Travel Medicine and Infectious Disease 54 (2023) 102607

Contents lists available at ScienceDirect

Travel Medicine and Infectious Disease

journal homepage: www.elsevier.com/locate/tmaid

Profiles of travelers to intermediate-high health risk areas following the reopening of borders in the COVID-19 crisis: A clustering approach

Nidia M. García-Marín [a,c], Gustavo A. Marrero [b,c,d,*], Ana Guerra-Neira [a,c], Almudena Rivera-Deán [a,c]

[a] International Vaccination Center at Santa Cruz de Tenerife, Spanish Ministry of Health, Spain
[b] University of La Laguna, Department of Economics, Spain
[c] University of La Laguna, Research Center of Social Inequality and Governance (CEDESOG), Spain
[d] University of La Laguna, IUDR, Spain

ARTICLE INFO

Keywords:
Traveler profiles
Risk perception
Infectious diseases
COVID-19
K-means

ABSTRACT

Background: The reactivation of international travel in 2021 has created a new scenario in which the profile of the traveler to medium-high health risk areas may well have changed. However, few studies have analyzed this new profile since the reopening of borders in that year.

Methods: We designed an ad hoc questionnaire that was administered face-to-face by our medical team during appointments with 330 travelers in the second half of 2021. Information was collected on the following topics: sociodemographic and socioeconomic status; type of travel and previous travel experience; health status and risk perception (of COVID-19 and tropical infectious diseases). Using all features simultaneously, an unsupervised machine learning approach (k-means) is implemented to characterize groups of travelers. Pairwise chi-squared tests were performed to identify key features that showed statistically significant differences between clusters.

Results: The travelers were clustered into seven groups. We associated the clusters with different intensities of perceived risk of acquiring COVID-19 and tropical infectious diseases on the trip. The perceived risk of both diseases was low in the group "middle or lower middle class young inexperienced male tourist" but high in the group "middle or lower middle-class young with children inexperienced business traveler".

Conclusions: Broadening our knowledge of the profiles of travelers to intermediate-high health risk areas would help to tailor the health advice provided by practitioners to their characteristics and type of travel. In a changing health context, the k-means approach supposes a flexible statistical method that calculates travelers' profiles and can be easily adapted to process new information.

1. Introduction

Characterizing the profile of travelers with an intermediate-high health risk associated with their trip (IHHRT) and providing valuable preventive information before departure is one of the most challenging public health tasks facing specialized travel medicine services such as travel clinics or international vaccination centers (IVC) [1–5]. IHHRT travelers represent an epidemiologically important population group, since they have a high risk of contracting infectious diseases and of importing them to their countries of origin [6–9].

Despite the existence of effective preventive measures for many travel-related health conditions [10–13], a significant number of travelers do not seek health advice before traveling, or, if they do, they fail to comply with the recommendations [12,14]. They are influenced in some cases by a low risk perception [14–17], or by a lack of effective communication of the possible risks on the part of medical advisors [17]. In order to provide travelers with more precise and personalized information, it is useful to determine their various profiles and then to recommend the most appropriate action in each case. However, defining accurate and updated traveler profiles is no easy task [18–20], as multiple features must be considered simultaneously. The main purpose of this paper is to characterize the profile of IHHRT travelers during a relevant and new period in the literature on travel medicine: between May and December 2021, just after the re-opening of borders to

* Corresponding author. University of La Laguna, Department of Economics and Research Center of Social Inequality and Governance, Camino La Hornera, 37, Facultad de Economía, Empresa y Turismo, 38200, San Cristóbal de La Laguna, S/C de Tenerife, Spain.
E-mail address: gmarrero@ull.edu.es (G.A. Marrero).

https://doi.org/10.1016/j.tmaid.2023.102607
Received 20 March 2023; Received in revised form 15 May 2023; Accepted 13 June 2023
Available online 22 June 2023

international travels during the COVID-19 crisis.

In our new context, several recent facts suggest that the profile of IHHRT travelers is changing [21–26]. Before the pandemic, the number of international travelers was rising all over the world, peaking at 1.46 billion people in 2019 [27]. Of these, close to 39% (that is, more than 500 million people) traveled to intermediate-high health-risk areas [6], and the importance of this group has increased in recent years [7,10,28]. Moreover, the diversity of the travelers has also increased [18–20, 27–29], in terms of their destinations, the type of travel, and their socioeconomic status [30–32]. Finally, the COVID-19 crisis and the reactivation of international travel since mid-2021 have created a new context for international travelers to health risk areas [26,33,34]: in addition to the diseases classically associated with certain international destinations, there is now a risk of contracting SARS-COV2 disease [35].

The fact that a major crisis, such as the COVID-19 pandemic, has changed the profile of travelers is not new. There is evidence, for example, that the 2008 financial crisis [36,37] and other health crises such as SARS-2003 [38] changed the profile of travelers, at least for several years. However, in this new scenario, few studies (and none at all in Spain) have analyzed the profile of an international IHHRT traveler just after the reopening of borders to international trips [21,39]. To the best of our knowledge, this is the first study to analyse the profiles of international travelers to IHHRT areas during this period using automatic clustering techniques. More specifically, this paper makes two relevant contributions.

First, we propose the use of an unsupervised machine learning tool (the k-means approach) [40] to generate traveler profiles by clustering information using simultaneously a broad set of features. We apply this technique to an original database of international travelers to IHHRT areas seen at the IVC in Santa Cruz de Tenerife (Spain) between May and December 2021. The approach creates groups of majority travelers who share a common set of characteristics related to the type of travel (destination, duration, reason for traveling, etc.), sociodemographic characteristics (age, gender, etc.), socioeconomic status (occupation, studies, place of residence, etc.) and aspects of health and risk perception (self-perceived health and perceived risk of contracting COVID-19 or tropical-infectious diseases at the destination).

Second, we exploit this information to assess the association between the different groups created and certain relevant features, such as the perceptions of the risk of acquiring COVID-19 or tropical-infectious diseases during the trip. This exercise would help consultants to provide more accurate advice to IHHRT travelers, alerting them to the health risks involved and increasing their compliance with medical recommendations and treatment [31,41].

The k-means clustering technique has frequently been applied in marketing, banking and tourist economics [42], but its use in the traveler literature is fairly novel. An exception is the recent application to generate an international traveler profile in Germany based on pre-pandemic information (between January 25th and February 11th, 2019) [31]. Most previous studies characterizing traveler profiles have typically used sample averages of the traveler features (i.e, considering each feature independently) [43] ignoring the multivariate dimension [44,45], or have grouped the individuals based on a predetermined target variable (i.e., being/not being vaccinated, having/not having contracted a disease, etc.) [44–48]. Most works included in this second approach, which analyses travelers' Knowledge, Attitudes and Practices (KAP) [44,46,48,49], provides valuable information for identifying gaps between groups in knowledge of public health issues and for designing effective communication and educational programs regarding the importance of travel health precautions. However, when using the KAP procedure to profile travelers, the profile will change with each predetermined target feature considered.

The k-means approach overcomes the concerns with these procedures, as it considers all features simultaneously and the algorithm itself automatically generates the groups using all the available information without focusing on a single feature. Depending on the sample used, the procedure generates groups of majority travelers who share common characteristics. In our case, our results are representative for a given time period (second half of 2021). Our generated traveler profile and conclusions can be used by IVCs that serve travelers with similar characteristics to ours (see Section 2).

2. Material and methods

2.1. The International Vaccination Centers in Spain

The International Vaccination Center (IVC) of Santa Cruz de Tenerife (S/C Tfe.) is part of the IVC network of the Spanish Ministry of Health. There are 29 IVCs distributed throughout the national territory. The Ministry, mainly through these IVCs, has the function of "organizing and guaranteeing the prevention of travelers' diseases and injuries and international vaccination services".

Our IVC is of medium size, with 3 doctors and 2 nurses in 2021, and it covers a population of slightly more than one million inhabitants (the population of the province of S/C Tfe.). Using data extracted from SISAEX (Information System of the Spanish Ministry of Health) for 2021, our center attended a total of about 1.5 thousand of people traveling to international destinations (0.14% of the total population). According to our size (persons attended and number of medical staff), we are very similar to other Spanish IVCs, although there are larger IVCs such as Madrid, Barcelona and Bilbao (see Table S1, Appendix 1).

The amount of information collected in our face-to-face interview (see section 2.2.) is not available for other IVC. However, we can compare them based on a small number of characteristics extracted from SISAEX, such as destination, motive, duration, gender, and age. Table S1 (Appendix 1) shows the distributions for each characteristic in 2021 for a representative number of IVC in Spain. We can conclude that, in general, the majority of centers (independent of their size) have distributions similar to ours.

There are other types of Travel Health Centers that do not belong to the Spanish Ministry of Health. There is no publicly available information about these centers, so we cannot directly compare them with the IVCs. However, there are some studies for specific centers in Barcelona and Las Palmas de Gran Canaria [50,51], which show that the majority of travelers attended also have similar characteristics to those attended in IVCs (at least for the basic characteristics included in Table S1).

2.2. The questionnaire

An anonymous ad hoc questionnaire was designed and conducted between May 1st and December 31st, 2021. It was administered face-to-face by the medical team members at the beginning of the appointment. Intermediate-high health risk destinations were classified as those in areas traditionally considered as tropical, with endemic transmission of vaccine-preventable diseases like hepatitis A, typhoid fever or yellow fever, and those with a moderate-high prevalence of mosquito-borne diseases such as malaria or dengue fever [6].

Our eligibility criteria were: travelers over 18 years of age, to areas of medium to high health risk associated with the trip; the motive cannot be a compulsory trip (i.e., we did not include "ship and aircraft crew" and similar workers); we did not include travelers on very short trips (less than 4 days), assuming very low risk during the trip, or travelers who reported very long trips (i.e., over one year in duration), if these were persons returning home and/or with permanent jobs abroad. Therefore, our sample is not strictly comparable with the information in SISAEX (Table S1), which includes all travelers to international destinations. The face-to-face approach improves the selection of travelers who meet the eligibility criteria.

Each questionnaire was completed in about 15 min. The information collected was fully homogeneous. The survey was tested with the first set of international travelers after the lockdown, between April 1st and April 30th, 2021. Data for the study were collected during the next seven

months. The final sample comprised 330 respondents, representing 73% of all eligible travelers that visited the medical center during that period.

Another advantage of the face-to-face interview is that it was easy to verify the reasons for non-response, which were not systematic reasons that could bias our results. The reasons included: not having time to complete the survey in a very busy day, forgetting to complete the survey, errors in transcribing the responses, the traveler being in a hurry, the absence of medical staff on a particular day, etc.

The questionnaire was made up of 29 questions regarding the type of travel (touris, work, etc.), previous trips (travel experience), risk perception (of COVID-19 and tropical-infectious diseases) and health status.[1] Information about socio-demographic and socio-economic status such as age, gender, civil status, cohabitants, education and occupation was also compiled. Details are shown in the Supplementary material, Appendix 2 (Table S2).

2.3. The clustering approach

The k-means algorithm [40] is the approach used to cluster our set of travelers according to common features. This algorithm divides the entire dataset into k-predefined and non-overlapped clusters, where each traveler can belong to only one group. The approach compiles information on all features and makes the intra-cluster individuals as similar as possible. Each cluster has a centroid, which corresponds to the arithmetical mean of all features corresponding to all travelers within the cluster. Thus, an individual belongs to a particular cluster because it minimizes the Euclidean distance of all his/her features to the associated centroid.

The k-means approach is applied to a specific sample, and the profiles generated would be the most representative and common for that sample. Moreover, the procedure does not force travelers to be grouped based on a single characteristic. Rather, grouping is based on all available characteristics, all of which are considered simultaneously. Thus, the procedure generates groups of majority travelers who share common characteristics, considering the multivariate dimension. This is a major advantage of the approach, making it very versatile and adaptable to any sample and any point in time. However, the clusters generated are sample-specific; hence if we want to extrapolate our results to everyday travel medicine consultations, we would need to apply the approach to a much larger and more representative sample.

Given the random initialization of the algorithm, different initializations may lead to different clusters [53]. We deal with this concern by running the procedure with different random initializations (300 in our case) and select the one that yields the lowest overall sum of squared distances. There is still one decision to make: the determination of the number of clusters, K. We determine K based on two statistical metrics: the Elbow [54] and the Silhouette [55]. See details in Supplementary material Appendix 3.

The clustering approach entails that, in general, clusters are statistically different when all features are considered simultaneously. However, this global significance does not necessarily mean that each individual feature differs between all clusters (i.e., gender may be statistically equal between a certain number of clusters and different between others). Once all travelers are allocated to a cluster, we perform pairwise chi-squared tests to determine the statistical differences between two clusters for each feature. Looking at all the pairwise tests, for all features, we have a clear picture of the predominant aspects in each cluster and we can easily differentiate one cluster from another. We also use a spider chart (Fig. 1) to illustrate a particular application in Section 3.3.

[1] The term "tropical diseases" describes "those diseases that are prevalent in, or unique to, tropical and subtropical regions" [52]. The face-to-face interview improves the reliability of certain responses, such as whether or not the illness was a travel-related tropical infectious disease. For example, if the traveler answered "Yes" to the question "To the best of your knowledge, have you ever been infected with a tropical infectious disease while traveling abroad?", the interviewer asked for details such as: do you remember the name of the disease? was it a disease confirmed by a laboratory test? what were the symptoms of the disease? In this way, we were able to differentiate whether the disease was a tropical infectious disease related to the trip or not.

3. Results and discussion

This section is divided into three parts. First, we describe the average characteristics of the travelers surveyed and highlight some differences compared to other averages before the COVID-19 crisis in Spain. Second, we show the results of the clustering procedure. Third, we focus on the differences in one specific aspect, which is particularly relevant for our purposes but is very difficult to measure: the perceived risk of contracting COVID-19 and tropical-infectious diseases during the trip.

3.1. Average features

In our sample of international travelers to IHHRT areas between May and December 2021 (see Table S2, Appendix 2), 53% of travelers were female, and 52% were aged 35 or older. About 71% were single, and 74% did not have children; 30% of the sample lived alone and 31% with a partner. As for socioeconomic status, 62% had completed higher education, and 83% were in employment; 40% worked in sectors related with a high social class.

More than half of the travelers (56%) were traveling to Sub-Saharian Africa (SSA), 60% on non-organized trips, with family-friends (37%) or with a partner (26%). The reason for travel was mainly tourism (63%) and the duration of the trip was less than 15 days in almost half of the sample. Most of the travelers (73%) had made fewer than five previous trips to IHHRT areas and were therefore classified as inexperienced.

Overall, travelers had no underlying diseases (84%) and had not acquired tropical-infectious disease on previous trips (91%). Most had a good (41%) or very good (49%) self-perceived health status; few (12%) had been diagnosed with COVID-19, and 84% had been vaccinated against the virus. Similar proportions of travelers perceived a low risk of COVID-19 (54%) and tropical-infectious diseases (53%) related to the trip.

Several relevant types of travelers that would be more common in other periods, such as the elderly, people with chronic diseases, or travelers reporting very poor health, are underrepresented in our sample (and in SISAEX, at least for 2021). We recognize that the absence of the most vulnerable travelers could be seen as a weakness of the sample, but in reality it is a finding. Our sample is not biased with respect to our particular time period (immediately after the reopening of international borders in 2021). The reason for this underrepresentation is that this type of persons hardly traveled in 2021, probably due to the effectiveness of massive information about the risk of suffering severe COVID-19 while traveling, which discouraged this group of people from traveling, especially, to areas of high health risk [56,57].

Looking at the existing (still scarce) literature in the tourism sector and comparing our results with related work done in the pre-COVID-19 period, the evidence is clear: the pandemic crisis has changed the profile of travelers (at least in the short term). As noted in the introduction, the fact that an economic and health crisis has changed the profile of travelers (at least for several years) is not new, as it happened previously with the 2008 financial crisis [36,37] and the SARS-2003 health event [38].

In the tourism literature, several papers show that the pandemic has changed tourists' travel behavior and their response to travel risk perception [21,25,26]. Recent studies also emphasize the underlying influence of financial and health crises on the tourist profile [58]. The pandemic has had a psychological and sociological impact on the way people make decisions during travel (at least in the short term). In general, people have become more cautious while traveling. Also, from

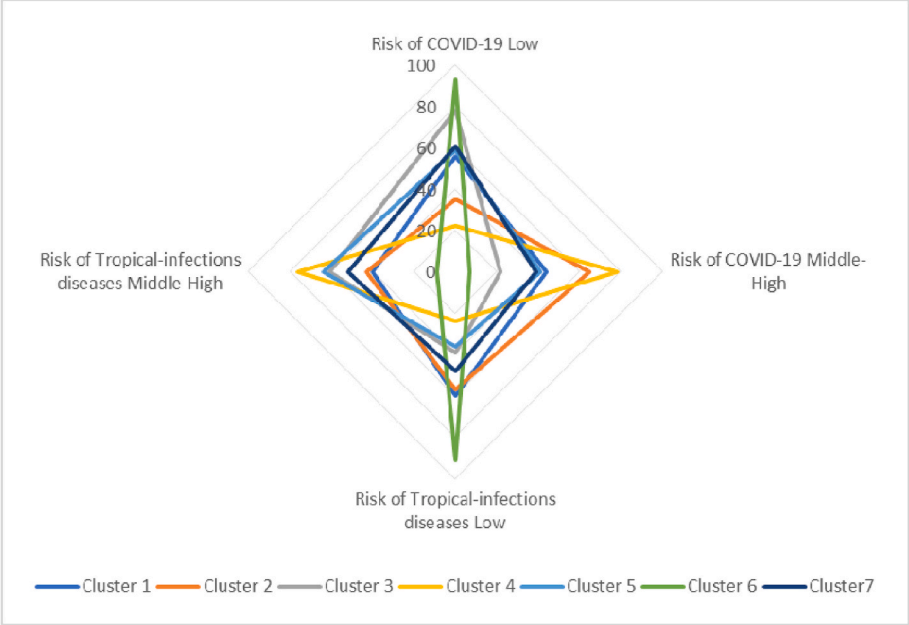

Fig. 1. Clusters and perception of risk of COVID-19 and risk of Tropical-infectious disease related to the trip.
Note: Author's elaboration based on the results of Tables 1 and 3 For each cluster (different colors), the percentage of travelers with a high (horizontal) or low (vertical) perceived risk of COVID-19 (top and right) and tropical infectious diseases (bottom and left) is shown.

an economic point of view, people increase their savings during the crisis and reduce the consumption in all type of goods and services, as well as they change the tourist sensitivity in short- and long-haul destinies.

Most papers (in Spain) analyzing the profile of IHHRT travelers before 2020 focus on a reduced set of characteristics (travel destination, motive, duration, age or sex) and none of the them use our methodology [29,44,59,60]. Aware of these differences, we can compare the averages of several key characteristics (using our results from Table S2). In our sample, travelers to the SSA region were the majority, while travelers to India, China, and almost all destinies to Southeast Asia, and to a lesser extent to South America, represented a small fraction in 2021. However, these latter destinations were at the top of the list in the pre-pandemic period. Also, before 2020, the majority of travelers were middle-aged males and the motive was more concentrated on tourist activities, while in our case the majority of travelers were younger and female, and although tourism is still the most important motive, its relative importance decreased in favor of work and cooperation reasons.[2] Other features included in our sample were either not considered in previous studies or results were not conclusive.

These differences seem to be more related with the sample than with the type of travel clinic considered. For instance, in mentioned papers: in Refs. [59,60], the studies were performed in the IVC of Malaga, simmilar to ours (Appendix 1); in Ref. [29], the analysis was carried out with travelers attending at the USAI BniM travel health clinic in Barcelona, not included in the Spanish Ministry of Health network but similar to our IVC. Finally, in Ref. [44], the survey was carried out in the two main Spanish international airports (Madrid and Barcelona).

3.2. Clustering analysis: travelers' profile

We generate groups of majority travelers who share common characteristics. From our discussion in Section 2.3, our results represent our specific sample, and the profiles generated are the most representative and common for that sample. Therefore, if the procedure does not generate a cluster containing a particular characteristic, it is because: there are not enough travelers in the sample who have this characteristic (i.e. the case of elderly couples, as already discussed in Section 3.1); or travelers are more or less evenly distributed among all clusters and are not concentrated in a particular one.[3]

Seven clusters were generated (see Supplementary material, Appendix 3 for details). Approximately 21% of travelers were in cluster 1 and 22% in cluster 2; cluster 3 was the smallest, with 7% of the sample; clusters 4, 5, and 6 each had approximately 13%, and cluster 7 had 10%. In Table 1, each centroid summarizes the individual features of each cluster and highlights the majority category in each case. In addition, the pairwise (between all clusters, for each feature) chi-squared tests provided information on the key features that presented significant differences (shaded cells) between clusters (details in Supplementary material, Table S3, Appendix 4). Using this information and following the recent paper in the travell tourist literature [31], each cluster was given a name summarizing the key individual and travel features (Table 2).

We describe the main characteristics of the individuals in each cluster, highlighting the differences. The travelers in cluster 1 (*middle-high class adult single inexperienced tourist woman to SSA*) were the most representative of the total averages described above, but with a higher presence of women over 35, single and without children, of a slightly higher social class, with less travel experience and better health status. Travelers in cluster 7 (*middle-high class adult single inexperienced tourist that had COVID-19*) were similar in many aspects to those in cluster 1, but there was one clear difference, namely that 100% had had COVID-19 before traveling. Cluster 2 travelers (*high class adult single inexperienced tourist to Latin America*) differed especially in terms of the type of trip: the group traveled less to SSA and more to other destinations (almost 70% to Latin America), their trips were longer and most were not organized.

[2] These differences in mean values are consistent with a recent but still unpublished manuscript, which compares 2018 with 2021 for travelers in Barcelona.

[3] In our sample, this could be the case for "married men with children who are not business travelers" or "single young women who do not travel to SSA".

Table 1
Proportion of travelers per cluster.

Features		Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
Gender	Male	30	46	43	60	52	65	42
	Female	70	54	57	40	48	35	58
Age	≥35	99	85	48	7	0	5	82
	<35	1	15	52	93	100	95	18
Civil Status	Single	97	93	82	18	41	58	94
	Married	3	7	9	47	39	26	3
	Other	0	0	9	35	20	16	3
Home Structure	Alone	21	32	39	25	50	25	21
	Couple	48	29	22	22	18	33	36
	Family-children	0	6	9	42	32	33	6
	Other	31	33	30	11	0	9	37
Children	None	100	93	91	29	52	45	94
	1–2	0	7	9	56	34	53	6
	≥3	0	0	0	15	14	2	0
Job Status	Unemployed	7	10	0	22	5	0	9
	Retired	0	0	4	9	2	9	0
	Student	9	14	0	0	2	0	3
	Active	84	76	96	69	91	91	88
Occupation Status	Low	16	22	17	76	7	26	12
	Mid	33	17	39	22	52	60	33
	High	51	61	44	2	41	14	55
Education	Primary and Low Sec.	7	1	4	42	0	14	0
	Upper Sec.	20	13	22	42	20	58	30
	Tertiary	73	86	74	16	80	28	70
Reason for trip	Tourist	77	67	69	29	43	81	64
	Business	7	17	9	38	43	12	12
	Cooperation	13	4	9	4	14	2	21
	VFR-others	3	12	13	29	0	5	3
Trip duration	≤15 days	64	36	48	24	66	67	55
	16–30 days	29	40	30	38	14	28	33
	>30 days	7	24	22	38	20	5	12
Trip destination	C.America & Caribbean	14	35	9	7	5	21	12
	Southamerica	23	34	13	24	0	11	21
	North Africa & Mid.East	0	0	4	0	4	5	3
	Subsaharian Africa	60	29	65	65	89	49	55
	Southeast Asia	3	1	0	0	2	14	6
	Rest Asia	0	1	9	4	0	0	3
Type of trip	Organized	54	32	30	16	57	37	52
	Non organized	46	68	70	84	43	63	48
Travel with	Alone	7	20	13	31	18	12	18
	Couple	37	26	26	16	16	23	37
	Family-friends	40	39	48	22	20	56	36
	Other	16	15	13	31	46	9	9
Number of previous trips to risk areas	≤5	87	78	35	89	39	77	82
	>5	13	22	65	11	61	23	18
Underlying diseases	Yes	16	3	17	40	14	16	12
	No	84	97	83	60	86	84	88
Previous Tropical-infectious disease	Yes	0	1	100	9	0	0	9
	No	100	99	0	91	100	100	91
COVID-19 disease	Yes	0	0	4	2	5	7	100
	No	100	100	96	98	95	93	0
Autoperceived Health last year	Bad-normal	10	10	9	15	2	9	12
	Good	41	30	26	56	32	56	49
	Very good	49	60	65	29	66	35	39
COVID-19 vaccine	Yes	83	83	65	89	93	84	82
	No	17	17	35	11	7	16	18
Perception of risk of COVID-19	Low	56	35	78	22	59	93	61
	Middle-High	44	65	22	78	41	7	39
Perception of risk of tropical-infections disease	Low	60	57	39	24	36	91	48
	Middle-High	40	43	61	76	64	9	52

Note: Author's elaboration based on data collected from an anonymous ad hoc questionnaire designed and conducted between May 1 and December 31, 2021. The final sample included 330 respondents: travelers to medium-high health risk destinations. K-means is the clustering approach used. Each column represents the "centroid" of each cluster. We use different random initializations and choose the one that gives the lowest total sum of squared distances. The number of clusters, K, is based on the Elbow and Silhouette metrics.

Most important features that distinguished travelers in cluster 3 (*middle-high class single experienced tourist that had Tropical-infectious disease)* were the fact that they are experienced travelers who mostly travel to SSA (on non-organized trips), and the lowest rate of vaccination against COVID-19, even though they had not had the disease; in contrast, they had all contracted tropical-infectious diseases on previous trips. Cluster 5 (*middle-high class young experienced business traveler*) were all young, upper-middle class (93%), had completed higher education (almost 80%), had experience of traveling to risk areas, and most were traveling for work (i.e., to attend a congress or to visit companies, etc.) or for a short period of time to SSA.

Travelers from clusters 4 (*middle-low class, young, with children, inexperienced business man*) and 6 (*middle-low class, young, inexperienced tourist*) were probably the most different from the other groups. In addition, as we will show and discuss below, they are the most different between them in terms of risk perception during the trip. In both cases,

Table 2
Main features of travellers by clusters.

# of Cluster	Brief description of the cluster	% female	% younth (<35)	% not have children	% tertiary educ.	% reason for trip: tourist	% more than 5 travels	% had COVID-19	% had Tropical-infectious disease
1	*Middle-high class adult single inexperienced tourist woman (to SSA)*	70%	1.4%	100%	73%	77%	13%	0%	0%
2	*High class adult single inexperienced tourist (to Latin America)*	54%	15%	93%	86%	67%	22%	0%	1%
3	*Middle-high class single experienced tourist that had Tropical-infectious disease*	56%	52%	91%	74%	69%	65%	4%	100%
4	*Middle-low class young with children inexperienced business traveler*	40%	93%	29%	16%	29%	11%	2%	9%
5	*Middle-high class young experienced business traveler*	47%	100%	52%	79%	43%	61%	5%	0%
6	*Middle-low class young inexperienced tourist man*	35%	95%	44%	28%	81%	23%	7%	0%
7	*Middle-high class adult single inexperienced tourist that had COVID-19*	57%	18%	94%	70%	63%	18%	100%	9%

Note: Author's elaboration based on the results of Table 1 and Table S3 (Appendix 4). For each cluster, we select the features that are generally more different for each cluster with respect to the others. See also note of Table 1.

they were mostly men, under 35 years old, inexperienced travelers, with a high percentage of children (more than 70% and 55%, respectively), the majority with primary or secondary education (around 85% and 70%, respectively). Their main differences are: duration and motive (more than 15 days on business or visiting relatives in group 4, while less than 15 days and for pleasure in group 6); ocuppational status was mostly low for cluster 4 (22% were unemployed), while it was mostly middle class for cluster 6 (none unemployed); more singles in cluster 4 than in cluster 6 (58% versus 18%); a much higher percentage of underlying disease and perceived risk of both COVID-19 and infectious diseases during the trip for travelers in cluster 4 than in cluster 6. We comment further on this issue in the next section.

3.3. Perceived risk of contracting COVID-19 and other infectious diseases

COVID-19 and infectious tropical diseases related to the trip are currently the two main sources of health risk that travelers face. One of the most important tasks of an IVC is to provide information on these risks. Detecting the groups that, a priori, perceive less risk regarding these diseases can improve the efficacy of the recommendations made by health professionals, as they can be more specific and more persuasive in targeting certain groups of travelers.

The results for the pairwise chi-squared tests in the last two rows in Table S3 (Appendix 4) is used to associate the different clusters with different intensities of the perceived risk of acquiring COVID-19 and tropical-infectious diseases on the trip. Table 3 classifies the seven clusters according with their different levels of perceived risks. Fig. 1 complements results in Table 3.

The group with the highest risk perception was cluster 4, who perceived a high risk of both types of disease. At the other extreme, individuals in cluster 6 perceived a low risk of both diseases. In between, cluster 1 did not present a clear pattern (undetermined risk for both diseases). The results for cluster 3 were interesting (and a priori unexpected) in that they perceived a low risk of COVID-19 but a high risk of infectious diseases; group 7 presented a low perceived risk for COVID-19 and an unclear level of risk of infection diseases. Finally, cluster 5 perceived a high risk of infectious disease and unclear levels of risk in COVID-19, while individuals in cluster 2 presented the opposite trend.

Interestingly, none of the clusters showed a low perceived risk of infection diseases and an unclear or high risk in COVID-19. This is an important finding: perceiving a risk of infectious diseases seems to be a better predictor of perceiving a risk of COVID-19 (and probably of other diseases). In other words, considering the risk of contracting COVID-19 to be low does not imply that one perceives the risk of contracting a contagious disease to be low as well.

Following on from the description of the clusters (Tables 1 and 2), we can associate individual and traveler's characteristics for each cluster with this taxonomy of perceived risk of disease (Table 3). Only travelers in cluster 6 presented a low perceived risk of both infectious diseases and COVID-19; this was the case of more than 90% of members of this group, and it seems that this group would consider the risk of almost any disease to be low. The reason is not the destination, marital status, family structure or the gender; in fact, no single feature characterizes this group. For example, they were mostly men under 35 years old and mostly with childer, as those travelers in cluster 4, a cluster that showed a totally opposite perception of risk (as comented at the end of section 3.2).[4] It is the combination of characteristics that makes their perception of risk low: namely, they are young male tourists with little experience of traveling, on a short trip with family or friends, mostly single but some with children, middle class but none of them are unemployed, mostly with secondary education and in good health. This example illustrates the importance of a multidimensional approach to analyzing these types of issues.

There is an extensive literature analyzing the differences between travelers who seek and those who do not seek pre-travel health advice. In this literature, a recent systematic review ref [12] examined the reasons for not seeking travel advice over a ten-year period (2007–2017). It found that the most important reason for this attitude is a low perceived risk of infectious diseases during travel. In general, this review concludes that lower adherence rates to pre-travel advice (and thus lower risk perception during the trip) are associated with VFR or business travelers, being male, young and middle-aged, moe frequent travelers, being in good health, and traveling for a short period of time. With the exception of motive and experience (we found inexperienced tourist travelers to be more associated with perceived lower risk), these findings are consistent with the characteristics of our travelers in Cluster 6, who have the lowest risk perception in our sample.[5]

Compared to cluster 6, clusters 3 and 7 had a higher perception of the

[4] In Refs. [61,62], they conclude that in general men perceive less risk than women when traveling in a medium to high risk area. According to our results, the high or low perception does not only depend on gender, but on the combination of different characteristics.

[5] Some of these characteristics are also found in other papers [44] for a sample of Spanish travelers to high-risk areas [18]; for Swiss travelers between 2010 and 2012 [17]; for travelers to tropical and subtropical destinations in Switzerland in 2008–2009.

Table 3
Travel clustering and perception of risk of COVID-19 and Tropical-infectious diseases related to the trip.

		Perception of risk of COVID-19		
		Low	Undetermined	High
Perception of risk of tropical-infectious disease	*Low*	Cluster #6: *Middle-low class young inexperienced tourist man traveling with family/friends*	–	–
	Undetermined	Cluster #7: *Middle-high class adult single inexperienced tourist that had COVID-19*	Cluster #1: *Middle-high class adult single inexperienced tourist (to Subsaharin Africa) woman*	Cluster #2: *High class adult single inexperienced tourist (to Latin America)*
	High	Cluster #3: *Middle-high class single experienced tourist that had tropical disease*	Cluster #5: *Middle-high class young experienced business traveler*	Cluster #4: *Middle-low class young with children inexperienced business traveler*

Note: Author's elaboration based on the results of Tables 1 and 2 The main distinguishing characteristics are used to name each group.

risk of infectious diseases. The main difference between these clusters is that the latter groups were more likely to be: from a high social class, over 35 years old, woman, with a previous history of illness (mainly of tropical-infectious disease) and to be more experienced travelers. In this comparison it is worth mentioning the travelers in cluster 3. They perceived a high risk of tropical infectious diseases and a low risk of COVID-19, probably because all of them had contracted a tropical infectious disease on previous trips but had not had COVID-19; moreover, they seem to be the group with the least fear of contracting this disease, because was the group with the largest contingent of people not vaccinated against COVID-19 (35%), in a period when they could have already been vaccinated.[6]

For a given level of perceived risk of acquiring a tropical-infectious disease, the following factors are specific to clusters that perceived higher risk of COVID-19. For example, comparing clusters 1, 2 and 7 (with an undetermined perceived risk of tropical-infectious disease), the destination (Latin America) and the type of travel (non-organized, longer duration), and not having had COVID-19 seem to increase the perceived risk of this disease during the trip.[7]

Comparing clusters 3, 4 and 5 (with a high perception of the risk of tropical-infectious disease), there are few features that undoubtedly mark the difference. One is the reason for the trip being "business", which seems to increase the perceived risk of COVID-19, probably because of the higher social contact involved in this type of trip.

4. Conclusion

The approach described here makes it possible to identify clusters based on a specific set of features. Among these features, travelers' perceptions of the risk of certain diseases during the trip is particularly important and is the focus of the last part of the paper. The results are also useful as a reference point for future comparisons and for an analysis of the possible changes in the profile of IHHRT travelers as time elapses since the COVID-19 crisis.

The use of this innovative statistical method applied to travelers during the worst part of the COVID-19 pandemic (bearing in mind all the public health information available and its probable effect on travelers' perceptions of risk) may help us to target health-related advice to particular traveler groups. In today's rapidly changing health context, with the re-appearance of old diseases and the possible emergence of new ones, it is more important than ever to have a rigorous and flexible statistical procedure that is able to define travelers' profiles and can be easily adapted to new information as it becomes available.

As a future extension, we can use the k-means approach to generate clusters that can be extrapolated to daily travel medicine consultations (i.e., in Spain). However, to do this, we need to "train" the model with a representative sample of travelers from all IVCs in Spain and over many years. Currently, the SISAEX information from the Spanish Ministry of Health could be used for this purpose, but only for a reduced set of characteristics, not including risk perception or other relevant aspects such as health status, underlying diseases or vaccination history. The benefits of improving this type of analysis and databases, including richer and more homogeneous information, would be enormous in terms of the quality and effectiveness of the health information provided by any health travel clinic. These analyses can have a positive impact not only on the health of individuals during travel, but also on global health by reducing the rate of disease importation.

Fundings

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Nidia M. García-Marín: Conceptualization, Methodology, Formal analysis, Resources, Investigation, Data curation, Writing – original draft, Writing – review & editing. **Gustavo A. Marrero:** Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Supervision. **Ana Guerra-Neira:** Resources. **Almudena Rivera-Deán:** Resources.

Declaration of competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors have declared no conflicts of interest.

Acknowledgments

We thank the editor and two anonymous reviewers for their helpful comments. We also thank the participants of the XII Congress of the Spanish Society of Tropical Medicine and International Health, Navarra, Spain, 2022 for their comments. We are indebted to the General Subdictorate of Public Health of the Spanish Ministry of Health for providing us with the SISAEX database, to Juan Marrero González and Yeray Millares Aguiar (from the International Vaccination Center of Santa Cruz de Tenerife) for their help with data collection, and to Pablo Bencomo-Mesa for his assistance with the initial statistical analysis.

[6] In Ref. [31] the authors found a group they called "female globetrotters". The main reason why the majority of persons in this group were not vaccinated was their low perception of the risk of acquiring a disease while traveling, which is consistent with our cluster 3.

[7] Using a different clustering approach than ours [21], found that risk perception for COVID-19 is smaller for shorter trips and experienced travelers. A similar result is found in Ref. [63], which, for a set of undergraduate travelers (from University in Melbourne), they find that younger persons with greater travel experience are willing to take more risk in the trip because they perceive low risk of severe COVID-19.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tmaid.2023.102607.

References

[1] Leggat PA. Risk assessment in travel medicine. Trav Med Infect Dis 2006;4(3):127–34.
[2] Hill DR, Ericsson CD, Pearson RD, Keystone JS, Freedman DO, Kozarsky PE, et al. The practice of travel medicine: guidelines by the infectious diseases society of America. Clin Infect Dis 2006;43(12):1499–539.
[3] Herman J, Patel D. Advising the traveller. Medicine 2018 Jan;46(1):59–65.
[4] Rowe K, Chaves N, Leder K. Challenges to providing pre-travel care for travellers visiting friends and relatives: an audit of a specialist travel medicine clinic. J Trav Med 2017;24.
[5] Freedman DO, Chen LH, Kozarsky PE. Medical considerations before international travel. N Engl J Med 2016;375(3):247–60.
[6] CDC. Yellow book 2024: health information for international travel. Available at: https://wwwnc.cdc.gov/travel/page/yellowbook-home.
[7] Angelo K, Kozarsky P, Ryan E, Chen L, Sotir M. What proportion of international travellers acquire a travel-related illness? A review of the literature. J Trav Med 2017;24.
[8] Ryan ET, Wilson ME, Kain KC. Illness after international travel. N Engl J Med 2002;347(7):505–16.
[9] Freedman DO, Weld LH, Kozarsky PE, Fisk T, Robins R, von Sonnenburg F, et al. Spectrum of disease and relation to place of exposure among ill returned travelers. N Engl J Med 2006;354(2):119–30.
[10] Leder K, Torresi J, Libman MD, Cramer JP, Castelli F, Schlagenhauf P, et al. GeoSentinel surveillance of illness in returned travelers, 2007–2011. Ann Intern Med 2013;158(6):456–68.
[11] Steffen R, Behrens RH, Hill DR, Greenaway C, Leder K. Vaccine-preventable travel health risks: what is the evidence—what are the gaps? J Trav Med 2014;22(1):1–12.
[12] Kain D, Findlater A, Lightfoot D, Maxim T, Kraemer MUG, Brady OJ, et al. Factors affecting pre-travel health seeking behaviour and adherence to pre-travel health advice: a systematic review. J Trav Med 2019 Sep 2;26(6):taz059. https://doi.org/10.1093/jtm/taz059.
[13] LaRocque RC, Rao SR, Tsibris A, Lawton T, Barry MA, Marano N, et al. Pre-travel health advice-seeking behavior among US international travelers departing from Boston Logan International Airport. J Trav Med 2010;17(6):387–91.
[14] Lammert S, Rao S, Jentes E, Fairley J, Erskine S, Walker A, et al. Refusal of recommended travel-related vaccines among U.S. international travellers in Global TravEpiNet. J Trav Med 2016;24.
[15] Hartjes L, Baumann L, Henriques J. Travel health risk perceptions and prevention behaviors of US study abroad students. J Trav Med 2009;16:338–43.
[16] Zuckerman J. Shaping travel health and medicine for the future. Lancet Infect Dis 2001;1(5):296–7.
[17] Zimmermann R, Hattendorf J, Blum J, Nüesch R, Hatz C. Risk perception of travelers to tropical and subtropical countries visiting a Swiss travel health center. J Trav Med 2013;20(1):3–10.
[18] Bühler S, Rüegg R, Steffen R, Hatz C, Jaeger VK. A profile of travelers–an analysis from a large swiss travel clinic. J Trav Med 2014;21(5):324–31.
[19] Han CT, Flaherty G. Profile of travelers with preexisting medical conditions attending a specialist travel medicine clinic in Ireland. J Trav Med 2015;22(5):312–7.
[20] Farnham A, Baroutsou V, Hatz C, Fehr J, Kuenzli E, Blanke U, et al. Travel behaviours and health outcomes during travel: profiling destination-specific risks in a prospective mHealth cohort of Swiss travellers. Trav Med Infect Dis 2022;47:102294.
[21] Neuburger L, Egger R. Travel risk perception and travel behaviour during the COVID-19 pandemic 2020: a case study of the DACH region. Curr Issues Tourism 2021;24(7):1003–16.
[22] Nazneen S. COVID-19 crises and tourist travel risk perceptions. SSRN Electron J 2021.
[23] Rahman MK, Gazi MAI, Bhuiyan MA, Rahaman MA. Effect of Covid-19 pandemic on tourist travel risk and management perceptions. PLoS One 2021;16(9):e0256486.
[24] Gobbi F, Noharet R, Abreu C, Nuñez Maria del Mar Lago, Canale A, Oñorbe MF, et al. South Europe perspective of COVID-19 impact on travel medicine. J Trav Med 2021. taab143.
[25] Kovačić S, Cimbaljević M, Tretyakova TN, Syromiatnikova YA, García Henche B, Petrović MD, et al. How has COVID-19 changed the way we travel? Exploring tourist personality, reactions to the perceived risk and change in travel behavior. Sustainability 2023;15(3).
[26] Agyeiwaah E, Adam I, Dayour F, Badu Baiden F. Perceived impacts of COVID-19 on risk perceptions, emotions, and travel intentions: evidence from Macau higher educational institutions. Tour Recreat Res 2021;46(2):195–211.
[27] UNWTO. Available at: https://www.e-unwto.org.
[28] Rack J, Wichmann O, Kamara B, Günther M, Cramer J, Schönfeld C, et al. Risk and spectrum of diseases in travelers to popular tourist destinations. J Trav Med 2005;12(5):248–53.
[29] Valerio L, Martínez O, Sabrià M, Esteve M, Urbiztondo L, Roca C. High-risk travel abroad overtook low-risk travel from 1999 to 2004: characterization and trends in 2,622 Spanish travelers. J Trav Med 2005;12(6):327–31.
[30] Glaesser D, Kester J, Paulose H, Alizadeh A, Valentin B. Global travel patterns: an overview. J Trav Med 2017 Jul 1;24(4). https://doi.org/10.1093/jtm/tax007.
[31] Bravo C, Castells VB, Zietek-Gutsch S, Bodin P, Molony C, Frühwein M. Using social media listening and data mining to understand travellers' perspectives on travel disease risks and vaccine-related attitudes and behaviours. J Trav Med 2022;29(2).taac009.
[32] LaRocque RC, Rao SR, Lee J, Ansdell V, Yates JA, Schwartz BS, et al. Global TravEpiNet: a national consortium of clinics providing care to international travelers-analysis of demographic characteristics, travel destinations, and pretravel healthcare of high-risk us international travelers, 2009-2011. Clin Infect Dis 2012;54(4):455–62.
[33] UNWTO-data COVID. Available at: https://www.unwto.org/tourism-data/international-tourism-and-covid-19.
[34] Prapaso S, Luvira V, Lawpoolsri S, Songthap A, Piyaphanee W, Chancharoenthana W, et al. Knowledge, attitude, and practices toward COVID-19 among the international travelers in Thailand. Tropical diseases, travel medicine and vaccines 2021;7(1):29.
[35] Ma T, Heywood A, MacIntyre CR. Travel health risk perceptions of Chinese international students in Australia - implications for COVID-19. Infect Dis Health 2020 AUG;25(3):197–204.
[36] Roure S, Pérez-Quílez O, Vallès X, Martínez-Cuevas O, Sabrià M, Valerio L. La crisis económica condiciona la aparición de un nuevo perfil de riesgo en viajeros internacionales? Rev Clínica Española 2015;215(8):439–45.
[37] Khalid U, Luke EO, Shafiullah M. The effects of economic and financial crises on international tourist flows: a cross-country analysis. J Trav Res 2020;59(2):315–34.
[38] Wilder-Smith A. The severe acute respiratory syndrome: impact on travel and tourism. Trav Med Infect Dis 2006 Mar;4(2):53–60.
[39] Kusumaningrum D, Wachyuni S. The shifting trends in travelling after the COVID-19 pandemic. International Journal of Tourism & Hospitality Reviews 2020;7:31–40.
[40] Steinley D. K-means clustering: a half-century synthesis. Br J Math Stat Psychol 2006 May;59(Pt 1):1–34.
[41] Schlagenhauf P, Weld L, Goorhuis A, Gautret P, Weber R, von Sonnenburg F, et al. Travel-associated infection presenting in Europe (2008-12): an analysis of EuroTravNet longitudinal, surveillance data, and evaluation of the effect of the pre-travel consultation. Lancet Infect Dis 2015 Jan;15(1):55–64.
[42] Bonhomme S, Manresa E. Grouped patterns of heterogeneity in panel data. Econometrica 2015;83(3):1147–84.
[43] Herck KV, Zuckerman J, Castelli F, Van Damme P, Walker E, Steffen R. Travelers' knowledge, attitudes, and practices on prevention of infectious diseases: results from a pilot study. J Trav Med 2003 Mar;10(2):75.
[44] Lopez-Velez R, Bayas J. Spanish travelers to high-risk areas in the tropics: airport survey of travel health knowledge, attitudes, and practices in vaccination and malaria prevention. J Trav Med 2007;14(5):297–305.
[45] Duval B, Serre G, Shadmani R, Boulianne N, Pohani G, Naus M, et al. A population-based comparison between travelers who consulted travel clinics and those who did not. J Trav Med 2003;10:4–10.
[46] Toovey S, Jamieson A, Holloway M. Travelers' knowledge, attitudes and practices on the prevention of infectious diseases: results from a study at johannesburg international airport. J Trav Med 2004;11(1):16–22.
[47] Hamer DH, Connor BA. Travel health knowledge, attitudes and practices among United States travelers. J Trav Med 2004 Jan;11(1):23–6.
[48] Van Herck K, Van Damme P, Castelli F, Zuckerman J, Nothdurft H, Dahlgren A, et al. Knowledge, attitudes and practices in travel-related infectious diseases: the European airport survey. J Trav Med 2004 Jan;11(1):3–8.
[49] Laver SM, Wetzels J, Behrens RH. Knowledge of malaria, risk perception, and compliance with prophylaxis and personal and environmental preventive measures in travelers exiting Zimbabwe from harare and victoria falls international airport. J Trav Med 2001 Nov;8(6):298.
[50] Valerio L, BalanzÃ³ Xd. La consulta de medicina tropical. An del Sist Sanit Navar 2006;29:97–104.
[51] Jaén-Sánchez N, Suárez-Hormiga L, Carranza-Rodríguez C, Hernández Cabrera PM, Pisos Álamo E, García-Reina L, et al. Características demográficas, quimioprofilaxis antimalárica e inmunoprofilaxis en 6.783 viajeros internacionales atendidos en una unidad monográfica. Revista Española de Quimioterapia; 2016.
[52] Manson's Tropical Infectious Diseases. Twenty-third ed.
[53] Cao F, Liang J, Jiang G. An initialization method for the K-Means algorithm using neighborhood model. Comput Math Appl 2009;58(3):474–83.
[54] The determination of cluster number at k-mean using Elbow method and purity evaluation on headline news. - 2018 international seminar on application for technology of information and communication. 2018.
[55] Wang F, Franco-Penya H, Kelleher J, Pugh J, Ross R. An analysis of the application of simplified silhouette to the evaluation of k-means clustering validity. 2017.
[56] Breza E, Stanford FC, Alsan M, Alsan B, Banerjee A, Chandrasekhar AG, et al. Effects of a large-scale social media advertising campaign on holiday travel and COVID-19 infections: a cluster randomized controlled trial. Nat Med 2021;27(9):1622–8.
[57] Yu M, Li Z, Yu Z, He J, Zhou J. Communication related health crisis on social media: a case of COVID-19 outbreak. Curr Issues Tourism 2021;24(19):2699–705.
[58] Senbeto DL, Hon AHY. The impacts of social and economic crises on tourist behaviour and expenditure: an evolutionary approach. Curr Issues Tourism 2020;23(6):740–55.

[59] López-Gigosos R, Segura Moreno M, Mariscal-Lopez E, Gutierrez-Bedmar M, Mariscal A. Prevalence of chronic diseases among international travelers seeking pretravel medical advice in 2018 at Malaga, Spain. Am J Trop Med Hyg 2020:102.
[60] Segura Moreno M, López-Gigosos R, Mariscal-Lopez E, Gutierrez-Bedmar M, Mariscal A. Trends in the travelers' demand for pre-travel medical advice at a Spanish International Vaccination Center between 2000 and 2017. PLoS One 2019; 14:e0217588.
[61] Floyd M, Pennington-Gray L. Profiling risk perceptions of tourists. Annals of Tourism Research - ANN TOURISM RES 2004;31:1051–4.
[62] Lepp A, Gibson H. Tourist roles, perceived risk and international tourism. Ann Tourism Res 2003 JUL;30(3):606–24.
[63] Akritidis J, McGuinness SL, Leder K. University students' travel risk perceptions and risk-taking willingness during the COVID-19 pandemic: a cross-sectional study. Trav Med Infect Dis 2023;51:102486.