was considered significant.

## Results

The results of this research study were evaluated at two levels. In the first step, along with crosstabing characteristics of the subjects, the univariate analysis of factors related to infants' growth was examined. In the second step, logistic regression analysis was performed to determine the net effect of risk factors on infants FTT, after adjusting for potential confounding variables.

Of the 445 eligible infants who were investigated, 180 (40.5%) had FTT and the remainder (59.5%) were

healthy. The findings showed that the mean age of mothers was 27 ± 5.75 years ranging from 16 to 48. The majority of them (95.5%) were housewives and the rest (4.5%) were working mothers. About 40% of the mothers had high school level education and the rest were from other educational levels. The majority of households (62.2%) had a monthly income of lower than USD300.

The findings highlighted that there were significant associations between undesirable growth of the subjects and their mothers' age (p-value = 0.041), mothers' education level (p-value = 0.007), number of gestation (p-value = 0.006), and family income (p-value = 0.001). These relationships were confirmed by the results of the Pearson's Chi- square test ($p < 0.05$) (Tab. I).

The present investigation indicated that 93 percent of subjects' birth weight was more than 2,500 gr, and 96.6% of infants Apgar score were more than seven at the time of birth. The majority of participants (96.4%) had been breastfed, and 73.7% of them had started the complementary feeding at the age of six months or more. Univariate analyses showed that there exist significant linkages between participants growth failure and birth weight (p-value = 0.002), duration of breastfeeding (p-value = 0.001), and birth order (p-value = 0.04). These relationships were confirmed by the results of the Pearson's Chi- square test ($p < 0.05$) (Tab. II).

To determine the real effect of the studied factors on growth failure, we applied a multivariate logistic regression model. Its results indicated that all the significant variables in univariate analyses, except for the mothers' age and number of gestations, have retained in the regression model and were statistically significant ($p < 0.05$) (Tab. III).

Education level of infants' mothers indicated a statistically significant linkage with undesirable growth. Those infants who had low literate mothers were 1.4 times more likely to have undesirable growth as compared to those children whose mothers were high literate [Adjusted Odds Ratio (AOR) = 1.421, 95% CI (1.172, 1.724)] (p-value = .000). The family's monthly income and child growth were found to have a statistically significant relationship, that is, children who had low family income were 1.5 times more likely to have undesirable growth in comparison with those who had high fam-

**Tab. I.** Sociodemographic characteristics of the study participants in two groups (n = 445).

| Variable | Groups | | #Chi-square value | *p*-value |
|---|---|---|---|---|
| | Cases | Controls | | |
| | N (%) | N (%) | | |
| **Gender of infant** | | | | |
| male | 85 (47.2) | 139 (52.5) | 1.17 | 0.28 |
| female | 95 (52.8) | 126 (47.5) | | |
| **Birth weight/BW** | | | | |
| < 2500 gr | 19 (10.6) | 9 (3.4) | 9.31 | 0.002** |
| ≥ 2500 gr | 161 (89.40) | 256 (96.6) | | |
| **Apgar score** | | | | |
| ≥ 7 | 174 (96.7) | 256 (96.6) | 0.001 | 0.97 |
| < 7 | 6 (3.3) | 9 (3.4) | | |
| **Birth age** | | | | |
| pre-term | 6 (3.3) | 6 (2.3) | 0.62 | 0.73 |
| term | 172 (95.6) | 257 (97.0) | | |
| post-term | 2 (1.1) | 2 (0.8) | | |
| **Breast feeding** | | | | |
| yes | 175 (97.2) | 254 (95.8) | 0.58 | 0.44 |
| no | 5 (2.80) | 11 (4.2) | | |
| **Duration of breast feeding** | | | | |
| < 12 months | 105 (58.33) | 112 (42.26) | 11.62 | 0.001** |
| ≥ 12months | 75 (41.67) | 153 (57.7) | | |
| **Hospitalization** | | | | |
| yes | 35 (19.4) | 50 (18.9) | 0.69 | 0.70 |
| no | 145 (80.6) | 214 (80.8) | | |
| **Starting complementary feeding** | | | | |
| < 6 months | 50 (27.8) | 67 (25.3) | 0.34 | 0.55 |
| ≥ 6 months | 130 (72.2) | 189 (74.7) | | |
| **Birth order** | | | | |
| first | 76 (42.2) | 145 (54.7) | 8.2 | 0.03* |
| second | 73 (40.6) | 88 (33.2) | | |
| third | 24 (13.3) | 23 (8.7) | | |
| above | 7 (3.9) | 9 (3.4) | | |

* $p < 0.05$, ** $p < 0.01$ is significant, # the Pearson's Chi- square test

**Tab. II.** Sociodemographic characteristics of family of the study participants in two groups (n = 445).

| Variable | Groups | | #Chi-square value | *p*-value |
|---|---|---|---|---|
| | Cases | Controls | | |
| | N (%) | N (%) | | |
| **Marriage age of mother (in years)** | | | | |
| < 15 | 21 (11.7) | **23 (8.7)** | 1.14 | 0.56 |
| 15-25 | 144 (80.0) | **221 (83.4)** | | |
| > 25 | 15 (8.3) | **21 (7.9)** | | |
| **Mother occupation** | | | | |
| employment | 5 (2.8) | **15 (5.7)** | 2.32 | 0.31 |
| housewife | 175 (97.2) | **250 (94.3)** | | |
| **Family monthly income** | | | | |
| < 300 USD | 124 (68.9) | **136 (51.3)** | 13.69 | 0.001** |
| 300-600 USD | 50 (27.8) | **117 (44.2)** | | |
| > 600 USD | 6 (3.3) | **12 (4.5)** | | |
| **Maternal age (in years)** | | | | |
| > 20 | 21 (11.7) | **14 (5.3)** | 9.98 | 0.04* |
| 20-24 | 43 (23.9) | **86 (32.50** | | |
| 25-29 | 65 (36.1) | **86 (32.5)** | | |
| 30-34 | 30 (16.7) | **54 (20.4)** | | |
| > 34 | 21 (11.7) | **25 (9.4)** | | |
| **Type of delivery** | | | | |
| vaginal | 116 (64.4) | **157 (59.2)** | 1.8 | 0.40 |
| cesarean | 64 (35.6) | **108 (40.8)** | | |
| **Maternal education** | | | | |
| illiterate | 11 (6.1) | **8 (3.0)** | 15.85 | 0.007** |
| elementary | 36 (20.0) | **26 (9.8)** | | |
| guidance | 28 (15.6) | **37 (14.0)** | | |
| high school | 68 (37.8) | **110 (41.5)** | | |
| university | 37 (20.6) | **84 (31.7)** | | |
| **Number of gestation/NG** | | | | |
| 1-2 | 115 (63.9) | **192 (72.5)** | 10.08 | 0.03* |
| 3-4 | 54 (30.0) | **70 (26.4)** | | |
| 4 and above | 11 (6.1) | **3 (1.1)** | | |
| **Gestation interval (in years)** | | | | |
| 1-2 | 110 (61.1) | **141 (53.2)** | 2.80 | 0.42 |
| 3-4 | 28 (15.6) | **48 (18.1)** | | |
| 5-6 | 20 (11.1) | **38 (14.3)** | | |
| > 6 | 22 (12.2) | **38 (14.3)** | | |
| **Number of family children** | | | | |
| 1 | 90 (50.0) | 130 (49.1) | 1.96 | 0.58 |
| 2 | 61 (33.9) | 98 (37.0) | | |
| 3 | 20 (11.1) | 30 (11.3) | | |
| ≥ 4 | 9 (5.0) | 7 (2.6) | | |

* p < 0.05, ** p < 0.01 is significant, # the Pearson's Chi- square test

ily income [AOR = 1.492, 95% CI (1.117, 2.230)] (p-value = 041). Furthermore, birth weight of the infants had a statistically significant association with their undesirable growth. Those infants with a low birth weight were 2.7 times more likely to have undesirable growth in comparison with those who had a high birth weight [AOR = 2.777, 95% CI (1.276, 7.166)] (p-value = .035). Moreover, duration of breastfeeding was associated with growth of infants. That is, children breastfed for a short-term period were1.8 times more likely to have undesirable growth and vice versa [AOR = 1.859, 95% CI (1.212, 2.852)] (p-value = .001). Birth order also had a statistically significant association with the undesirable growth of the infants [AOR = .741, 95% CI (.573-.958)] (p-value = .022) (Tab. III).

## Discussion

Growth failure in children, especially among infants under two years of age, is in a state of emergency. Health care providers should focus on this issue and find the relevant factors. With this work, we hoped to aid health experts in managing and controlling FTT problem before it develops into dangerous and threatening conditions in children. The study aimed to explore factors influencing growth failure among infants aged 6-24 months in Urmia, northwest of Iran in 2013.

The results revealed that the prevalence of growth failure in female infants was somewhat more than that in the males. This finding was consistent with the study of Mohammad poor Asl et al. [3] and was also supported

**Tab. III.** Results of logistic regression analysis on growth status of infants (n = 445).

| Variable | B | SE | Wald | df | AOR[#] (95.0% C.I)[£] | *p*-value |
|---|---|---|---|---|---|---|
| Gender of infant | -0.083 | 0.210 | 0.154 | 1 | 0.921 (0.610- 1.391) | 0.695 |
| Birth weight/BW | 1.021 | 0.484 | 4.461 | 1 | 2.777 (1.076-7.166) | 0.035* |
| Apgar score | 0.243 | 0.660 | 0.136 | 1 | 1.276 (0.350-4.648) | 0.712 |
| Birth age | 0.144 | 0.608 | 0.056 | 1 | 1.155 (0.351-3.803) | 0.813 |
| Breast feeding | 0.198 | 0.619 | 0.102 | 1 | 1.219 (0.362-4.102) | 0.749 |
| Duration of breast feeding | 0.620 | 0.218 | 8.060 | 1 | 1.859 (1.212-2.852) | 0.001** |
| Hospitalization | 0.144 | 0.285 | 0.256 | 1 | 1.155 (0.660-2.021) | 0.613 |
| Starting complementary feeding | 0.184 | 0.239 | 0.596 | 1 | 1.202 (0.753-1.920) | 0.440 |
| Birth order | -0.300 | 0.131 | 5.229 | 1 | 0.741 (0.573-0.958) | 0.022* |
| Marriage age of mother (in years) | 0.072 | 0.289 | 0.062 | 1 | 1.074 (0.610-1.894) | 0.804 |
| Mother occupation | 0.146 | 0.409 | 0.128 | 1 | 1.157 (0.520-2.579) | 0.721 |
| Family monthly income | 0.400 | 0.196 | 4.191 | 1 | 1.492 (1.017-2.190) | 0.041* |
| Maternal age (in years) | 0.025 | 0.143 | 0.030 | 1 | 1.025 (0.775-1.355) | 0.863 |
| Type of delivery | 0.084 | 0.218 | 0.148 | 1 | 1.088 (0.709-1.668) | 0.700 |
| Maternal education | 0.351 | 0.098 | 12.740 | 1 | 1.421 (1.172-1.724) | 0.000** |
| Number of gestation/NG | -0.233 | 0.159 | 2.235 | 1 | 0.587 (0.353-0.975) | 0.060 |
| Gestation interval (in years) | 0.189 | 0.132 | 2.043 | 1 | 1.208 (0.932-1.565) | 0.153 |
| Number of family children | 0.196 | 0.216 | 0.822 | 1 | 1.216 (0.797-1.857) | 0.365 |

* $p < 0.05$, ** $p < 0.01$ is significant, [#] Adjusted Odds Ratio, [£] Confidence interval

and confirmed by the results of Hajian's study [20]. However, our findings were in contrast with that of Vaghari's study [21]. One interpretation would be that such a difference may arise from the social and cultural discrimination between genders. That is, some families pay more attention to nutrition of male children compared to that of the females, which in turn may potentiate growth failure and other health problems in female infants [3].

The present investigation indicated that factors such as education level of the mother, duration of breastfeeding, birth order, birth weight, and family income were associated with the failure growth of the studied infants. Among these variables, education level of mothers was the first and most important predictor of growth failure of children. Infants with illiterate or low literate mothers experienced greater growth failure compared to other children.

Previous studies have also highlighted that in comparison with illiterate or low literate mothers, the high literate mothers had infants with a more desirable growth [22]. Maternal education, even without considering other factors affecting the growth, has a significant impact on improving child growth. In developing countries, insufficient knowledge of parents, and not child malnutrition, is the determining factor [23]. The study conducted by Hameida et al. on Lybian children showed that illiterate mothers had a major role in malnutrition and growth failure of their children [24].

In two independent studies, underweight was more prevalent in children under the age of two whose mothers were illiterate [25]. Likewise, underweight was negatively associated with the parent's education level among children aged 6-30 months [26]. Results of these studies were consistent with our research findings.

The logistic regression results showed that duration of breastfeeding was another risk factor influencing growth failure. Infants who were breastfed for a short-term period experienced more growth failure compared to those breastfed for a long-term period. With the cessation of breastfeeding, babies were fed with supplementary food. Therefore, if healthy food choices are not made for children or if supplementary feeding of children does not start at the right time, it will have a major effect on weight loss of children. Another study has also indicated and confirmed the relationship between cessation of breastfeeding and growth retardation [16]. There is a concordance between findings of these studies and this part of our results.

Breast milk is uniquely suited to the human infant's nutritional needs and is a live substance with immunological and anti-inflammatory properties that protect both the mother and her child against a host of illnesses and diseases [27]. Increasing the rate of breastfeeding can help reduce the prevalence of various illnesses and health problems, which in turn results in lowering health care costs [28].

Children with low birth weight are more susceptible to diseases, experience breastfeeding failures more frequently, and are at a higher risk of growth disorders, as compared to children with appropriate weight at birth [29]. In the present investigation, the infants with a low birth weight (lower than 2,500 gr) experienced more growth failure in comparison with those who had a birth weight higher than 2,500 gr. A low birth weight was found to be a risk factor for growth of children, according to regression results. This finding was confirmed and supported by the study of Marilia de Carvalho Lima et al. They found that there exists a significant correlation between malnutrition and children's birth weight. The largest risk for malnutrition was found in children with a low birth weight; a risk that was around six times higher than that for children with a birth weight of 3,500 gr or

more [17]. Similarly, the study conducted in a rural area of Western Ethiopia showed that children with a low birth weight (LBW) were about 2.4 times more likely to be underweight as compared to high birth weight children [1]. The findings of these studies were in accordance with our research findings. The findings of the current study also showed that household income had a significant relationship with growth failure in children. This was consistent with results obtained in other studies [18]. For this reason, it seems that most children with growth failure or retardation of growth live in families with low economic status.

Ultimately, the results of the present study showed a statistically significant association between birth order of children and their growth status. Indeed, birth order appeared to be a protective factor for growth of children. Infants who were in the first and second order of birth experienced more growth failure compared to children of the third order and so on. This result was consistent with findings of the study by Mohamad Por Asl et al. [30].

The present study, like other research studies, has some limitations. Firstly, a case-control design is not a strong method for understanding cause and effect relationships between variables. It is necessary to assess and identify these associations by applying stronger epidemiological methods. Secondly, the data of this study were collected using a self-report questionnaire. Participants may have underestimated or overestimated risk factors related to FTT, which may have affected the study findings. Thirdly, although weight for age index is the simplest way to identify FTT of children, it is not a comprehensive index. It is recommended that a combination of growth monitoring indices for identifying infants with FTT be used.

## Conclusions

The present study highlighted that education level of the mother, duration of breastfeeding, birth order, birth weight, and family income were independently associated with growth status of the studied infants. Thus, national public health intervention programmers and health care providers working on improving child growth should focus on these determinants to reduce growth failure of children. In addition, taking into account the importance of healthy growth of children, it seems that educating mothers/caretakers would be beneficial in preventing dangerous diseases in children.

### Acknowledgment

We would like to thank the mothers, interviewers, and personnel of health centers of Urmia who participated in this study and all the people who kindly helped us in conducting this research.

### References

[1] Asfaw M, Wondaferash M, Taha M, Dube L. *Prevalence of undernutrition and associated factors among children aged between six to fifty nine months in Bule Hora district, South Ethiopia.* BMC Public Health 2015;15:41.

[2] Young MF, Martorell R. *The public health challenge of early growth failure in India.* Eur J Clin Nutr 2013;67:496-500.

[3] Mohammadpoorasl A, Sahebihag MH, Rostami F, et al. *Factors related to undesirable growth of 6 month – 2 years old children in Tabriz-Iran*. J Gorgan Uni Med Sci 2010;12: 45-50. [Article in Persian]

[4] Mahendra D. Shah, M.D. *Failure to Thrive in Children*. J Clin Gastroenterol 2002;35:371-4.

[5] Baucher H. *Failure to thrive*. In: Kliegman RM, Behrman RE, Jenson HB, et al. (eds). *Nelson Textbook of Pediatrics*, 18th ed. Philadelphia: Saunders Elsevier 2007, p. 184-7.

[6] Daniel M, Kleis L, Cemeroglu AP. *Etiology of failure to thrive in infants and toddlers referred to a pediatric endocrinology outpatient clinic.* Clin Pediatr (Phila) 2008;47:762-5.

[7] National Committee for Children Nutrition. *Improved nutrition and growth of children.* 6th ed. Tehran: Health ministry and medical education 2001, p. 39-51.

[8] Grummer-Strawn LM, Scanlon KS, Fein SB. *Infant feeding and feeding transitions during the first years of life*. Pediatrics 2008;122:36-42.

[9] Cole SZ, Lanham JS. *Failure to Thrive: An Update*. Am Fam Physician 2011;83:829-34.

[10] Kholdi N, Zayeri F, Akbarzadeh Bagheban A, et al. *A study of growth failure and its related factors in children from 0 to 2 years in Tehran, Iran*. Turk J Pediatr 2012;54:38-44.

[11] Krugman SD, Dubowitz H. *Failure to Thrive*. Am Fam Physician 2003;68:879-84.

[12] Natale V, Rajagopalan A. *Worldwide variation in human growth and the World Health Organization growth standards: a systematic review*. BMJ Open 2014;4:1-7.

[13] Hosseini MK, Norbala M, Ahmed A. *Trend of growth change (weight and height) of Children 2-18 years old in Iran, during nearly one month (years 1990 to 1999).* Hakim Research Journal 2006; 9(1):1-10.

[14] Victora CG, Adair L, Fall C, et al. *Maternal and child undernutrition: consequences for adult health and human capital.* Lancet 2008;371:340-57.

[15] Raynor P, Rudolf MC. *Anthropometric indices of failure to thrive*. Arch Dis Child 2000;82:364-5.

[16] Bloss E, Wainaina F, Bailey RC. *Prevalence and predictors of underweight, stunting, and wasting among children aged 5 and under in western Kenya*. J Trop Pediatrics 2004;50:260-9.

[17] de Carvalho Lima M, Farias Almeida Motta ME, Cavalcanti Santos E, et al. *Determinants of impaired growth among hospitalized children – a case-control study*. Sao Paulo Med J 2004;122:117-23.

[18] Victora CG, Vaughan JP, Kirkwood BR, et al. *Risk factors for malnutrition in Brazilian children: the role of social and environmental variables.* Bull World Health Organ 1986;64:299-309.

[19] Lawshe CH. *A quantitative approach to content validity*. Pers Psychol 1975;28:563-75.

[20] Hajian K, Sajjadi P, Alipour A, et al. *A study on the growth of infants from birth up to the age of two years in rural communities of Babol, 1998*. JMUMS 2002;34:57-65. [Article in Persian]

[21] Vaghari GR, Ahmadpour M, Vakili MA. *Assessment of height and weight in children under 6 years in rural areas of gorgan, 1998*. JMUMS 2002;12:66-72. [Article in Persian]

[22] Waters H, Saadah F, Surbakti S, et al. *Weight-for-age malnutrition in Indonesian children, 1992-1999*. Int J Epidemiol 2004;33:589-95.

[23] Hossianzadeh K, Yousefi H, Por Mehr A, et al. *Growth trend of infants aged 0-24 months in health care centers of Hormozghan, Iran.* Medical journal of Hormozghan 2006;9:29-33.

[24] Hameida J, Billot L, Deschamps JP. *Growth of preschool children in the Libyan Arab Jamahiriya: regional and sociodemographic differences*. East Mediterr Health J 2002;8:458-69.

[25] Jesmin A, Yamamoto SS, Malik AA, et al. *Prevalence and determinants of chronic malnutrition among preschool children: a cross-sectional study in Dhaka City, Bangladesh.* J Health Popul Nutr 2011;29:494-9.

[26] Yusuke Kamiya. *Socioeconomic Determinants of Nutritional Status of Children in Lao PDR: Effects of Household and Community Factors*. J Health Popul Nutr 2011;29:339-48.

[27] Lawrence RA, Lawrence RM. *Breastfeeding: a guide for the medical profession*. 7th ed. Philadelphia: Saunders 2010.

[28] U.S. Department of Health and Human Services. *Healthy people 2010*. 2nd ed. Washington, DC: U.S. Government Printing Office 2000.

[29] Lunn PG. *The impact of infection and nutrition on gut function and growth in childhood.* Proc Nutr Soc 2000;59:147-54.

[30] Mohamad Por Asl A, Sahebi Hagh MH, Rostami F. *Factors related to undesirable growth of 6 month – 2 years in Tabriz, Iran*. GOUMS 2010;12:45-50. [Article in Persian]

■ Received on July 16, 2015. Accepted on September 22, 2015.

■ Correspondence: Alireza Didarloo, Department of Health and Preventive Medicine, School of Medicine, Urmia University of Medical Sciences, Serow Highway, Nazlou, Urmia, Iran - Tel. +98 4432752372 - Postal code: 5715799313 - Fax +98 4432780801 - E-mail: didarloo_a@yahoo.com

ORIGINAL ARTICLE

# Nitrates in drinking water: relation with intensive livestock production

M. GIAMMARINO[1], P. QUATTO[2]
[1] Department of Prevention, Asl CN 1, Veterinary Service, Area Animal Sanity, Racconigi (CN), Italy; [2] Department of Economics, Management and Statistics, University of Milano-Bicocca, Milano, Italy

Key words

Nitrates • Drinking water • Fish

Summary

**Introduction**. *An excess of nitrates causes environmental pollution in receiving water bodies and health risk for human, if contaminated water is source of drinking water. The directive 91/676/CEE [1] aims to reduce the nitrogen pressure in Europe from agriculture sources and identifies the livestock population as one of the predominant sources of surplus of nutrients that could be released in water and air. Directive is concerned about cattle, sheep, pigs and poultry and their territorial loads, but it does not deal with fish farms. Fish farms effluents may contain pollutants affecting ecosystem water quality.*
**Methods**. *On the basis of multivariate statistical analysis, this paper aims to establish what types of farming affect the presence of nitrates in drinking water in the province of Cuneo, Piedmont, Italy. In this regard, we have used data from official sources on nitrates in drinking water and data Arvet database, concerning the presence of intensive farming in the considered area. For model selection we have employed automatic variable selection algorithm.*
**Results and discussion**. *We have identified fish farms as a major source of nitrogen released into the environment, while pollution from sheep and poultry has appeared negligible. We would like to emphasize the need to include in the "Nitrate Vulnerable Zones" (as defined in Directive 91/676/CEE [1]), all areas where there are intensive farming of fish with open-system type of water use. Besides, aquaculture open-system should be equipped with adequate downstream system of filtering for removing nitrates in the wastewater.*

## Introduction

Organic wastes are utilized in agriculture for improving the soil properties and as nutrient for growing crops [2]. Nitrogen surpluses increase levels of nitrogen compounds in surface and groundwater and become a health risk for animals [3] and humans, when polluted water is used to produce of drinking water [4]. In humans nitrate is reduced to nitrite that convert hemoglobin to methemoglobin, unable to transport oxygen [5]. The most appropriate means of controlling nitrate concentrations is the prevention of contamination [6]. Nitrate concentrations have increased in Europe in the last years and have doubled over the past 20 years. In the United Kingdom an average annual increase of 0.7 mg/l has been observed in some rivers [7]. In Denmark and the Netherlands nitrate concentrations are increasing by 0.2-1.3 mg/l per year [8]. In order to reduce excess of nitrogen, due to assessed agricultural sources and livestock farming, the Directive 91/676/CEE [1] was adopted by the European Commission. Such "Nitrates Directive" defines the "Nitrate Vulnerable Zones" as "areas draining into waters which are or could be affected by pollution and which contribute to pollution by intensive use of fertilizers or intensive livestock production". Nitrates Vulnerable Zones are areas where nitrate concentration in ground and in drinking water amount to 50 mg/l or more. In these zones the intensive use of fertilizers and the high animal densities contribute to environmental nitrate pollution, as acknowledged by European Water Framework Directive [9]. For reducing nitrate concentrations, specific and programmatic agricultural practices are adopted. To this purpose the Directive cares to report densities of some animal species farmed in Europe: cattle and dairy cattle, swine, poultry and sheep. However, the Directive does not seem to consider the potential environmental impact of fish farms, despite the fact Ling et al. [10] have shown that aquaculture is the major contributor to increasing the level of ecosystem pollution. According to Troell et al. [11], fish farms represents continuous or intermittent source of pollution. As a matter of fact, effluents may contain pollutants which have significant effects on water quality of ecosystem, especially when "flowthrough" o "open-system type" is adopted, discharging residues into water bodies surrounding [12]. Nitrogen compounds are the most widespread contaminants released by these types of breeding. Ryther and Dunstan [13] have shown that this type of pollution is a limiting factor in the growth of phytoplankton in costal and estuarine habitats and then it may lead to eutrophication. Therefore, it is reasonable to assume that the groundwater may be contaminated, thus having important effects on water for human consumption.

## Methods

The analysis has focused on the drinking water supplied by water companies of 86 local municipalities in the province of Cuneo (44° 23' 00'' N, 7° 33' 00'' E). Data refer to 2012 and are obtained from Alpi Waters Spa website [14] and from the site www.cheacquabeviamo.it [15].
The studied area is characterized by relatively high degree of agricultural and breeding activity. Natural water is extracted from underground by about a thousand springs or wells, drilled in the mountains in the plains and starts in the ducts through tanks and reservoir. A small percentage comes from surface water. We have investigated the relationship between the content of nitrates in drinking water and the number of farms and the animal densities. These data refer to 2012 and are obtained from the Arvet dataset of the local Veterinary Service (ASL CN1).
Statistical analysis were performed using SAS program.

## Results

The quantity of nitrate in the considered municipalities ranges from a minimum of 0.6 mg/l of drinking water and a maximum of 43.9 mg/l of drinking water. The mean, median and mode are respectively 11,01 mg/l, 6,95 mg/l, 1,5 mg/l (standard deviation 10,35; Skewness 1,45; Kurtosis 1,329). Using PROC UNIVARIATE we have rejected the null hypothesis of normal distribution (Shapiro-Wilk: W = 0,815 p < 0,0001; Kolmogorov-Smirnov: D = 0,198, p < 0,0100). For this reason, we have applied the logarithmic transformation to the data and then we have accepted the hypothesis of normality for the transformed dataset (Shapiro-Wilk: W = 0,984 p = 0,3865; Kolmogorov-Smirnov: D = 0,060, p > 0,1500).
According to the literature, we have investigated the relationship between the transformed Nitrate variable and the pollution from fertilizers and livestock or industrial wastewater. For this purpose, we have performed a simple regression analysis using the elevation of the municipalities, as the use of fertilizers and the presence of the waste decrease with the increase of altitude.
We have obtained the following equation:

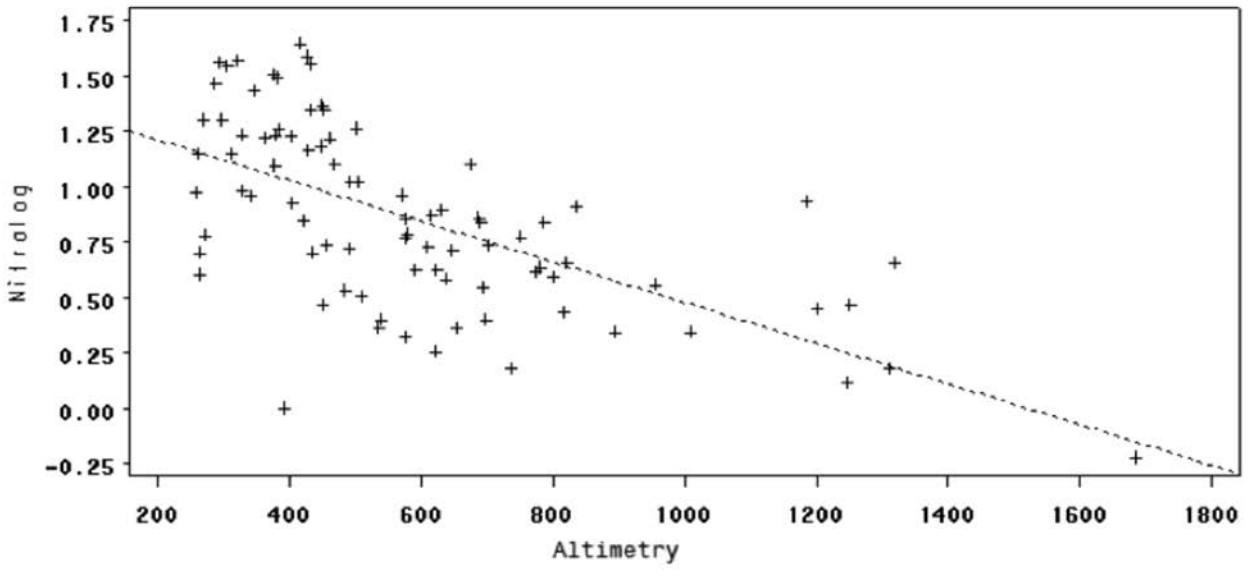
$$\text{Nitralog} = 1.39340 - 0.00091659*(\text{altimetry})$$

from which we can conclude that the logarithm of Nitrate concentration (Nitralog) negatively depends on the altimetry (Fig. 1). The t-test are significant ($p < 0.0001$) for both parameters.
Then we have considered several parametric models for predicting the dependence of the Nitralog variable on different types of farm. In particular, we have considered the following independent variables: altimetry, densities of cattle, sheep, horse, pig, poultry, rabbit and the density of fish farms. For the model selection we have employed the stepwise procedure of PROC REG. Using the adjusted R-squared (= 0.6016) and the Cp method (Cp = 3.8949), we have obtained the following model:

$$\text{Nitralog} = + (\text{cattle density}) + (\text{fish farms density}) + (\text{altimetry})$$

For the three independent variables the F-test are significant ($F = 35.68$, $p < 0.0001$; cattle density: $p < 0.0001$; fish farms density $p = 0.0183$; altimetry: $p = 0.0002$).

**Fig. 1.** Regression analysis: Altimetry vs Nitralog.



Tab. I. Estimates of the parameters with standard errors and p-values.

| Parameter | Estimate | Standard error | *p*-value |
|---|---|---|---|
| β0 | 0.95882 | 0.09911 | < 0.0001 |
| β1 | 0.00171 | 0.00033369 | < 0.0001 |
| β2 | 2.34190 | 0.96341 | 0.0183 |
| β3 | -0.00049899 | 0.00012573 | 0.0001 |

In the Table I we have reported estimates, their standard errors and p-values of the parameters.
The resulting model is given by:

$$\text{Nitralog} = 0.95882 + 0.00171*(\text{cattle density}) + 2.34190*(\text{fish farms density}) - 0.00049899*(\text{altimetry})$$

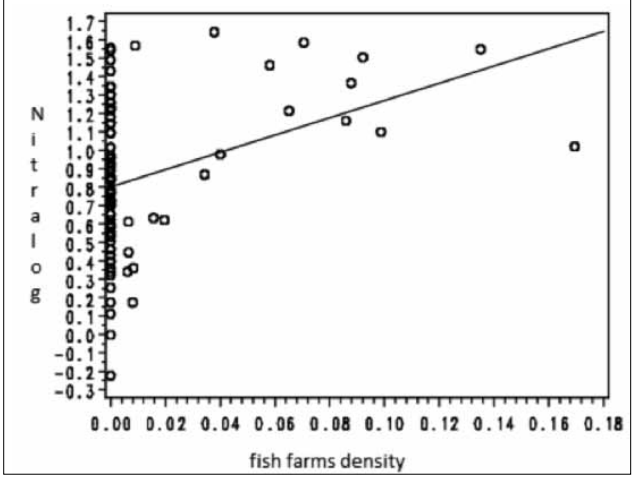
and explains 59% of the variability of Nitralog (= 0.5878; adjusted = 0.5727).

By means of PROC CORR we have shown in Figure 2 a positive correlation between the presence of nitrates in drinking water and density of fish farms, despite the fact that the correlation indices are not high (Pearson = 0.35854; Spearman = 0.24225).

## Discussion and conclusions

Nitrate pollution is widespread in the lowlands where livestock farms are more common than in mountain and foothill areas. About 59% of the nitrate variability is explained by the proposed model. On the basis of available data, it may be that there are other sources of nitrate pollution not included in the model (e.g. industrial and domestic effluents). However, we think that fish farms could play a role in increasing this type of pollution. As a matter of fact, large quantities of water that has not been properly filtered and has hosted a high concentration of farmed fish with protein feed, can release into to the environment a significant amount of nitrates and pollute the groundwater.
So, we believe that authorities responsible for controlling nitrates in surface and deep waters and for identifying the "Nitrate Vulnerable Zones", should also consider all areas where aquaculture is widespread. Moreover, such areas should be protected, making sure that the fisheries that adopt "open" farming method are equipped with appropriate filters for the uptake of nitrate and with systems for removing excess nutrients and fecal material before releasing the exploited water into the environment. This measure could be included as a part of good agricultural practices aimed at reducing the concentration of nitrates in water bodies of ecosystems.

Fig. 2. Regression analysis: Fish farms density vs Nitralog.




### References

[1] European Commission. 1991. *Council Directive Concerning the Protection of Waters Against Pollution Caused by Nitrates from Agricultural Sources* (91/676/EEC). http://europa.eu.int/comm/environment/water/water-nitrates/index_en.html.

[2] Westerman PW, Bicudo JR. *Management considerations for organic waste use in agriculture*. Bioresour Technol 2005;96:215-21.

[3] Costagliola A, Roperto F, Benedetto D, et al. *Outbreak of fatal nitrate* toxicosis (Foeniculum vulgare) *in cattle farmed in Campania region (southern Italy).* Env Sci Pollut Res Int 2014 Jan 23. [Epub ahead of print]

[4] Grossman MR. *Nitrates from agriculture in Europe: the EC Nitrates Directive and its implementation in England*. BC Envtl Aff L Rev 2000;27:567-630.

[5] Winton EF, Tardiff RG, McCabe LJ. *Nitrate in drinking water*. J Am Water Works Assoc 1971;63:95-8.

[6] Schmoll O. *Protecting groundwater for health. Managing the quality of drinking-water sources*. London: IWA Publishing, on behalf of the World Health Organization 2006.

[7] Young CP, Morgan-Jones M. *A hydrogeochemical survey of the chalk groundwater of the Banstead area, Surrey, with particular reference to nitrate*. J Inst Water Eng Scient 1980;34:213-36.

[8] WHO. *Health hazards from nitrate in drinking-water.* Report on a WHO meeting, Copenhagen, 5-9 March 1984. Copenhagen: WHO Regional Office for Europe (Environmental Health Series 1985;1).

[9] European Parliament and Council. *Parliament and Council Concerning the Establishment of a Framework of Community Action in the Field of Water Policy* (2000/60/EC).

[10] Ling C, Weimin W, Yi Y, et al. *Environmental impact of Aquaculture and countermeasures to aquaculture pollution in China*. Env Sci Pollut Res 2007;14:452-62.

[11] Troell M, Ronnback P, Halling C, et al. *Ecological engineering in aquaculture: use of seaweeds for removing the nutrients from the intensive mariculture.* J Appl Phycol 1999;11:89-97.

[12] Balderston WL, Sieburth j Mcn. *Nitrate removal in closed-system aquaculture by columnar denitrification applied and environmental microbiology*. Applied and Environmental Microbiology 1976, p. 808-18.

[13] Ryther JH, Dunstan WM. *Nitrogen, Phosphorous and eutrophication in the coastal marine Environment.* Science 1971;171:1008-13.

[14] www.alpiacque.it

[15] www.cheacquabeviamo.it

■ Received on July 6, 2015. Accepted on July 28, 2015.

■ Correspondence: Mauro Giammarino, Department of Prevention, Asl CN 1, Veterinary Service, Area Animal Sanity, Racconigi, piazza Santa Maria 1, 12038 Racconigi (CN) Italy - E-mail: mauro.giammarino@aslcn1.it

ORIGINAL ARTICLE

# Occurrence of Acrylamide in breakfast cereals and biscuits available in Italy

R. CAPEI, L. PETTINI, A. LO NOSTRO, G. PESAVENTO
Department of Health Sciences, University of Florence, Italy

Key words

Acrylamide • Biscuits • Breakfast cereals

Summary

**Introduction.** *Acrylamide, produced during thermal processing of carbohydrate-rich foods, is classified as "probably carcinogenic to humans" by the International Agency for Research on Cancer.*
**Objective.** *The aim of this study was to determine acrylamide levels in biscuits and breakfast cereals considering the widespread use of these products for all ages.*
**Method.** *Acrylamide determination was carried out in 56 samples by HPLC-UV technique.*
**Results and discussion.** *The results showed a considerable variability in the contents of acrylamide in the samples analysed, most likely due to differences in industrial processing and ingredients. The percentages of contaminated samples tested were very high (95.5% of the biscuits and 75% of the breakfast cereals) with a wide range of contamination: from 30 µg/kg to 940 µg/kg. Our results showed that 22.7% of biscuits and 33% of breakfast cereals exceeded the indicative values recommended by EC 2013/647 set at 500 µg/kg and 200-400 µg/kg (according to the composition) respectively.*
**Conclusions.** *Our findings suggest concern about the risk for human health.*

## Introduction

Acrylamide (AA) was first detected in heat-treated foods, especially potato chips, at high concentrations (mg/kg) by a group of Swedish researches [1-3]. This finding aroused great interest in all the world because of the well-known toxicity of AA, classified as "probably carcinogenic to humans" (Group 2A) by the International Agency for Research on Cancer [4] and as an A3 carcinogen (confirmed animal carcinogen with unknown relevance to humans) by *The American Conference of Governmental Industrial Hygienists* (as cited in *Agency for Toxic Substances and Disease Registry* [5]). Experimental studies in animals have shown that acrylamide causes chronic nephrotoxicity, adverse reproductive effects, and is characterized by genotoxic and carcinogenic activity mainly in glands: thyroid, mammary, adrenal, pituitary etc. [6-9]. The genotoxic activity of acrylamide seems to be partly mediated by its main metabolite, glycidamide [10-12], which is mutagenic [13-15].
The risk of cancer due to AA intake is unclear; the epidemiological studies on humans up to present have reported conflicting data and do not provide any substantial evidence that dietary exposure to AA is associated with cancer in humans, therefore more data are necessary to better assess the risk from dietary intake [7, 16], whereas the neurotoxicity activity of AA in humans is rather well known in accidental or occupational exposure through inhalation or dermal absorption [17].
AA is produced during thermal processing, at temperatures of 120°C or higher and low moisture, mainly via the Maillard reaction, responsible for the flavour and colour of cooked foods, between amino acids (mainly free asparagine) and reducing sugars, such as glucose and fructose [1, 7, 18, 19]. However, several other pathways and precursors have been proposed to contribute to AA formation as reviewed by Keramat et al. [20]. Most of AA is formed rapidly during final phases of baking, grilling or frying processes [21] of carbohydrate-rich foods such as fried potato and bakery products, biscuits, breakfast cereals, cocoa and coffee.
After ingestion, AA is rapidly absorbed and distributed to different body tissues and organs such as muscle, liver, heart, thymus, brain, kidneys, placenta and breast milk, thus also representing a significant risk to the health of foetus and infants [22, 23].
AA is metabolized by the liver to the reactive genotoxic epoxide, glycidamide. AA and glycidamide are then conjugated with glutathione, following the classical detoxification scheme, and eliminated in urine as mercapturic acid derivatives [24]. Since the urinary excretion is about 50% in the first 24 hours and 90% within 7 days, the AA metabolites could be used as an intake biomarker as recent exposure to AA [16, 23, 25, 26].
As stated above, with concern for the potential risk on human public health, and in particular in babies and adolescents, since 2003, at the request of the European Commission [27], many studies have been carried out worldwide both on the development of analytical methods and on the determination of AA in food. According to this Recommendation, furthermore, Member States should carry out AA monitoring annually and provide

the relative data to the *European Food Safety Agency* (EFSA) by June 1st of the following year.
The recent EFSA Scientific Report [28] showed the results of the monitoring of AA in specific food categories from 2007 to 2010 and the estimation of the trend in AA levels by comparing results from 2007, 2008, 2009 and 2010 in main food categories and sub-categories. Data collected showed only a few changes in the trend of AA values: a decrease in 'processed cereal based foods for infants and young children' and an increase in 'coffee and coffee substitutes'. The highest AA levels have been found in fried potato products, bread, bakery products and coffee. In 2010 the AA mean values ranged from 31 μg/kg in 'other processed cereal based foods for infants and young children' to 1350 μg/kg in 'coffee substitutes', while breakfast cereals and biscuits had mean values of 138 μg/kg and 289 μg/kg respectively. The last Scientific Opinion on AA in Food by EFSA [29] published in 2015, confirms that the major contributing foods to AA exposure are potato products, coffee and bakery products (biscuits and bread) and relates that dietary AA intake is estimated in the range of 0.3-1.9 μg/kg b.w. per day for the general population, and for the 95th percentile exposure from 0.6 to 3.4 μg/kg b.w. per day. EFSA concludes that "AA in food potentially increases the risk of developing cancer for consumers in all age groups".
Within the *European Prospective Investigation into Cancer and Nutrition* (EPIC) study, Freisling et al. [30] estimated the AA intake in 10 European countries: the study, involved a total of 39 994 participants, aged 35-74, showed that the mean intake ranged from 12 to 39 μg per day for women and from 13 to 47 μg for men.
Moreover, a substantial probability that the intake by children and adolescents is remarkably higher than that of adults must be considered, due to their more frequent consumption of certain foods such as fried potato products and biscuits.
A large group of products contributing to AA intake is cereal-based food. In Italy, to our knowledge, the data regarding dietary AA exposure principally concerned potato chips, which are estimated to contribute about 30% of the total diet [31-33] and coffee [34], whereas cereal-based foods have been investigated less in recent years.
A maximum limit for AA in drinking water (0.5 μg/L) has been established by WHO [35], but no limit is established in food.
In 2011, the European Commission [36] established indicative values, as amended by Recommendation EC of 2013 [37], for the food categories cited in Recommendation EC of 2010 [38] including 'french fries', 'potato crisps', 'soft bread', 'breakfast cereals', 'biscuits, crackers, wafers, crisp bread and similar', 'roast coffee', 'instant coffee', 'biscuits and rusks for infants and young children', 'processed cereal based foods for infants and young children' and 'baby foods, other than processed cereal based foods', with the exception of the products for home cooking and 'other products'. This Recommendation stresses that 'these values are to be understood as indicative of the need for an investigation and not as safety thresholds'.
In Italy, the second largest item of expenditure in the cereals derivatives segment is products for breakfast (30% of total). 95% of Italians buys biscuits at least once a year. In 2013 an increase in the volume of purchases of bakery products was observed (+ 3.2%) compared to 2012 [39]. Furthermore, although in Italy the breakfast cereals market is not yet as developed as in other countries, over the last few years we have witnessed a process of strong sales growth in volume. In particular, from 2004 to 2007, the share of household spending on breakfast cereals has increased. In 2007, fortified cereals and grains held the vast majority of sales quotas by volume, with 46% and about 44% respectively [40].
The occurrence of AA in commercially available food is a matter of interest for public health due to continuous and prolonged exposure to a probable carcinogenic substance. As the ingestion of contaminated food is the primary route of potential human exposure to AA, the aim of this research was to determine AA levels in cereal-based food, in particular biscuits and breakfast cereals, considering the widespread use of these products for all ages.

## Materials and methods

### Samples

A total of 56 samples of biscuits and breakfast cereals of the most consumed brands in Italy (17 and 5 respectively) were randomly collected in the two major supermarkets in Florence, as indicated by the European Commission Recommendation n° 307 [38], and tested for AA. The samples collected included 12 breakfast cereals and 44 biscuits whose composition consisted of maize, rice, malt, oatmeal, wheat, spelt, barley, rye and cocoa in different combinations.

### AA determination

AA determination was carried out following the method of Wang et al. [41] except for some modifications both in the extraction procedure and in the instrumental working conditions [42]. The HPLC method adopted is fast and cost-effective and it can be applied to screen the occurrence of AA in food samples ensuring a good degree of reliability.

### Instruments and reagents

All organic solvents, methanol and acetonitrile, were from J.T. Baker (Deventer, Holland) HPLC-grade. Water was purified by a Milli-Q-RO4 water purification system (Millipore, Bedford, MA, USA) with a resistivity of 10 MΩ·cm. The standard AA was purchased from Fluka (Deisenhofen, Germany). Potassium hexacyanoferrate (II) trihydrate and zinc sulfate heptahydrate were obtained from Merck (Darmstad, Germany). Oasis HLB 200 mg/6 mL SPE cartridges were obtained from Waters (Milford, MA, USA). Bond Elut-Accucat (200 mg/3 mL) SPE cartridges were purchased from Varian (Chi-

cago, IL, USA) and used with a SPE vacuum manifold (Visiprep, Supelco, Bellefonte, PA, USA). Amber glass auto-sampler vials with septum screw caps were obtained from Agilent Technologies (Wilmington, DE, USA). The analytical column (Atlantis dC18, 250 × 4.6 mm/5) with pre-column and syringe filters (0.45 µm PVDF) were from Waters (Milford, MA, USA).
A standard AA stock solution (1.0 mg/mL) was prepared by dissolving 1.0 mg of AA in 1.0 mL of Milli-Q water. To weigh AA, an electronic analytical balance (Radwag, PCE group, 0.01 mg Radom, Poland) was used. The AA stock solution was diluted in amber glass volumetric flasks to prepare calibration standards at 50, 100, 200, 500, 800 and 1000 ng/mL respectively and stored at 4°C until use.

### Sample preparation

100 g of each sample was finely pulverized, homogenized and dehydrated at 103°C for 4 hours, prior to sampling. Two 1.0 g solid samples were accurately weighed into centrifuge tubes.
A volume of 10 µl of the internal standard solution, AA at 100 µg/mL, was added to one centrifuge tube for recovery study. AA was extracted by adding 8 mL of water. The samples were shaken for 3 min on a Vortex and centrifuged (15 min at 3100 x g). The aqueous phase was collected and the extraction process was repeated twice, each time adding 2 mL of water, shaking for 2 min, and centrifuging for 15 min at 3100 x g. Aliquots of 1 ml of a 0.355 mol/l potassium hexacyanoferrate (II) trihydrate solution (Carrez I) and 1 ml of a 1.04 mol/l zinc sulfate heptahydrate solution (Carrez II) were added to the supernatant for protein precipitation. The samples were shaken for 2 min, kept cool at 4°C for 30 min, then centrifuged (15 min at 3100 x g). The samples were filtered through a 0.45 µm Millex syringe filter. Oasis HLB SPE cartridges and Bond Elut-Accucat SPE cartridges, conditioned with methanol and washed by water, were used in sequence to remove the first eluted fraction containing interfering substances. The final eluted fractions were concentrated to 1 mL by evaporation under vacuum at 45°C kept in amber glass vials at 4°C before injection.

### HPLC-UV analysis of AA

HPLC-UV analysis was performed with a HPLC instrument equipped with a vacuum degasser, a binary pump, an autosampler, and a DAD detector (Waters 1525 binary pump /Waters 996 PDA detector), at room temperature. The injection volume of calibration standards and sample extracts was 20 µl; AA was detected at two UV wavelengths of 200.5 and 210 nm. An isocratic elution was performed at a flow rate of 0.5 ml/min; the mobile phase consisted of water/acetonitrile 96:4 (v/v).
The retention time of AA was approximately 10.1 min, and the total run time was 35 min.
Under these chromatographic conditions, AA and the food components in the samples were all separated and eluted. AA content is resulting from an average of three measurements and the relative standard deviation of each sample was below 10%.
The AA recovery was on average 87%.
Table I reports the analytical parameters calculated in the range of 0.05-2 µg/g. The Limit of Detection (LOD) (calculated using the criterion signal to noise, s/n, of 3:1) and Limit of Quantification (LOQ) (calculated using the criterion s/n of 7:1). The intra-day and inter-day measurements show a good repeatability and reproducibility of the analytical methodology.

## Results and discussion

The results of the survey are shown in Tables II and III. According to Commission Recommendation n° 307 (2010) the main ingredients are reported for each sample. Our results are comparable to those of other studies [43-45]. Of 56 samples tested, only five had AA concentrations below the detection limit (10 µg/kg). The percentage of positive samples (≥ LOD) we detected in both kinds of food analysed was higher in biscuits (95.5%) than in breakfast cereals (75%) with a wide range of contamination and a considerable variability, as indicated from the high levels of standard deviation (294.1 and 287.75 respectively), also within each brand especially in biscuits (Tab. II). In detail, the standard deviation of the samples of brand 6 was 297.85 (AA levels ranging from < LOD to 820 µg/kg) while that of brand 17 was 298.57 (AA levels ranging from 30 µg/kg to 700 µg/kg). The biscuits of three brands (number 1, 9 and 16) showed the lower values, ranging from 30 to 50 µg/kg, from < LOD to 30 µg/kg and from 30 to 40 µg/kg respectively.
In the biscuit category, compared to that of breakfast cereals, we noted also the highest value, in spite of the fact that the average value and the median are slightly lower. The range of contamination of biscuits was from 30 µg/kg to 940 µg/kg. Only ten samples (22.7%) of biscuits showed a concentration above 500 µg/kg, set as the indicative value recommended [37]. The maximum level of 940 µg/kg was found in the sample of brand 14 (Tab. II), probably due to the industrial process conditions and to the presence of many different ingredients, such as

**Tab. I.** Analytical data of the acrylamide standard solution analysed by HPLC-UV: linear regression equation (peak area/conc.) and correlation coefficient ($R^2$) calculated in the linearity range 0.05-2 µg/g at wavelengths 210 and 200.5 nm; limit of detection (LOD, s/n = 3/1) expressed as µg/kg; limit of quantification (LOQ, s/n = 7/1) expressed as µg/kg; intra-day RSD and inter-day RSD (n = 3).

| Linear regression equation | $R^2$ | LOD (µg/kg) | LOQ (µg/kg) | Intra -day [a] RSD | Inter-day [b] RSD | Wavelength |
|---|---|---|---|---|---|---|
| y = 116616 x -4415. 8 | 0.998 | 10 | 28 | 3-6.6 | 6.8 | 210 nm |
| y = 250771 x -2640. 9 | 0.999 | 10 | 28 | 4.2-7.2 | 1.9 | 200.5 nm |

[a] Values are referred to six independent samples analyzed in one day
[b] Values are referred to six independent samples analyzed three times in three different days

**Tab. II.** Acrylamide (AA) content (µg/kg) in biscuits.

| Biscuits (N. samples) | Main ingredients [a] | AA µg/kg |
|---|---|---|
| Brand 1 (2) | Grain | 50 |
| | Wheat /barley | 30 |
| Brand 2 (4) | Wheat/rice/barley | 66 |
| | Wheat/oat | 200 |
| | Wheat/cocoa | 200 |
| | Wheat/maize | 30 |
| Brand 3 (1) | Wheat /cocoa | 310 |
| Brand 4 (3) | Wheat/maize/rice/oat | 560 |
| | Grain/maize/oat | 840 |
| | Grain | 50 |
| Brand 5 (1) | Wheat /cocoa | 400 |
| Brand 6 (7) | Wheat | < LOD |
| | Wheat/oat | 300 |
| | Whole grain | 90 |
| | Wheat | 80 |
| | Hazelnuts/ cocoa | 590 |
| | Wheat /cocoa | 200 |
| | Wheat /cocoa | 820 |
| Brand 7 (1) | Wheat | 260 |
| Brand 8 (1) | Whole grain | 450 |
| Brand 9 (2) | Wheat/maize | < LOD |
| | Wheat | 30 |
| Brand 10 (2) | Whole grain/oat | 430 |
| | Wheat | 280 |
| Brand 11 (2) | Grain | 840 |
| | Whole grain | 720 |
| Brand 12 (2) | Grain /cocoa | 840 |
| | Wheat | 50 |
| Brand 13 (3) | Maize/barley | 280 |
| | Wheat | 190 |
| | Wheat /cocoa | 90 |
| Brand 14 (1) | Wheat/oat/barley/ Rice/maize/rye | 940 |
| Brand 15 (2) | Wheat | 100 |
| | Wheat | 180 |
| Brand 16 (3) | Wheat /cocoa | 140 |
| | Wheat/rice | 30 |
| | Wheat | 40 |
| Brand 17 (7) | Wheat | 40 |
| | Grain | 60 |
| | Wheat /rice | 30 |
| | Grain/barley/rice/ Oats/rye | 530 |
| | Wheat /cocoa | 240 |
| | Wheat/cocoa | 490 |
| | Wheat /cocoa | 700 |
| mean [b] | 297.18 | |
| median [b] | 200 | |
| LB - UB | < LOD - 940 | |

[a] The first is the ingredient present in the largest quantity
[b] Results lower than LOD were assigned ½ LOD
LB: lower bound
UB: upper bound

**Tab. III.** Acrylamide (AA) content (µg/kg) in breakfast cereals.

| Breakfast cereals (N. samples) | Main ingredients [a] | AA µg/kg |
|---|---|---|
| Brand 1 (1) | Maize/rice | 360 |
| Brand 2 (3) | Oat/wheat/maize/rice | < LOD |
| | Rice/wheat | < LOD |
| | Wheat/maize/rice/cocoa | 220 |
| Brand 3 (5) | Wheat/rice/oat | 280 |
| | Wheat/cocoa | 60 |
| | Maize/barley | < LOD |
| | Maize/oat/rice/cocoa | 110 |
| | Grain/oat/rice/cocoa | 450 |
| Brand 4 (2) | Wheat/rice/cocoa | 810 |
| | Spelt | 780 |
| Brand 5 (1) | Wheat/cocoa | 80 |
| mean [b] | 263.75 | |
| median [b] | 165 | |
| LB-UB | < LOD - 810 | |

[a] The first is the ingredient present in the largest quantity
[b] Results lower than LOD were assigned ½ LOD
LB: lower bound
UB: upper bound

wheat, oat, barley, rice, maize, rye. Ingredients indeed play an important role in AA formation, as different cereals have different amounts of free asparagine and reducing sugars available for the Maillard reaction [23].

EC Recommendation n° 647 [37] amended the previous indicative values for breakfast cereals according to the composition: for bran products and whole grain cereals the indicative value is 400 µg/kg, for wheat and rye based products it is 300 µg/kg, while for maize, oat, spelt, barley and rice based products the value is 200 µg/kg.

In the category of breakfast cereals tested, the AA occurrence, ranging from 60 to 810 µg/kg, is less frequent (75%) than in biscuits, although the percentage of samples with AA concentration above the indicative values recommended (33%) was higher than that stated in the EFSA report [28]. The mean value obtained was 263.75 µg/kg and the median value was 165 µg/kg; both these values were lower than those of the biscuits (297.18 and 200 µg/kg respectively) and in agreement with data recorded in Europe and other countries [7, 21, 28].

The highest values were found in two samples of brand 4, one consisting of spelt (780 µg/kg) and the other consisting of wheat, rice and cocoa (810 µg/kg): this also supports the hypothesis that the industrial process of preparation is the determinant key factor.

Of 56 samples tested, 18 (32.1%) contained cocoa: 12 (27.3%) biscuits and six (50%) breakfast cereals. None of these had AA concentrations below the detection limit, and they showed a wide range of contamination, from 60 µg/kg in a breakfast cereal to 840 µg/kg in a biscuit sample. The average value of all samples containing cocoa was 375 µg/kg and the median 275 µg/kg. In comparison to the rates related to the AA contamination in the samples without cocoa (mean = 242.6 µg/kg, median = 95 µg/kg), we suppose that the presence of cocoa

probably contributes to the increase of AA concentration. Some studies [46, 47] reported that AA concentration in cocoa is variable and can reach concentrations of about 600 µg/kg.

In the scientific report of EFSA [28] on AA levels in food from monitoring years 2007 to 2010, the average concentration of AA in breakfast cereals and biscuits ranged from 138 and 289 µg/kg to a maximum of 1290 and 5849 µg/kg respectively, and the median values are 99 and 91 µg/kg (data referring to 2010). We noted that our results show mean and median values higher than those reported by EFSA [28], while the maximum values are clearly lower.

Our results, compared to the AA indicative values recommended [37], showed that 22.7% of biscuits and 33% of breakfast cereals exceeded the indicative values set at 500 µg/kg for biscuits and from 200 to 400 µg/kg for breakfast cereals, depending on the composition. Our findings are interesting to evaluate the impact of AA on dietary intake as the overall consumption of breakfast cereals by Italian consumers is 14.8 g/d, while the overall consumption of biscuits is more relevant because it is rated at 27.3 g/d. For food consumption data Authors referred to a survey carried out by the Italian Institute for Nutrition (INRAN) [48] that is highly representative of Italian population for age, gender, geographical area and the wide number of subjects involved.

## Conclusions

In conclusion our data, not unlike that from the most recent research carried out in Italy and other countries, showed widespread contamination by AA with high frequency of positive samples (91% of all samples tested ≥ LOD) and considerable variability in concentration values, most likely due to industrial processing and the presence of various ingredients.

The good news is that 41.1% of samples tested, 40.9% biscuits and 41.7% breakfast cereals, showed an AA concentration below 100 µg/kg, which can be considered a low value, even if low concentrations may not imply zero risk, especially in view of neoplastic effects based on animal evidence.

Since cereal based products are contributing substantially to human exposure, reducing AA levels in foods should be a high priority to decrease the risk for human health.

### References


[1] Friedman M. *Chemistry, biochemistry, and safety of acrylamide: A review.* J Agr Food Chem 2003;51:4504-26.

[2] Stadler RH, Blank I, Varga N, et al. *Acrylamide from Maillard reaction products.* Nature 2002;419:449-50.

[3] Tareke E, Rydberg P, Karlsson P, et al. *Analysis of acrylamide, a carcinogen formed in heated foodstuffs.* J Agr Food Chem 2002;50:4998-5006.

[4] International Agency for Research on Cancer [IARC]. *Acrylamide.* In: *IARC monographs on the evaluation of carcinogen risk to humans.* Some industrial chemicals 1994;60:389- 433. Lyon: France. Available at: http://monographs.iarc.fr/ENG/Monographs/vol60/mono60.pdf [last accessed: 15/12/2014]

[5] Agency for Toxic Substances and Disease Registry [ATSDR]. *Toxicological profile of acrylamide.* U.S. Department of Health and Human Services. Public Health Service. December 2012. Available at: http://www.atsdr.cdc.gov/toxprofiles/tp203.pdf [last accessed: 09/01/2015].

[6] Friedman MA, Dulak LH, Stedham MA. *A lifetime oncogenicity study in rats with acrylamide.* Fund Appl Toxicol 1995;27:95-105.

[7] Joint Expert Committee on Food Additives [JECFA]. *Acrylamide.* In: *Evaluation of Certain contaminants in food, seventy-Second report of the Joint FAO/WHO Expert Committee on Food Additives*. WHO Technical Report Series 2011;959:9-21. Geneva, Switzerland: Available at: http://whqlibdoc.who.int/trs/who_trs_959_eng.pdf [last accessed: 03/02/2015].

[8] Johnson KA, Gorzinski SJ, Bodner KM, et al. *Chronic toxicity and oncogenicity study on acrylamide incorporated in the drinking water of Fisher 344 rats.* Toxicol Appl Pharm 1986;85:154-68.

[9] Raoufi A, Jamshidi K. *Acrylamide induced chronic nephrotoxicity in Rats.* Toxicol Lett 2013;221S:S59-S256.

[10] Baum M, Fauth E, Fritzen S, et al. *Acrylamide and glycidamide: genotoxic effects in V79-cells and human blood.* Mut Res 2005;580:61-9.

[11] Doerge DR, Gamboa da Costa G, McDaniel LP, et al. *DNA adducts derived from administration of acrylamide and glycidamide to mice and rats.* Mut Res 2005; 580:131-41.

[12] Gamboa da Costa G, Churchwell MI, Hamilton LP, et al. *DNA adduct formation from acrylamide via conversion to glycidamide in adult and neonatal mice.* Chem Res Toxicol 2003;16:1328-37.

[13] Beland FA, Mellick PW, Olson GR, et al. *Carcinogenicity of acrylamide in B6C3F1 mice and F344/N rats from a 2-year drinking water exposure.* Food Chem Toxicol 2013;51:149-59.

[14] Blasiak J, Gloc E, Wozniak K, et al. *Genotoxicity of acrylamide in human lymphocytes.* Chem-Biol Interact 2004;149:137-49.

[15] Ghanayem BI, Witt KL, El-Hadri L, et al. *Comparison of germ cell mutagenicity in male CYP2E1-null and wildtype mice treated with acrylamide: evidence supporting a glycidamide mediated effect.* Biol Reprod 2005;72:157-63.

[16] Riboldi PB, Vinhas AM, Moreira JD. *Risks of dietary acrylamide exposure: A systematic review.* Food Chem 2014;157:310-22.

[17] World Health Organization [WHO]. *Summary report of the sixty-fourth meeting of the Joint FAO/WHO Expert Committee on Food Additives (JECFA*). Rome, 8-17 February 2005;7-17. Available at: ftp://ftp.fao.org/es/esn/jecfa/jecfa64_summary.pdf [last accessed: 10/12/2014].

[18] Mottram DS, Wedzicha BL, Dodson AT. *Acrylamide is formed in the Maillard reaction.* Nature 2002;419:448-9.

[19] Zyzak DV, Sanders RA, Stojanovic M, et al. *Acrylamide formation mechanism in heated foods.* J Agr Food Chem 2003;51:4782-7.

[20] Keramat J, Le Bail A, Prost C, et al. *Acrylamide in foods: Chemistry and analysis. A review.* Food Bioprocess Tech 2011;4:340-63.

[21] Joint Expert Committee on Food Additives [JECFA]. *Acrylamide.* In: *Evaluation of certain food contaminants, sixty-fourth report of the Joint FAO/WHO Expert Committee on Food Additives.* WHO Technical Report 2006;930:8-26. Geneva, Switzerland. Available at: http://whqlibdoc.who.int/trs/WHO_TRS_930_eng.pdf [last accessed: 09/01/2015].

[22] Annola K, Keski-Rahkonen P, Vähäkangas K, et al. *Simultaneous determination of acrylamide, its metabolite glycidamide and antipyrine in human placental perfusion fluid and placental tissue by liquid chromatography-electrospray tandem mass spectrometry.* J Chromatograp B 2008;876:191-7.

[23] Capuano E, Fogliano V. *Acrylamide and 5-hydroxymethylfurfural (HMF): A review on metabolism, toxicity, occurrence in food and mitigation strategies.* LWT-Food Sci Technol 2011;44:793-810.

[24] Arribas-Lorenzo G, Morales FJ. *Recent insights in Acrylamide as carcinogen in foodstuffs.* Adv Mol Toxicol 2012;6:163-93.

[25] Bjellaas T, Stølen LH, Haugen M, et al. *Urinary acrylamide metabolites as biomarkers for short-term dietary exposure to acrylamide.* Food Chem Toxicol 2007;45:1020-6.

[26] Boettcher MI, Bolt HM, Drexler H, et al. *Excretion of mercapturic acids of acrylamide and glycidamide in human urine after single oral administration of deuterium-labelled acrylamide.* Arch Toxicol 2006;80:55-61.

[27] European Commission [EC]. *Recommendation of 3 May 2007 on the monitoring of acrylamide levels in food.* 2007/331/EC. OJ L 123/33, 12.5.2007. Available at: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32007H0331&from=EN [last accessed: 10/03/2015].

[28] European Food Safety Authority [EFSA]. *Update on acrylamide levels in food from monitoring years 2007 to 2010.* EFSA Journal 2012;10:2938. doi:10.2903/j.efsa.2012.2938. Available at: http://www.efsa.europa.eu/it/efsajournal [last accessed: 20/01/2015].

[29] European Food Safety Authority. EFSA CONTAM Panel (EFSA Panel on Contaminants in the Food Chain), 2015. *Scientific Opinion on acrylamide in food.* EFSA Journal 2015;13:4104, 321. doi:10.2903/j.efsa.2015.4104. Available at: www.efsa.europa.eu/efsajournal [last accessed: 05/06/2015].

[30 Freisling H, Moskal A, Ferrari P, et al. *Dietary acrylamide intake of adults in the European Prospective Investigation into Cancer and Nutrition differs greatly according to geographical region.* Eur J Nutr 2013;52:1369-80.

[31] Carrieri G, Anese M, Quarta B, et al. *Evaluation of acrylamide formation in potatoes during deep-frying: The effect of operation and configuration.* J Food Eng 2010;98:141-9.

[32] Marchettini N, Focardi S, Guarnieri M, et al. *Determination of acrylamide in local and commercial cultivar of potatoes from biological farm.* Food Chem 2013;136:1426-8.

[33] Tateo F, Bononi M, Gallone F. *Acrylamide content in potato chips on the Italian market determined by liquid chromatography tandem mass spectrometry.* Int J Food Sci Tech 2010;45:629-34.

[34] Banchero M, Pellegrino G, Manna L. *Supercritical fluid extraction as a potential mitigation strategy for the reduction of acrylamide level in coffee.* J Food Eng 2013;115:292-7.

[35] World Health Organization [WHO]. *Guidelines for Drinking-Water Quality, incorporating 1st and 2nd addenda,* Vol. 1, Recommendations. 3rd ed. 2008; Geneva, Switzerland. Available at: http://www.who.int/water_sanitation_health/dwq/fulltext.pdf?ua=1 [last accessed: 12/01/2015].

[36] European Commission [EC]. *Recommendation of 10 January 2011 on investigations into the levels of acrylamide in food,* C (2010) 9681 final 2011. Available at: http://ec.europa.eu/food/food/chemicalsafety/contaminants/recommendation_10012011_acrylamide_food_en.pdf [last accessed: 19/01/2015].

[37] European Commission [EC]. *Recommendation of 8 November 2013 on investigations into the levels of acrylamide in food.* 2013/647/EU. OJ L 301/15, 12.11.2013. Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:301:FULL:EN:PDF [last accessed: 05/12/2014].

[38] European Commission [EC]. *Commission recommendations of 2 June 2010 on the monitoring of acrylamide levels in food.* 2010/307/EU. Official Journal of the European Union, L137, 3.6.2010; 4-10. Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:137:0004:0010:EN:PDF [last accessed: 0512/2014].

[39] Istituto di Servizi per il Mercato Agricolo Alimentare [ISMEA]. *Food consumptions – Italian household consumptions* [in Italian]. No. 1/13. 2013. Available at: http://www.ismea.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/8499 [last accessed: 12/01/2015].

[40] Varacca A, Sckokai P. *Product differentiation and competition between brands in market for breakfast cereals* [in Italian]. Agriregionieuropa 2012;8:31. Available at: http://hdl.handle.net/10807/39180 [last accessed: 12/01/2015].

[41] Wang H, Feng F, Guo Y, et al. *HPLC-UV quantitative analysis of acrylamide in baked and deep-fried Chinese foods.* J Food Compos Anal 2013;31:7-11.

[42] Gökmen V, Morales FJ, Ataç B, et al. *Multiple-stage extraction strategy for the determination of acrylamide in foods.* J Food Compos Anal 2009;22:142-7.

[43] Anese M, Quarta B, Peloux L, et al. *Effect of formulation on the capacity of l-asparaginase to minimize acrylamide formation in short dough biscuits.* Food Res Int 2011;44:2837-42.

[44] Russo MV, Avino P, Centola A, et al. *Rapid and simple determination of acrylamide in conventional cereal-based foods and potato chips through conversion to 3-[bis(trifluoroethanoyl) amino]-3-oxopropyl trifluoroacetate by gas chromatography-coupled with electron capture and ion trap mass spectrometry detectors.* Food Chem 2014;146:204-11.

[45] Yusa V, Quintas G, Pardo O, et al. *Determination of acrylamide in foods by pressurized fluid extraction and liquid chromatography-tandem mass spectrometry used for a survey of Spanish cereal-based foods.* Food Addit Contam 2006;23:237-44.

[46] Ren Y, Zhang Y, Jiao J, et al. *Sensitive isotope dilution liquid chromatography/electrospray ionization tandem mass spectrometry method for the determination of acrylamide in chocolate.* Food Addit Contam 2006;23:228-36.

[47] Sirot V, Hommetb F, Tard A, et al. *Dietary acrylamide exposure of the French population: Results of the second French Total Diet Study.* Food Chem Toxicol 2012; 50:889-94.

[48] Leclercq C, Arcella D, Piccinelli R, et al. on behalf of the INRAN-SCAI 2005-06 Study Group. *The Italian food consumption survey INRAN-SCAI 2005–06: main results in terms of food consumption*. Public Health Nutr 2009;12:2504-32. Available at: http://nut.entecra.it/710/I_consumi_alimentari__INRAN-SCAI_2005-06.html. [last accessed: 22/06/2015].

■ Received on June 9, 2015. Accepted on November 9, 2015.

■ Correspondence: Raffaella Capei, Department of Health Sciences, University of Florence, viale Morgagni 48, 50134 Florence, Italy - Tel. +39 055 2751078 - E-mail: raffaella.capei@unifi.it

ORIGINAL ARTICLE

# Prevalence of acute alcohol intoxication in Borgo Trento Hospital Emergency Department (Verona)

S. MAJORI[1], G. RICCI[2], F. MARCHIORI[3], M. BOCCHI[3], M. ZANNONI[2]
[1] Department of Public Health and Community Medicine, Hygiene and Environmental Occupational Preventive Medicine Division, University of Verona, Italy; [2] Accident and Emergency Department of "Borgo Trento" Hospital, Azienda Ospedaliera Universitaria Integrata Verona, Italy; [3] Specialisation School in Hygiene and Preventive Medicine, University of Verona, Italy

Key words

Alcohol • Acute intoxication • Emergency department

Summary

**Introduction**. *Alcohol is an important factor weighting towards global disease burden, premature death and Disability-Adjusted Life Years. This study examines the burden imposed on the Borgo Trento Verona Hospital Emergency Department (ED) Italy by patients with Acute Alcohol Intoxication (AAI).*
**Methods**. *A 6-year retrospective study was performed by reviewing medical records in all patients (≥ 16 years old) diagnosed with AAI. Clinical criteria for inclusion in the sample followed those defined in the "Alcohol Intoxication Symptoms" section of DSM-IV-TR. Ambiguous cases presenting traumas potentially related to AAI were confirmed positive using a Blood Alcohol Level (BAL) test before inclusion in the sample. Socio-demographic data, case history, timing of admission/discharge and outcome were collected for each patient.*
**Results**. *One thousand five hundred forty-seven patients (males:females = 6.6:1) were included in the study. With regard to marital status, the crude rate within the sample shows that single and married subjects have a significantly higher prevalence ($p < 0.01$). When demographic data of residents was taken into consideration, divorced/separated and single categories were significantly more represented ($p < 0.05$).*
*The admissions appeared fairly constant throughout the week with a higher prevalence of patients aged between 25-55 years, and a significant peak of the youngest age class on weekends. Higher rates of admission were recorded during the late afternoon and night.*
*The triage code of admission appeared uniformly distributed along the week, with the highest prevalence of green code (67.7%), followed by yellow one (25.8%). The analysis of clinical symptoms and BAL highlights that moderate clinical severity concerns almost 50%, and clinically critical severity the 6.6% of cases of access.*
*Admissions rates of foreign patients resulted to be 3 times higher compared to those of natives when adjusted to demographic data.*
**Discussion**. *Social disadvantages such as foreign provenance and social difficulties as unemployment, low money intake, perception of loneliness, and dysfunction in family life were frequently identified in the sample. An exception was found within the youngest age group, where the AAI are mostly related to alcohol abuse during social outings over the weekend.*
*EDs play a crucial role in helping patients with AAI, but prevention of alcohol abuse though implementation of social and sanitary health policies on all ages (but especially among the youngest) is essential.*

## Introduction

Alcohol is the third risk factor for premature death and disability in the world: the WHO estimates that inappropriate use of alcohol causes approximatively 3.3 million deaths per year (about 6% of total world deaths) and over 5% of global disease burden (mainly cancer, vascular, and liver disease; traffic accidents, suicide, and violence), and represents an important cause of Disability-Adjusted Life Years (DALYs) [1, 2].
Europe appears to be a region in the world with recurrent heavy drinking, and stands out as the area with the most significant increase in patterns of alcohol consumption with a corresponding increase of cases of Acute Alcohol Intoxication (AAI) [3].
As a result, in this region alcohol consumption is an important and a growing public health problem: every year 195,000 people die for harmful use of alcohol (12% of all male and 2% of all female premature mortality cases), which represent about the 3.8% of all causes of death and the 4.6% of DALYs [4-6].
In Italy in 2010 were recorded 16,829 deaths (age > 15 years; rate M:F = 2.27) [7, 8] and the prevalence of at-risk alcohol consumers older than 11 years is estimated to be around 8,000,000 subjects (rate M:F = 1:3.27) [9].
The incidence of Acute Intoxications (AI) represents the 10.0/1000 of the total admittance to Emergency Departments (EDs) and among them about 57% were alcohol related, representing between the 0.6 and 40% of all the ED accesses as referred in various reports. [10-15]
As the ED play a crucial role for patients affected by acute Alcohol Use Disorders (AUDs), a study was performed to evaluate the burden imposed by patients with AAI to the Emergency Department of Borgo Trento Hospital in Verona (Italy).

## Methods

This retrospective study involved a medical record review spanning a 6-year period (1 January 2009 - 31 December 2014) of all consecutive patients admitted to the ED due to alcohol intoxication. Admissions averaged approximately 98,800 patients per year.
The Italian "privacy" law and the Helsinki Declaration on human subjects testing were fully respected and the protocols were approved by the local ethical committee.
All the patients included in the study were over 15 years old (patients under 16 years of age are served by a separate pediatric ED), and had a clinical criteria for inclusion following diagnosis of AAI as defined by the DSM-IV-TR "Diagnostic Criteria for Alcohol Abuse and Dependence" [16, 17] which include "a) recent ingestion of alcohol; b) clinically significant maladaptive behavioral or psychological changes (e.g., inappropriate sexual or aggressive behavior, mood liability, impaired judgment, impaired social or occupational functioning) that developed during, or shortly after, alcohol ingestion; c) one (or more) of the following signs, developing during, or shortly after, alcohol use: slurred speech, incoordination, unsteady gait, nystagmus, impairment in attention or memory, stupor or coma."
All the samples were taken and coded with anonymous matching bar codes, non-specific to the patient, and all the information necessary for the study was associated.
To determine the presence and the level of alcohol intoxication of each patient a blood sample was obtained and analyzed at the central laboratory of the same hospital of the Emergency Department (ED) using a commercial enzymatic "alcohol dehydrogenate" method with the UV spectrophotometer as a blood alcohol level (BAL) test. The minimum measurable amount was of 0.1 g/L.
The severity of AAI in patients was classified by the ED into four categories considering both the severity of the clinical symptoms and the BAL: clinical symptoms (c.s.) less than slight and BAL < 17 mmol/L; c.s. slight and BAL between 17 and 53 mmol/L; c.s. moderate and BAL between 54 and 87 mmol/L; c.s. critical and BAL > 87 mmol/L.
Patients admitted for other clinical reasons such as various kind of traumas, burns, and mental disorders, and which could initially have led to a suspicion of excessive alcohol consumption, were only included if BAL tests resulted positive.
Day span was defined as beginning at 00:01 and ending at 24:00 h. The weekend started at 00:01 h on Saturday and finished at 24:00 h on Sunday.
For each admitted patient the following data were extracted from the database of the ED: identification code, sociodemographic data (sex, date of birth, age, nationality, place of residence, marital status), timing of admission/discharge (day and hour of admittance, day and time of discharge, period of stay), and case history (triage category, anamnestic data, physical examination, BAL and severity of clinical symptoms, and diagnosis). The outcome was divided into "admission to a hospital ward", "sent to a general practitioner" (i.e. going to a family doctor for further treatment), "discharge" (i.e. returning back home), and "refusal of treatment" (i.e. decision of the patient not be treated).
Collected data were registered within a Microsoft Office Excel database.

### Statistical Analysis

In order to report the results of the analysis, categorical variables were presented as numbers and percentages; the comparison between subgroups was carried out using Chi-square test for qualitative data. Linear trend was evaluated with the Cochran-Armitage test for linear trend.
P values were considered significant when $< 0.05$.

## Results

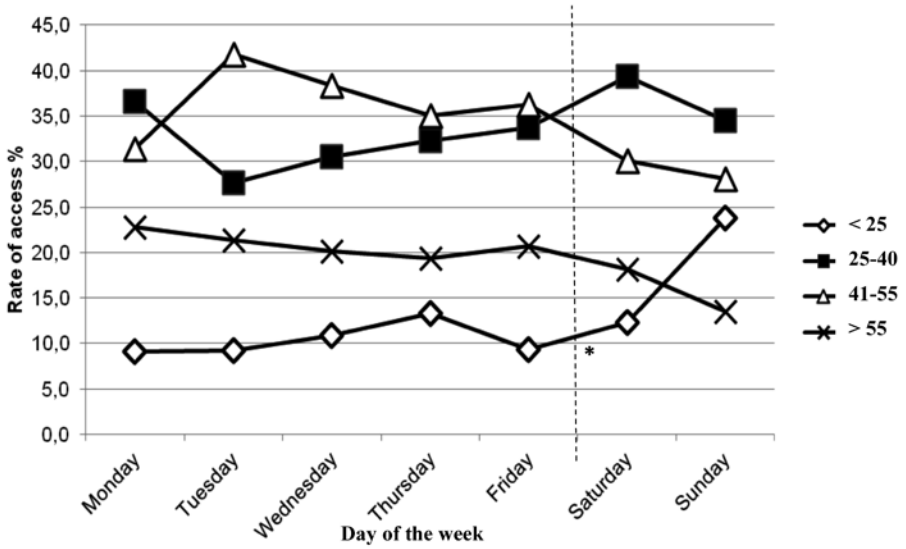
The study included 1,547 patients, representing 0.26% of all ED visits occurred during the duration of the study. Males represented 85.2% of subjects, with a male-to-female ratio of 6.6:1. The age range was 16-92 years for males, 16-90 years for females, with a mean age of 42.6 and of 40.8 years respectively.
The crude distribution rate of marital status (Tab. I) highlights that single and married subjects show a significantly higher prevalence ($p < 0.001$). When demographic data of residents was taken into consideration, divorced/separated and single categories were significantly more represented ($p < 0.05$).
Admissions appear fairly constant along the week (Fig. 1), with a predominance of patients aged between 25-55 years. On Saturday and Sunday admissions were respectively of 16.9% and 18.4% of the sample total, with exception of the youngest age class (< 25 years), which showed the largest and most significant increase (more than the 52%) during weekend compared to the other age classes ($p < 0.001$).
Admissions to the ED resulted mostly clustered (Fig. 2) around the late afternoon with a significantly ($p < 0.001$) increasing trend during the night (between hours 20 and 2). A remarkable reduction of patients was registered in the morning.
The triage code distribution evaluated along the week (Fig. 3) shows a uniform and not significantly trend. Average Green Code pertained to 67.7% of the patients; Yellow to 25.8%; White to 4.6% and Red to 1.9%.
The length of stay in the ED under medical observation was different according to the urgency level: < 1 hour for patients for White Codes, 1-2 hours with Green Code, 3-4 hours for Yellow Codes, and 5-7 hours for Red Codes.
The distribution of severity of clinical symptoms and BAL is reported in Tab. II. Notably, moderate severity was found in almost 50% of patients (prevalent age: 21-50 years), and 6.6% resulted clinically critical (prevalent age: 31-50 years). Moderate clinical symptoms resulted significantly prevalent ($p < 0.001$) compared to other classifications of clinical classes.

**Tab. I.** Marital status: distribution in admitted patients to ED according to crude rate and to the respective distribution in resident population (Italian plus foreigner).

| | Admitted patients | |
|---|---|---|
| **Marital status** | **Crude rate (%)** | **Rate according to marital status of resident population (%)** |
| Single | 38.1* | 0.5 & |
| Married | 22.2*, ** | 0.3 &, && |
| Divorced/separated | 5.7** | 0.6 && |
| Widow | 0.4 | 0.1 |
| Not declared | 33.6 | 0 |

* and **: $p < 0.01$; & and &&: $p < 0.05$

**Fig. 1.** Weekly trend of admissions because of alcohol intoxication to ED according to age class and day of the week.




**Age < 25** vs other age classes during WE = $p < 0.01$

**Fig. 2.** Distribution rate (%) of admissions to Emergency Department according to daytime.

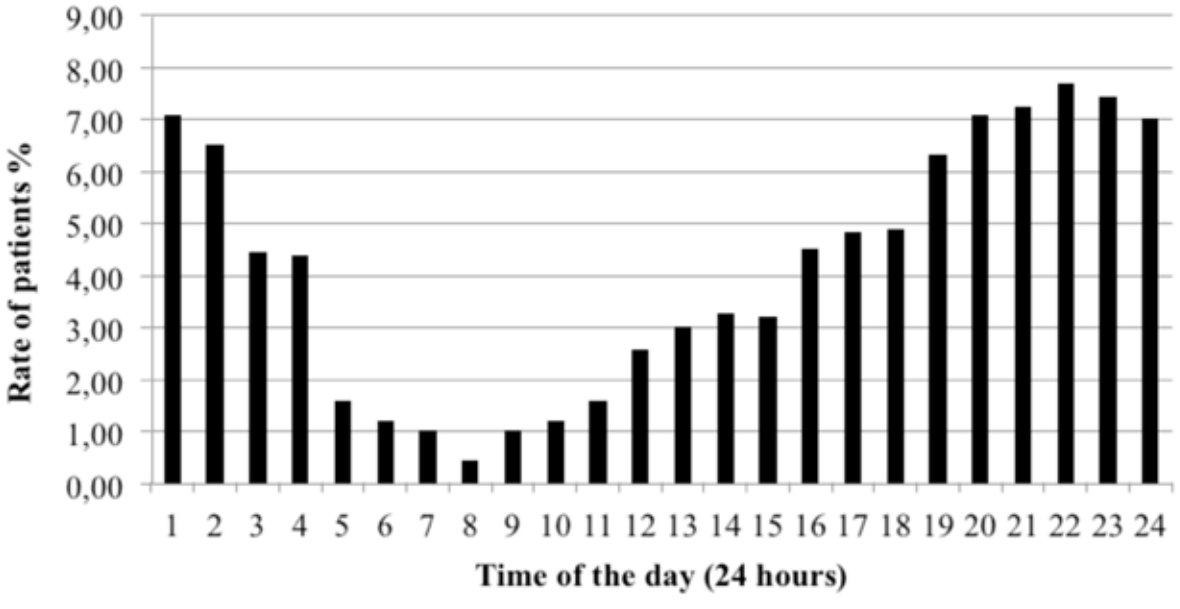
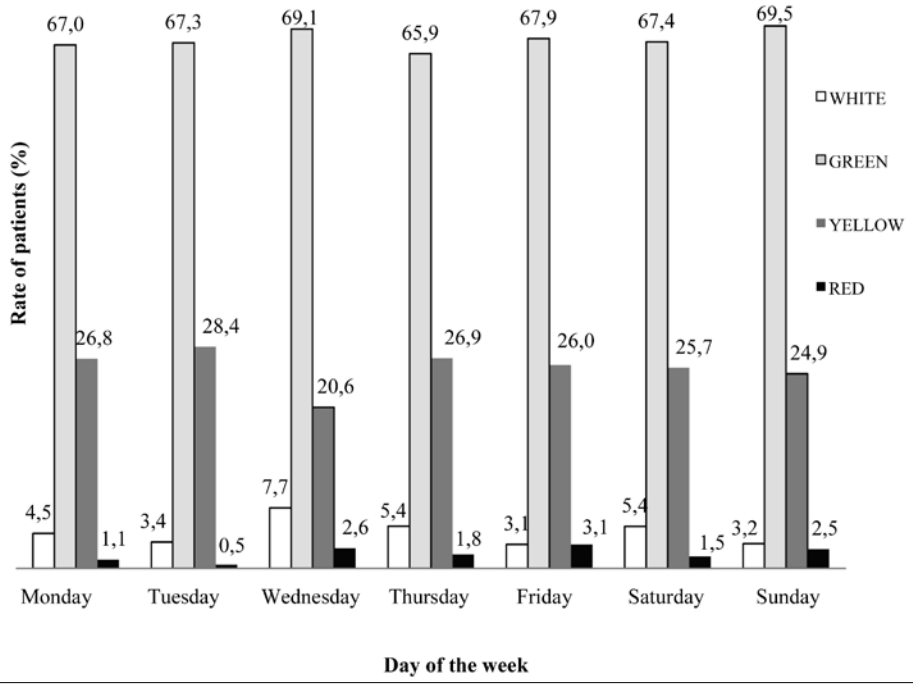


**Fig. 3.** Distribution rate (%) of admissions according to day of attendance and triage.




The majority of admissions are represented by subjects aged between 21-50 (p: n.s.) with a significantly predominance among ages 41-50 (27.9% of cases) ($p < 0.05$). Analysis based on clinical symptoms and BAL shows that distribution of admissions by age is different according to clinical condition: younger subjects generally had a non-relevant or a moderate AAI, while patients aged 31-50 resulted experiencing prevalently and significantly moderate ($p < 0.05$) or critical ($p < 0.05$) BAL levels.

The majority of patients (crude rate: 68.6%) admitted to the ED were native Italian subjects.

However, when the analysis of these data was evaluated proportionally to demographic information (local municipal data), the admission rate of foreign patients results in $123.2/10^5$ inhabitants, which is almost three times greater than rate for the Italian patients ($46.4/10^5$ inhabitants).

Among foreign patients African and Asian subjects show the greatest admission rate, being collectively responsible for 32.7% of admissions, followed by Eastern European (26.1%) and Western European (3.9%) subjects.

The main distributions of foreign country provenance of the most prevalent patients (first ten countries) admitted to the ED and the proportion, according to nationality, relative to the respective resident foreign population are reported in Table III. The data highlight that the majority of admitted patients came from Morocco and Sri Lanka,

**Tab. II.** Distribution rate (%) of patients with AAI according to clinical symptoms, (BAL) and age class.

| | Severity of clinical symptoms and level of BAL in Acute Alcohol Intoxication (AAI) | | | | |
|---|---|---|---|---|---|
| Clinical Symptomps<br>BAL | Non Relevant<br>*(< 17 mmol/L)* | Slight<br>*(17-53 mmol/L)* | Moderate<br>*(54-87 mmol/L)* | Critical<br>*(> 87 mmol/L)* | Total<br>results |
| AGE CLASS | & | & | | | |
| < 20 | 1.5% | 4.7% | 2.1% *, ** | 0.3% | 8.6 % |
| 21-30 | 5.3% | 5.1% | 8.8% *, ** | 0.6% | 19.8% |
| 31-40 | 4.7% | 1.9% | 10.1% *, ** | 1.7% +, ++ | 18.4% |
| 41-50 | 9.2% | 2.6% | 13.3% *, ** | 2.8% +, ++ | 27.9% £ |
| 51-60 | 2.6% | 2.3% | 4.9% *, ** | 0.8% | 10.6 % |
| >60 | 3.2 % | 1.9% | 9.2% *, ** | 0.4% | 14.7% |
| TOTAL | 26.5 % | 18.5% | ***48.4 %*** °° | 6.6% | 100% |

&: Chi-square test among age classes (from < 20 to > 60) among non relevant and among slight clinical symptoms: p = n.s.
£: Total results: Age 41-50 vs all other age classes: $p < 0.05$ (Chi-square test);
* and **: Moderate clinical symptoms: evaluation of age classes 31-40 and of 41-50 and other age classes: $p < 0.05$ (Chi-square test)
+ and ++: Critical clinical symptoms: $p < 0.05$ (Chi-square test)
°° : Moderate clinical symptoms vs other clinical classes: $p < 0.001$ (Chi-square test)

Tab. III. Foreign patients: countries of provenance (first ten countries), crude rate of admission to ED and rate of foreign patients according to the respective resident foreign population in Verona.

| Country of origin | Crude rate of admission to ED for foreign patients according to nationality out of total admitted foreigners (%) | Rate of foreign patients according to nationality in relation to the respective resident foreign population in Verona (%) |
|---|---|---|
| Morocco | 20.9 | 4.8 |
| Sri Lanka | 17.0 | 1.1 |
| India | 14.5 | 11.0 |
| Romania | 13.6 | 0.7 |
| Poland | 4.1 | 4.9 |
| Moldavia | 3.6 | 0.5 |
| Ghana | 3.6 | 1.2 |
| Senegal | 2.3 | 2.8 |
| Algeria | 1.6 | 2.5 |
| Albania | 1.1 | 0.3 |

but that the demographically highest proportion came from India.

Slightly more than half of the patients (59.3%) left the ED before concluding of all checks and the final visits. Among the remaining patients, one in three was discharged and sent to his GP and 0.4% was sent directly to Community Psychiatric Services.

Only 4.4% of patients were transferred to a hospital ward: 39.7% was admitted to the ED ward, 35.3% to the Medical Ward, 17.6% to the Surgical Ward, and 7.34% to the Psychiatric ward.

## Discussion

By examining the population served by the Borgo Trento ED, this study sheds a different light onto the prevalence of AAI, which has otherwise only been appreciated within the contexts of mortality rate or hospitalization statistics.

During the six years studied 0.26% of all ED admittances were due to AAI. Among them the crude rate shows that almost 70% of admitted patients were native Italians. However, when the data were compared to the proportion of nationalities of Verona residents, it became apparent that this phenomenon concerns foreigners almost tree times more severely than native Italians ($123.2/10^5$ vs $46.4/10^5$, respectively).

Recent Italian data report harmful uses of alcohol in 6.3% of Italians [18] and a higher rate (13.2%) in foreigners [19], which may explain the greater prevalence of AAI among foreign people found in this report.

Foreigners show diversified behaviors of alcohol consumption depending on their country of provenance, all of which are styles that differ from the Italian traditional one [20, 21]. This fact could be related to a general and widespread greater social disadvantage in comparison with the indigenous population, to difficulties with integration (language, racism), to socio-economic difficulties (finding a job, food, housing, and obtaining documents), to a need to decrease anxiety, to increased availability of alcohol compared to countries of origin (due to legal, religious, or economic reasons). All of these reasons can facilitate alcohol use and abuse, leading to episodes of acute intoxication and/or to chronic alcoholism [22].

In comparison to the Epicentro Italian study [23], in which a greater prevalence of AAI was found among Romanian, Albanian, and Polish people, our data reports a higher prevalence, as crude rate, in patients coming from Morocco, Sri Lanka, India, and Romania.

From the analysis of age groups, a constant higher prevalence along the week emerged in patients aged 21 to 55. A significant peak was observed over the weekend only among younger patients (aged < 21 years), who instead show the lowest prevalence between Monday and Friday. According to National data (Istat, 2013) [7, 18], this weekend peak, as well as the higher prevalence during evenings and nights, could reflect the predominant lifestyle adopted by young people during recent years, which includes an increase in socially shared and emulated transgressions and risky behaviors [24].

Furthermore, it is necessary to consider that both occasional and chronic alcohol abuse may be a sign of social and personal difficulties, such as those related to work, social, or personal problems, all of which appear to have increased in the past few years [25].

Data are in accordance with what has emerged from the BALs calculated for accesses during the week relative to age: the majority of younger patients (< 20 years) were classified into the lower level category, and had probably resorted to ED care either because they were less experienced and more disconcerted by their symptoms, or perhaps due to a rising tendency among younger generations to seek professional care no matter the extent of symptomatic distress. In accordance with national data regarding the higher prevalence of alcohol consumption [24], alcohol intoxication with clinical symptoms (moderate and critical and BALs) is mostly found among 31- to 40-year-olds, and even more among 41- to 50-year-olds, probably, as above discussed, mainly due to social, economic, and personal causes. It was within these age groups, as a matter of fact, that we found the majority of divorced/separated and single patients with AAI, which might consequentially have a greater perception of loneliness, and of economic and social dif-

ficulties. On the other hand we must also consider the reverse hypothesis, i.e. that alcohol abuse may lead to break-ups and dysfunction within family life, resulting in a deterioration of the psychological and living situation of the patient. Thirdly, a two way process with both factors influencing each other may be at play, making it hard to discern a clear cause-and-effect relationship. An exception is found in the youngest age group over the weekend, which represents the moment of the week when social gatherings, and hence the tendency to abuse alcohol, are more frequent [26]. Although a small number of patients were classified and treated with Code Red (1.9%), which is certainly the most challenging and expensive in sanitary terms, the higher prevalence of Green (67.7%) and Yellow (25.8%) Codes are still responsible for the largest portion of healthcare costs for the ED in terms of space, personnel involved, and specific therapeutic interventions [27].

## Conclusions

Access to the ED is only the tip of the iceberg of the acute and chronic alcohol abuse that represents a major health and social problem among the population [28-30].
The significantly high costs encountered are related to a number of factors: medical interventions (use of emergency services and medical specialists, and diagnostic and therapeutic care); social burden (due to road, occupational, and domestic accidents, risky sexual behaviors, fetal alcohol spectrum disorders, criminal acts, social problems and psychiatric distress, and increase in alcohol-related diseases); and years of life lost due to premature mortality and disability attributable to alcohol (DALYs) [25, 31, 32].
The findings highlighted by this study concern only ED accesses, but in fact the phenomenon is much more serious when one considers that in Italy official statistics report a minimum of 30,000 hospitalizations and 800 deaths per year due to AAI [15].
EDs undoubtedly play a very important role when it comes to providing relief to the phenomenon of AAI.
The worrisome entity of the impact of alcohol abuse, however, calls for an improved implementation of public health activities on a national scale, in order to prevent excessive alcohol use throughout the country and across all the ages. Policies should include a substantial involvement of the school system, of pediatricians and family doctors, of Local Health Departments, and of mass communication platforms (including e-channels).

### References


[1] *Global status report on alcohol and health – 2014ed.* WHO Library Cataloguing-in-Publication Data. http://apps.who.int/iris/bitstream/10665/112736/1/9789240692763_eng.pdf_

[2] *Management of substance abuse. Alcohol.* (WHO). http://www.who.int/substance_abuse/facts/alcohol/en/

[3] Rehm J, Mathers C, Popova S, et al. *Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders.* Lancet 2009;373:2223-33.

[4] ISTISAN ALCOOL 2014/1 http://www.iss.it/binary/alco4/cont/Rapporto_Istisan.pdf

[5] Anderson P., Møller L., Galea G. *Alcohol in the European Union Consumption. Harm and policy approaches (WHO, 2012)* http://www.euro.who.int/__data/assets/pdf_file/0003/160680/e96457.pdf

[6] *Alcol e patologie correlate.* Dipartimento delle dipendenze – Regione del Veneto, AULSS 20. http://alcol.dronet.org/epidemiologia/europa.html

[7] ISTAT. *Uso e abuso da alcol in Italia, ed. 2012.* http://www.istat.it/it/archivio/59033.

[8] Scafato E, Ghirini S, Gandin C, et al. *La mortalità alcol-attribuibile in Italia e nelle Regioni. Analisi delle cause di morte nei contesti territoriali e nelle differenti generazioni.* Istituto Superiore Di Sanità. Osservatorio Nazionale Alcol (ONA) – CNESPS. Reparto Salute della Popolazione e suoi Determinanti. Rapporto 2014 http://www.epicentro.iss.it/alcol/apd2014/OK%20SCAFATO%20REVISIONE%20FACSHEET%20mortalit%C3%A0%20alcol%20attribuibile.pdf

[9] Gandin C, Galluzzo L, Martire S, et al. *Epidemiologia e monitoraggio alcol-correlato in Italia e nelle Regioni.* Valutazione dell'Osservatorio Nazionale Alcol-CNESPS sull'impatto del consumo di alcol ai fini dell'implementazione delle attività del Piano Nazionale Alcol e Salute. Rapporto 2014. http://www.iss.it/binary/publ/cont/14_1web.pdf

[10] DSM-IV-TR. *Criteria for Alcohol Intoxication.* http://www.hamsnetwork.org/dsm_intoxication/

[11] National Institute on Alcohol Abuse and Alcoholism. *Alcohol Use Disorder: A Comparison Between DSM-IV and DSM-5.* http://pubs.niaaa.nih.gov/publications/dsmfactsheet/dsmfact.pdf

[12] Pirmohamed M, Brown C, Owens L, et al. *The burden of alcohol misuse on an inner-city general hospital.* QJM 2000; 93:291-5.

[13] Charalambous MP. *Alcohol and the accident and Emergency Department: a current review.* Alcohol 2002;37:307-12.

[14] Pletcher MJ, Maselli J, Gonzales R. *Uncomplicated alcohol intoxication in the Emergency Department: an analysis of the National Hospital Ambulatory Medical Care Survey.* Am J Med 2004;117:863-7.

[15] Botti P, Cipriani F, Dannaoui B, et al. *Acute intoxications and poisonings in Italian Emergency Rooms.* Ann Ist Super Sanita 2006;42:287-97.

[16] *The Surgeon General's Call to Action To Prevent and Reduce Underage Drinking. Appendix B: DSM-IV-TR Diagnostic Criteria for Alcohol Abuse and Dependence.* http://www.ncbi.nlm.nih.gov/books/NBK44358/

[17] *Alcohol Dependence and Alcohol Abuse/Harmful Use: DSM-IV and ICD-10 Diagnostic Criteria.* http://www.estreatment.org/wordpress/wp-content/uploads/2012/10/adc.pdf[18] Tavole Istat, http://www.istat.it/it/archivio/117897

[19] http://www.epicentro.iss.it/alcol/apd2014/APD2014_perez_last1.pdf

[20] Pelliccia A. *Alcol e immigrazione. Uno sguardo transculturale*. Istituto di Fisiologia Clinica Sezione di Epidemiologia e Ricerca sui Servizi Sanitari CNR 2006. https://www.epid.ifc.cnr.it/AreaDownload/Pubblicazioni/ALCOL_E_IMMIGRAZIONE_%20Uno_sguardo_transculturale.pdf

[21] Salvioni P, Soria A, Feltrin R. *Tossicodipendenza e immigrazione: analisi sociodemografica degli utenti di un servizio a bassa soglia d'accesso a Milano*. Boll Farmacodip e Alcoolis 2002;XXV. http://www.unicri.it/min.san.bollettino/bulletin_it/2002-1%202/p101.pdf

[22] Carta MG, Bernal M, Hardoy MC, et al. *Migration and mental health in Europe (the state of the mental health in Europe working group: appendix 1.* Clin Pract Epidemiol Ment Health 2005;1:13-8.

[23] Dota F, Di Giorgio G, Perez M, et al. *Il consumo di alcol tra i cittadini stranieri.* Istat. Alcol Prevention Day, Roma, 9 Aprile 2014. http://www.epicentro.iss.it/alcol/apd2014/APD2014_

perez_last1.pdf[24] Mancinelli R, Fidente RM, Draisci R. Donna e alcol: aggiornamenti in tema di ricerca clinica e preclinica. Rapporti ISTISAN 13/36 http://www.iss.it/binary/alco4/cont/Rapporto_Istisan.pdf.

[25] *Separazioni e divorzi in Italia. Anno 2012.* Statistiche Istat. http://www.axerta.it/news/124/statistiche_separazioni_divorzi_italia.htm

[26] Cherpitel CJ. *Emergency room and primary care services utilization and associated alcohol and drug use in the United States general population.* Alcohol 1999;34:581-9.

[27] Corser WD, Huebner M, Zhu Q. *Predictors of 2010-2011 Michigan Medicaid Beneficiary Predictors of 2010-2011.* Michigan Medicaid Beneficiary Adverse E-Code Health Care Encounters. J Patient Saf 2015:215-9.

[28] Cherpitel CJ, Ye Y. Risky *Drinking, Alcohol Use Disorders, and Health Services Utilization in the U.S. General Population: Data from the 2005 and 2010 National Alcohol Surveys.* Alcohol Clin Exp Res 2015:305-10.

[29] Cherpitel CJ, Ye Y. *Drug use and problem drinking associated with primary care and emergency roomutilization in the US general population: data from the 2005 national alcohol survey.* Drug Alcohol Depend 2008;97:226-30.

[30] Cherpitel CJ. *Changes in substance use associated with emergency room and primary care services utilization in the United States general population: 1995-2000.* Am J Drug Alcohol Abuse 2003;29:789-802.

[31] Alcohol. Media Centre. *Fact sheet.* January 2015 http://www.who.int/mediacentre/factsheets/fs349/en/

[32] *Alcol. Informazioni generali. EpiCentro.* http://www.epicentro.iss.it/alcol/alcol.asp

■ Received on August 31, 2015. Accepted on September 23, 2015.

■ Correspondence: Silvia Majori, Department of Public Health and Community Medicine, Hygiene and Environmental Occupational Preventive Medicine Division, University of Verona, Istituti Biologici II, strada Le Grazie 8, 37134 Verona, Italy - Tel. +39 045 8027653 - E-mail: silvia.majori@univr.it