BMJ Open

What is the impact of health literacy on healthcare costs? A systematic review and evidence synthesis

Francesca Tusoni, Annalisa Iagnemma, Riccardo Mastrantonio, Mario Muselli, Leila Fabiani, Stefano Necozione

To cite: Tusoni F, Iagnemma A, Mastrantonio R, *et al*. What is the impact of health literacy on healthcare costs? A systematic review and evidence synthesis. *BMJ Open* 2025;**15**:e108816. doi:10.1136/bmjopen-2025-108816

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2025-108816).

Received 02 August 2025
Accepted 24 November 2025

Check for updates

Department of Life, Health and Environmental Sciences, University of L'Aquila, L'Aquila, Abruzzo, Italy

Correspondence to
Riccardo Mastrantonio;
riccardo.mastrantonio@univaq.it

ABSTRACT

Objectives Higher levels of individual health literacy have been associated with better health outcomes, greater medication adherence and improved self-management of chronic conditions. Hence, higher health literacy levels are expected to be indirectly associated with lower healthcare costs. The aim of this review is to identify and synthesise available studies on the relationship between individual health literacy and healthcare costs.

Design Systematic review with qualitative evidence synthesis.

Data sources MEDLINE, Scopus, Web of Science Core Collection and CINAHL were searched up to 7 March 2025.

Eligibility criteria We considered only studies that investigated and quantified the relationship between individual health literacy and healthcare costs sustained by individuals, insurance companies or health providers.

Data extraction and synthesis Article screening and data extraction were performed by two authors independently. We critically appraised the identified study by using the AXIS checklist and evaluated the methodology adopted for cost analysis. Finally, we performed a qualitative synthesis of the study results.

Results Of a total of 5801 articles identified, 23 studies met the inclusion criteria. Almost half of the studies were conducted in the USA and about one-third in European countries. The included studies showed fair average quality and great heterogeneity in health literacy measures and cost analyses. The analyses considered general medical, treatment-related, inpatient, outpatient and emergency costs, out-of-pocket expenses and financial hardship. Fifteen studies reported statistically significant results, estimating the association between health literacy and costs or evaluating the difference in costs incurred by different health literacy subgroups. All study results supported the hypothesised negative association between health literacy levels and healthcare costs.

Conclusion Individual health literacy was found to be negatively associated with a range of healthcare costs, although the supporting evidence was not always robust. Interventions aimed at containing healthcare expenditure should consider this association, while further research is needed to define its nature.

PROSPERO registration number The review has been registered in the PROSPERO International prospective register of systematic reviews (registration code CRD42023435502).

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ Analysis of health literacy measurement strategies highlighted the discretion often used in the application of the tools.
⇒ Analysis of economic evaluations provided an overview of the methodologies adopted to quantify healthcare costs.
⇒ The considerable heterogeneity of the studies in terms of populations and outcome measures precluded quantitative synthesis.
⇒ Differences among the healthcare systems of the various countries did not allow for a comparison between the studies.

INTRODUCTION

Health literacy (HL) has been defined by the WHO as 'the cognitive and social skills which determine the motivation and ability of individuals to gain access to, understand and use information in ways which promote and maintain good health'.[1] The updated version of the *Health Promotion Glossary of terms* (2021) gives an expanded definition of HL that considers how those skills accumulate (ie, 'through daily activities, social interactions and across generations').[2] The glossary also highlights the link between HL and empowerment, suggesting that HL promotes personal and community empowerment by improving people's ability to access and use health information and services effectively.[1 2] Empowerment is a process that leads people to gain greater control over their actions and choices. Therefore, HL is critical for informed decision-making and essential to a shared decision-making approach to care.[2 3]

Previous studies showed that HL is associated with better self-management of chronic conditions, increased adherence to prescribed medication and treatments, and improved health outcomes and health-related quality of life.[4–11] Furthermore, HL correlates positively with healthy behaviours, such as practising physical activities and consuming fruits and vegetables; on the other hand, it correlates negatively with the frequency of utilisation of medical and surgical services and with referrals to emergency care for minor health problems.[11] As a consequence, HL can indirectly impact healthcare costs (HCC).

The connection between HL and the socio-economic dimension has been partially explored. The European Health Literacy Population Survey, conducted by the M-POHL consortium between 2019 and 2021, shed light on this connection, revealing that financial deprivation correlates negatively with general HL.[11] At the same time, previous studies investigated the economic implications of low HL either directly or indirectly by analysing the impact of HL on healthcare utilisation.[12 13] It is worth noting that the scientific production on HL and associated factors has increased considerably in recent years, so that literature reviews on this subject have also increased. However, studies assessing the association between HL and HCC have not yet been systematically analysed and summarised.

Hence, the present study aims to bridge this gap by systematically reviewing and synthesising the studies assessing the association between HL and costs sustained by citizens, insurance companies or healthcare providers.

Review question

This review aimed to answer the following question: *What is the impact that individual health literacy has on the healthcare costs sustained by citizens, insurance companies and healthcare systems?*

MATERIALS AND METHODS

The present study is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist.[14] The review was registered in the PROSPERO register of systematic reviews. With respect to the registered protocol, we used a different checklist for critical appraisal, as the selected studies were mainly cross-sectional.

Eligibility criteria

We searched for studies that evaluated the association between health literacy levels (HLL) and HCC. No restrictions were applied for publication dates or reporting language. However, only observational studies were included in the review. We deemed eligible studies that:

- Investigated the presence of an association between HLL and HCC.
- Quantified the association between HLL and HCC.

We included both studies that considered HCC as a dichotomous variable and studies that quantitatively analysed HCC. Similarly, we considered both studies that treated HL as a continuous score and those that stratified HL levels. We included studies that assessed the difference in costs incurred by groups with different HL levels, as well as studies that quantified the relationship between HL and HCC using measures of association. Studies measuring literacy or disease-related knowledge instead of HL were not considered for inclusion. We also exclude studies assessing indirect costs (eg, health outcomes, frequency of access to healthcare services, etc.).

Information sources and search strategy

We systematically searched relevant electronic databases from inception to 7 March 2025. We considered the following databases, which only index peer-reviewed journals: MEDLINE (via PubMed), CINAHL (via EBSCOhost), SCOPUS and Web of Science Core Collection. The initial search string was constructed for PubMed after several search attempts using combinations of keywords. The string developed was then adapted to the other databases. The database search was performed by selecting 'all fields' and without applying any filters. Details of the search strategy are provided in online supplemental material (Search strings). We did not consider grey literature for our research.

Study selection

The study selection was carried out by three reviewers with a background in Public Health (FT, RM and AI), in compliance with the inclusion and exclusion criteria. FT searched the selected databases, collected the results and removed duplicates. The three reviewers made an initial selection based on titles and abstracts (first screening phase). The selected articles were then retrieved and subjected to a full-text screening to determine their eligibility (second screening phase). For articles that were not available on the website of open access, hybrid or subscription journals, we used the university's interlibrary loan service and/or wrote to the article authors to retrieve the full texts. Each article was screened by two reviewers independently, and any conflict was resolved by consensus in both the screening phases. Finally, a list of eligible studies was compiled, and their references were reviewed to identify other studies to be included. We did not search for articles that cited the selected studies. The reviewers carried out all the selection process manually by using Rayyan and Zotero software.

Data collection

One reviewer (FT) collected data from all the articles included, then two reviewers (RM and AI) double-checked the collected data. Descriptive characteristics were extracted, together with characteristics of outcome measures, type of covariates and main findings. Descriptive characteristics included authors, year of publication, study design, country, time of data collection, population and sample characteristics. Characteristics of outcome measures included tools and scoring systems used to quantify HLL, as well as sources and type of considered costs.

Critical appraisal

To critically appraise the selected studies, we used the Appraisal tool for Cross-Sectional Studies (AXIS tool).[15] The AXIS tool is a 20-item checklist developed to assess the quality of study design and reporting, and the potential sources of bias in cross-sectional studies. Each item is a question that can be answered *Yes*, *No* or *Do not know*. The AXIS tool does not provide an overall score for the studies to be evaluated, as it only assesses whether the study design or report addressed issues that may affect the study quality. Therefore, we assigned a '+' when an issue was properly addressed by the study, a '-' when an issue was not properly addressed by the study, and a '?' when it was not possible to determine whether or how the study addressed that issue. We also assigned a '+/-' when an issue was partially addressed. We then scored each study by dividing the number of '+'s assigned to a study by the total number of items that were judged as applicable to that specific study. Then, we calculated the mean score of each study. The quality of the studies was rated as 'above average' if they exceeded the mean score or 'below average' otherwise. A '+/-' was not considered as a '+'; we adopted this coding to account for the differences among the various degrees of appropriateness of the methodologies adopted. Please note that we reversely scored item 13 ('Does the response rate raise concerns about non-response bias?') and item 19 ('Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?').

Although some studies included in the review reported not having a cross-sectional design (ie, a prospective/retrospective design), we did not consider applying a checklist for cohort or case-control studies, as HL cannot be considered properly an exposure factor but only a construct (ie, a competence) that can be measured. Therefore, checklists designed for these types of studies would not be suitable for critically evaluating the studies included in our review. Furthermore, the items of the AXIS tool can be easily applied to prospective and retrospective studies. However, unlike the cross-sectional design, the cohort design allows researchers to make causal inferences. Therefore, we considered this aspect in the discussion section. Two investigators (RM and AI) critically appraised the selected studies, while another cross-checked the results of the critical analysis (FT).

Furthermore, due to the typology of studies included in the review, we considered assessing the study approach to cost evaluation. As no existing tool was suitable for this purpose, we chose a selected number of items from a checklist developed for the critical appraisal of cost-of-illness (COI) studies.[16] The COI checklist by Schnitzler and colleagues is the first checklist specifically designed to assess those studies aimed at identifying, measuring and valuing the economic burden of an illness or disease. Although health literacy is not really a disease, it is a construct that affects health outcomes. Therefore, we considered four items of the COI checklist by Schnitzler and colleagues to identify study perspective, costing approach, time horizon and inflation adjustment.

Synthesis method

Data pooling of study results was not recommended due to the great heterogeneity in study populations, outcome measures and methods of cost analysis.[17] Hence, we performed a qualitative synthesis of the study results. We reported study findings using a synthesis table that summarises all types of HCC evaluated by the single studies and whether they were found to be associated with HLL in adjusted analysis. When an association was found and quantified, we also reported the value of the association measure. In the main article text, we included an extract of the synthesis table which reports only statistically significant results, along with a detailed description of its contents. The extended synthesis table—showing results from all the analysed studies—is available as supplementary material.

RESULTS

Study selection

The database search identified 5801 articles: 98 in PubMed, 5208 in Scopus, 192 in Web of Science and 303 in CINAHL. We identified and removed 463 duplicates, then reviewed the remaining articles by title and abstract. We selected 323 articles that were suitable for full-text review. Full-text screening resulted in the exclusion of 298 articles, while two full texts were not retrieved. The remaining 23 articles were included in the final study.[18–40] No additional studies were identified by checking the references of the eligible articles. The selection process is shown in the PRISMA flowchart (figure 1).[14]

Study characteristics

We collected descriptive data for each study included in the review. Except for one study published in 2005,[18] the publication year ranged from 2011[19] to 2025.[20 21] Many studies were conducted in the USA (n=11),[18 19 22 25 26 28 29 31 33 37 40] seven studies were conducted in the European region,[23 24 27 30 32 38 39] four in the Asian region[21 34–36] and one in the African region.[20] Most studies had a cross-sectional design (n=16),[18–21 25 26 30 32–40] six had a cohort design (three prospective[23 24 29] and three retrospective)[22 28 31] and one was a cost-of-illness study.[27] One of the retrospective studies was a replication of a previous study.[23 24] One study reported a population-level analysis, with an undefined sample size,[25] one study reported a county-level analysis that involved 3137 counties, for a total estimated population of over 300 million.[26] In the other studies, sample size ranged from 121[27] to 92 749.[28] Both female and male gender/sex were represented, with a greater prevalence of female gender/sex in studies conducted on child caregivers[20 29] and a greater prevalence of male gender in a study conducted on veterans.[28] Based on the reported data, the age of participants ranged from 15 years[20] to over 85 years.[28]

The study samples included different types of population. Seven studies were conducted on the general population,[21 25 26 30 31 35 36] five studies included individuals with

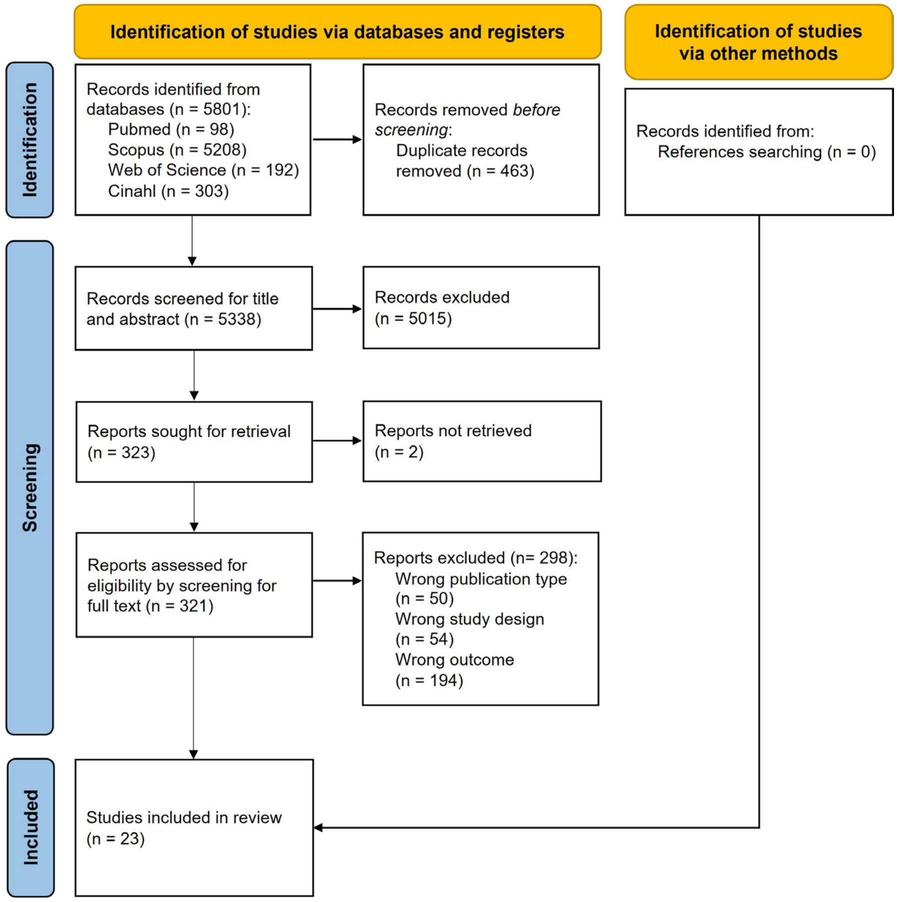

Figure 1 Flowchart of the screening process (Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 flow diagram).

a medical condition (ie, diabetes,[32] cancer,[33] chronic pain,[27] COPD,[39] surgical patients),[22] three studies included child caregivers[20 29 40] and two studies included workers.[23 24] Four studies were conducted on individuals with a specific health plan, of which two included only elderly people.[18 19 37 38] Table 1 shows the descriptive characteristics of the studies.

Regarding outcome measures, a variety of tools was used for measuring HL. The most used were the Three Health Literacy Screening Questions (3HLSQs) developed by Chew and colleagues and further validated by Wallace and colleagues.[41 42] The 3HLSQs is a set of questions, each of which can be used as an independent HL screener, although they are sometimes used together. The 3HLSQs were administered in a modified version in two studies,[19 27] while four studies used only one of the three questions.[32 33 37 38] None of the studies used the tool in its full original version. The European Health Literacy Survey Questionnaire (HLS-EU-Q)[43 44] was used in its full version (47 items) in two studies,[34 39] and in its 16-item version in two other studies.[21 30] The BRIEF Health Literacy Screening Tool (BRIEF)[45] by Haun and colleagues was used in two studies,[22 28] as was the 12th item of the Musculoskeletal Health Questionnaire (MSK-HQ).[23 24 46] The Short Test of Functional Health Literacy in Adults (S-TOFHLA), the Newest Vital Sign (NVS), the Rapid Estimate of Adult Literacy in Dentistry (REALD-30), the Single Item for Literacy Screener (SILS) and the Health Literacy Questionnaire (HLQ) were used each in only one study.[47–51] In general, only one HL-measuring tool was used in each study, except one study, where both NVS and REALD-30 were administered.[29] Together with HL, one study measured health-related knowledge by using 5 out of 13 questions proposed by two of the authors in a previous research study.[52]

Other studies used HL data derived from national surveys, for which ad hoc questionnaires were originally developed: the National Assessment of Adult Literacy (NAAL, USA) and the Chinese Health Literacy Scale (CHLS, China). The NAAL is the national US questionnaire aimed at measuring the literacy levels of adult people.[36] Since 2003, it contains a component designed specifically to measure HL. The CHLS was developed for the purpose of the Chinese national health literacy surveillance system, established by the National Health Commission in 2012.[53 54] Two studies used data driven from the NAAL survey,[25 31] and the other two studies used data from the Chinese national survey.[35 36] Finally, one study used HL scores derived from the Health Literacy Map produced by the University of North Carolina at Chapel Hill.[26]

Table 1 Descriptive characteristics of the studies

Author	Study design	Country	Population	Sample size (n)	Age (years)
Howard *et al*[18]	Cross-sectional study	USA	Non-institutionalised elderly individuals enrolling in a Medicare managed care plan	3260	Mean range=71.6–75.6
Hardie *et al*[19]	1 year cross-sectional study	USA	Individuals enrolled in a Consumer-Driven Health Plan (CDHP) by two national employers	4130	Range=18–65
Cimasi *et al*[25]	Population-level secondary analysis	USA	Missouri residents	Not defined	Not defined
Vann *et al*[29]	Prospective cohort study	USA	Child-caregiver dyads from the Carolina Oral Health Literacy (COHL)	1132 dyads	Mean=27 (parents); 19 months (children)
Franzen *et al* (2014)[38]	Cross-sectional study	Switzerland	Insured persons of the basic health insurance plan of an insurance group	493	Range=35–70
Haun *et al*[28]	Retrospective study	USA	US veterans	92 749	Ranges: <55; 55–64; 65–74; 75–84; ≥85
Mantwill *et al*[32]	Cross-sectional study	Switzerland	Individuals with type 2 diabetes	391	Mean=63.8; Range=38–70
Rasu *et al*[31]	Retrospective study	USA	Individuals aged ≥18 years old	22 599	Mean=48.7
MacLeod *et al*[37]	Cross-sectional study	USA	Individuals having been enrolled in AARP Medicare Supplement plans for at least 6 months	7334	Min=65
McDougall *et al*[33]	cross-sectional survey	USA	Individuals with localised or regional cancer of the colon or rectum	277	Mean=57
Mayer *et al*[27]	Bottom-up cost-of-illness study	Austria	Individuals with chronic pain for 3 months or more	121 (whole sample); 54 (cost-of-illness analysis)	Mean=48.8 (whole sample); 49.6 (cost-of-illness analysis)
Hou *et al*[34]	Cross-sectional study	Taiwan	Community participants	648	Min=65
Shi *et al*[26]	county-level cross-sectional study	USA	People living in all urban and rural counties in the USA	3137 counties; 303 897 545 individuals	LHL and HHL ranges: <64; 65–74; >75
Ture *et al*[39]	Descriptive study	Turkey	Individuals with chronic obstructive pulmonary disease (COPD)	316	Not defined
Gerich *et al*[30]	Cross-sectional study	Austria	General population	590	Range=18–80
Liu *et al*[36]	Secondary analysis of cross-sectional data	China	Residents	6316	Mean=49.1
Puzzitiello *et al*[22]	Retrospective study	USA	Patients who underwent elective inpatient reverse or anatomic shoulder arthroplasty	230	Mean=68.2
Goodwin *et al*[40]	Cross-sectional study	USA	Parents of Children with Medical Complexity	236	Max=19 (children)
Killingmo *et al*[23]	Secondary analysis	Norway	Workers on sick leave due to musculoskeletal disorders	549	Mean=50
Killingmo *et al*[24]	Secondary analysis (Replication study)	Norway	Workers on sick leave due to musculoskeletal disorders	509	Mean=49
Li *et al*[35]	Cross-sectional study	China	Non-institutionalised residents	49 967	Mean=45.3
Boateng *et al*[20]	Cross-sectional study	Ghana	Caregivers of children aged <5 years.	1234	Range=15–69
Lee *et al*[21]	Cross-sectional study	Korea	Participants in the Korea Health Panel Survey (KHPS)	5469	Range=19–65

HHL, high health literacy; LHL, low health literacy.

All but three studies classified HL scores, often based on categorisations used in previous studies. One study used continuous scores, stratifying the analysis for the various questionnaire items.[19] Table 2 describes HL tools and scoring approach for each study.

Cost evaluation

We analysed the strategy adopted for cost evaluation. According to selected items from the COI checklist, we examined the study perspective, the costing approach, the currency and any adjustments for inflation. We also examined the time frame for data collection and reported whether the study analysis was adjusted for covariates (table 3).

The perspective adopted for cost analysis was rarely clearly reported, nor was it easy to infer it by reading the article text. However, at least five studies analysed

Table 2 Health literacy tools and scoring

Author	Tool	Scoring
Howard *et al*[18]	Short Test of Functional Health Literacy in Adults	Three identified categories (inadequate HL, marginal HL and adequate HL), unclear cut-off.
Hardie *et al*[19]	Modified version of the Three Health Literacy Screening Questions by Chew and colleagues	Continuous (for the whole tool and the single items separately)
Cimasi *et al*[25]	Missouri Health Literacy Mapping Tool, based on the National Assessment of Adult Literacy (NAAL)	Below basic (NAAL≤184); basic (NAAL=185–225); intermediate (NAAL=226–309); proficient (NAAL=310–500)
Vann Jr *et al*[29]	Newest Vital Sign (NVS), Rapid Estimate of Adult Health Literacy in Dentistry (REALD-30)	Low HL (NVS<2, REALD-30<13); adequate/high HL (NVS≥2, REALD-30≥13)
Franzen *et al*[38]	Adapted version of one of the Three Health Literacy Screening Questions by Chew and colleagues	Five categories (based on the question answers)
Haun *et al*[28]	BRIEF Health Literacy Screening Tool	Inadequate HL (BHLST=4–12); marginal HL (BHLST=13–16); and adequate HL (BHLST=17–20)
Mantwill *et al*[32]	Adapted version of one of the Three health literacy screening questions by Chew and colleagues	Continuous
Rasu *et al*[31]	Performance test from the National Assessment of Adult Literacy (NAAL) 2003	Below basic or basic HL (HLS<226); above basic (HLS≥226)
MacLeod *et al*[37]	One of the Three Health Literacy Screening Questions by Chew and colleagues	Inadequate HL (answer=not at all/a little bit) and high HL (other answers)
McDougall *et al*[33]	One of the Three Health Literacy Screening Questions by Chew and colleagues	Low HL (answer=not at all/a little bit), medium HL (answer=somewhat, quite a bit) and high HL (answer=extremely)
Mayer *et al*[27]	Adapted version of the Three Health Literacy Screening Questions by Chew and colleagues	Continuous
Hou *et al*[34]	47-item European Health Literacy Survey Questionnaire (HLS-EU-Q)	Sufficient HL (HLS-EU-Q≥34); limited HL (HLS-EU-Q<33)
Shi *et al*[26]	HL scores derived from the Health Literacy Map produced by the University of North Carolina at Chapel Hill	Below basic/basic HL (LHL; HL score≤225) and above basic HL (HHL; HL score>225)
Ture *et al*[39]	Turkish translation of the European Health Literacy Survey Questionnaire (HLS-EU-Q).	Inadequate HL (HLS-EU-Q = 0–25), problematic/limited HL (HLS-EU-Q = 25–33), adequate HL (HLS-EU-Q = 33–42) and perfect HL (HLS-EU-Q = 42–50)
Gerich *et al*[30]	16-item short version of the HLS-EU-Q (subjective HL); 5 out of the 13 items taken from Gerich and Moosbrugger, 2018 (health-related knowledge).	Low subjective HL/high subjective HL (cut-off HLS_EU_Q=3), low health-related knowledge/high health-related knowledge (cut-off=4)
Liu *et al*[36]	Chinese Health Literacy Scale	Adequate health literacy (≥80% of the full score), inadequate HL (<80% of the full score)
Puzzitiello *et al*[22]	BRIEF Health Literacy Screening Tool	Low HL (BHLST score≤17); high HL (BHLST score>17)
Goodwin *et al*[40]	Single Item Literacy Screener (SILS)	For primary analysis adequate HL ('never'/'rarely'), inadequate HL ('sometimes'/'often'/'always'); for sensitivity analysis adequate HL ('never'), inadequate HL (all other responses)
Killingmo *et al*[23]	A single item (Q12) from the Musculoskeletal Health Questionnaire (MSK-HQ)	Three categories (completely/very well understanding, moderate understanding or slightly/no understanding).
Killingmo *et al*[24]	A single item (Q12) from the Musculoskeletal Health Questionnaire (MSK-HQ)	Three categories (completely/very well understanding, moderate understanding or slightly/no understanding)
Li *et al*[35]	Chinese Health Literacy Scale	Adequate (CHLS≥53), intermediate (CHLS=40–52), basic (CHLS=27–39) and below basic (CHLS≤26)
Boateng *et al*[20]	Health Literacy Questionnaire (HLQ)	Seven profiles ('profile 1' = highest mean scores in all nine subscales; 'profile 7' = lowest mean scores in all nine subscales)
Lee *et al*[21]	European Health Literacy Survey Questionnaire (HLS-EU-Q16)	Inadequate HL (HLS-EU-Q16<9 points), problematic HL (HLS-EU-Q16=9–12 points) and sufficient HL (HLS-EU-Q16>12 points).

HL, health literacy.

Table 3 Strategies adopted for cost evaluation

	Perspective	Costing approach	Cost source	Year(s) of data collection	Foreign currency	Inflation adjust	Cov. adjust
Howard *et al*[18]	Patient, healthcare service	Top-down	Prudential Healthcare's administrative claims files	1997	+ ($)	–	+
Hardie *et al*[19]	?	Top-down	Data warehouse of consumer-driven health plan (CDHP) members	2006	+ ($)	–	+
Cimasi *et al*[25]	?	Top-down	Missouri Information for Community Assessment (MICA)	2008	- ($)	–	+
Vann Jr *et al*[29]	Health system	Top-down	Medicaid claims	2008–2009	- ($)	+ (2010)	+
Franzen *et al*[38]	Insurance companies	Top-down	Insurer's administrative database	2012	+ (€)	?	+
Haun *et al*[28]	?	Top-down	Decision Support System National Data Extracts (NDE)	2007–2009	- ($)	–	+
Mantwill *et al*[32]	Insurance companies	Top-down	Insurer's administrative database	2012 (data relating to 2009–2011)	- (CHF)	–	+
Rasu *et al*[31]	Patient	Top-down	Medical Expenditure Panel Survey - Household Component (MEPS-HC).	2005–2008	- ($)	+ (2010)	+
MacLeod *et al*[37]	Patient, national insurance	Top-down	Administrative claims databases	2012–2013	- ($)	+ (2015)	+
McDougall *et al*[33]	Patient/patient's family	Bottom-up	Respondents	2015	n.a.	n.a.	+
Mayer *et al*[27]	Patient, healthcare system, society	Bottom-up	Department of Health Economics Unit Cost Online Database for Austria	2012–2014	- (€)	+ (2016)	–
Hou *et al*[34]	Medical/public health	Bottom-up	Respondents	2018	n.a.	n.a.	+
Shi *et al*[26]	National insurance (Medicaid)	Top-down	Dartmouth Atlas of Health Care	2010–2011	- ($)	No	+
Ture *et al*[39]	?	top-down	National Health Service	2017	+ ($)	No	–
Gerich *et al*[30]	?	Top-down	Register of the Upper Austrian sickness fund	2017	- (€)	No	+
Liu *et al*[36]	Patient	Bottom-up	Respondents	2019	- (CNY)	No	+
Puzzitiello *et al*[22]	?	Top-down	Estimated using time-driven activity-based costs (TDABC)	2018–2021	?	?	+
Goodwin *et al*[40]	?	Top-down	Pediatric Health Information System (PHIS) database	Not specified	- ($)	No	+
Killingmo *et al*[23]	?	Top-down	Norwegian Patient Registry (NPR); Municipal Patient and User Registry (KPR)	2018–2019	+ (€)	+ (2022)	+
Killingmo *et al*[24]	?	Top-down	Norwegian Patient Registry (NPR); Municipal Patient and User Registry (KPR)	2019–2020	+ (€)	+ (2022)	+
Li *et al*[35]	Patient	Bottom-up	Respondents	2020	- (CNY)	No	+
Boateng *et al*[20]	Patient	Bottom-up	Respondents	2017	+ ($)	No	+
Lee *et al*[21]	Patient	Bottom-up	Korea Health Panel Survey	2021	+ ($)	No	+

n.a., not applicable.

costs from the patient perspective. Seven studies used a bottom-up approach, as they collected cost-related data directly from respondents, whereas 16 studies retrieved data from available datasets. Consulted datasets included insurance company databases, national and local registries, and survey secondary data. The data collection time frame was reported by 22 studies and ranged from 1997[18] to 2021.[21 22]

In addition to the seven studies conducted in the USA, another five studies reported costs in US dollars ($). Five European studies reported costs in euros (€), although only 2 of them were conducted in a country that has adopted the European currency (ie, Austria). Two studies reported costs in Chinese Yuans (CNY), while one used Swiss Francs (CHF). As shown in table 3, six studies adjusted the costs for inflation (the year of the currency is reported in brackets). McDougall and colleagues[33] and Hou and colleagues,[34] respectively, considered financial hardship and self-paid vaccination as indicators of health-care costs, without quantifying the total amount.

Twenty-one studies out of 23 analysed the association between HLL and HCC adjusting for covariates. In most studies, covariates included the sociodemographic characteristics of the sample (eg, age, sex/gender, race/ethnicity, education, marital status, occupation, etc.), together with health status (mostly in studies conducted on persons with a medical condition) and health behaviours.

Critical appraisal

We critically analysed the studies included in the review by using the AXIS checklist. The analysis revealed that almost all studies clearly defined their aims (item 1, n=22) and the reference population (item 4, n=22), reported an accurate description of basic data (item 12, n=20), presented results for all the analyses described in the methods (item 16, n=21) or reported results that were internally consistent (item 15, n=20).

Although the target population was defined in most studies, the choice of the sample frame (item 5) and the sample selection process (item 6) were frequently at risk of selection bias. The chosen sample frame was evaluated as representative of the target population in eight studies, whereas in six studies it was difficult to determine whether the sample frame could have been representative. The remaining studies were scored negatively mainly because the sample was selected from a specific setting or context. The sample selection process was positively evaluated in 12 studies that included all subjects from the identified sample frame or used random selection. The selection process was rated as ‘unclear’ (n=9) when it was difficult to determine whether it would lead to a representative sample, due to restrictive inclusion criteria. Items 5 and 6 were judged as not applicable to two studies that analysed data at a population level.

Another issue affecting the representativeness of the sample is the management of non-responders. This issue was disregarded by many of the studies included in the review. Only two studies properly addressed and categorised non-responders (item 7)—one also providing information about non-responders (item 14)—whereas two studies did it partially. The response rate was barely reported (7 out of 13 studies) and it was very low in six studies (<38%). One study reported a response rate of 65%, which is a bit higher than the one reported by others, but still may raise concerns (item 13). We considered items 7, 13 and 14 as not applicable for those studies that used data from previous research or existing dataset.

Only five studies reported the methodology used for sample sizing or selected the number of variables included in the statistical analysis based on the size of the available sample (item 3). We did not apply this item to one population study.

Nearly half of the studies clearly described the methodology adopted for their investigation (item 11, n=17). However, only eight studies fully discussed the study aim, possible selection bias (when applicable), concerns related to non-respondents, confounding factors and non-significant results (item 17). The remaining studies only addressed some of the issues mentioned above in the discussions.

All 23 studies used variables that were appropriate for their aim (item 8). However, the instruments and methods used to measure these variables had in many cases not previously been tested or validated (item 9). Seven studies correctly measured both HL and HCC, whereas 12 studies used unvalidated outcome measures, estimated the outcome, or relied on self-reported data for HL or HCC. Four studies used slightly reliable measures for both HL and HCC.

Two studies did not report any declarations about funding sources or conflicts of interest, 1 study reported a conflict of interest for one of its authors, 1 study reported funding for two authors, while four studies reported being funded by societies or organisation that may have an interest in the study results (item19). All these studies (n=8) were scored ‘?’, as it was not possible to determine whether the fundings or conflicts of interest could have affected the interpretation of results.

Eleven studies reported being approved by an ethical committee and that consent was obtained from participants (item 20). We scored as unclear (‘?’) the remaining studies, as they did not clearly state whether ethical approval was obtained or necessary, although some of them were based on secondary data or on available datasets.

All studies adopted an appropriate design (item 2) and discussed the limitations of their investigation (item 18). Measures used to detect statistical significance (eg, p value, CIs) were deductible from all the studies; however, only 16 studies explicitly mentioned those measures in the methods (item 10).

Table 4 reports the results of the appraisal process.

We assigned a score to each study, based on the critical appraisal. We considered the number of issues properly addressed and the number of items that were applicable to every single study. As figure 2 shows, the average score was 0.68. Thirteen studies scored above average, with 5 scoring 0.8 or higher. Two studies got the highest score (0.88).

Evidence synthesis

The studies included in this review investigated the relationship between HLL and a variety of HCC. The cost

Table 4 Critical appraisal of included studies

	Intro	Methods										Results					Discussion		Other	
Items	1	2	3	4	5	6	7	8	9	10	11	12	13*	14	15	16	17	18	19*	20
Howard et al[18]	+	+	–	?	?	?	n.a.	+	+	+	+/-	+	–	+/-	+	+	+/-	+	?	?
Hardie et al[19]	+	+	–	+	?	?	–	+	–	(+)	+	+	?	–	+	+	+/-	+	?	?
Cimasi et al[25]	+	+	n.a.	+	n.a.	n.a.	n.a.	+	–	+	+/-	+	n.a.	n.a.	+	+	+/-	+	?	?
Vann Jr et al[29]	+	+	–	+	+	?	n.a.	+	+	+	+	+	n.a	n.a	+	+	+	+	+	+
Franzen et al[38]	+	+	–	+	?	+	+/-	+	+/-	(+)	+	+	–	+	+/-	+	+	+	?	+
Haun et al[28]	+	+	–	+	?	+	n.a.	+	+	+	+	+	n.a.	n.a.	+	+	+	+	+	+
Mantwill et al[32]	+	+	–	+	+	?	–	+	+/-	(+)	+	+	–	+/-	+	+	+/-	+	+	+
Rasu et al[31]	+	+	–	+	+	+	n.a.	+	–	+	+	+	n.a.	n.a.	+	+	+/-	+	+	n.a.
MacLeod et al[37]	+	+	–	+	+	+	+	+	+	(+)	+	+	–	–	+/-	+	+/-	+	?	+
McDougall et al[33]	+	+	–	+	+	?	+/-	+	+	+	+	+	–	+/-	+	+	+	+	+	+
Mayer et al[27]	+	+	–	+	–	?	–	+	–	+	+/-	+	?	–	+	+	+/-	+	+	+
Hou et al[34]	+	+	–	+	–	?	–	+	+/-	+	+/-	+	?	–	+	+	+/-	+	?	+
Shi et al[26]	+/-	+	–	+	n.a.	n.a.	n.a.	+	+/-	+	+/-	+	n.a.	n.a.	+	+	+	+	+	n.a.
Ture et al[39]	+	+	–	+	–	?	–	+	+/-	(+)	+/-	+/-	?	–	+	+/-	+/-	+	+	+
Gerich et al[30]	+	+	–	+	–	+	+	+	+	(+)	+	+	–	+/-	+	+	+	+	+	+
Liu et al[36]	+	+	+	+	+	+	–	+	+/-	(+)	+	+	?	–	+	+	+/-	+	?	+
Puzzitiello et al[22]	+	+	–	+	–	+	n.a.	+	+/-	+	+	+	n.a.	n.a.	+	+	+/-	+	?	+
Goodwin et al[40]	+	+	–	+	–	+	–	+	+	+	+	+/-	–	–	+	+	+	+	+	?
Killingmo et al[23]	+	+	+	+	?	+	n.a.	+	+/-	+	+	+	n.a.	n.a.	+	+	+	+	+	+
Killingmo et al[24]	+	+	+	+	?	+	n.a.	+	+/-	+	+	+	n.a.	n.a.	+	+	+/-	+	+	+
Li et al[35]	+	+	+	+	+	+	–	+	+/-	+	+	+	?	–	+	+/-	+/-	+	+	+
Boateng et al[20]	+	+	+	+	–	+	n.a.	+	+/-	+	+	+	n.a.	n.a.	–	+	+/-	+	+	+
Lee et al[21]	+	+	–	+	+	?	n.a.	+	+/-	+	+	+/-	n.a.	n.a.	+	+	+/-	+	+	n.a.

*Reversely scored.
?, unclear; ±, issue partially addressed; +, issue properly addressed; -, issue not properly addressed; n.a, item not applicable.

analyses considered generic medical costs (n=3 studies), treatment-related costs (n=9), outpatient costs (n=4), inpatient costs (n=9), emergency costs (n=6), out-of-pocket expenses (n=4) and financial hardship (n=1). Treatment-related costs were due to medication, restorative oral care and malaria treatment. Outpatient costs also included office visits and preventive costs (vaccination and preventive oral care). Inpatient costs were either generic or disease-specific (lung disease), stage-specific (end-of-life care) or setting-specific (intensive care unit). Emergency costs also included expenses due to emergency oral care.

Fifteen studies reported statistically significant results, five of which computed the difference in HCC between different HL subgroups, whereas the remaining nine analysed the association between HLL and HCC. Table 5 reports only significant results (see online supplemental materials for all results).

Treatment costs

Regarding generic medical costs, Hardie *et al*[19] found a negative association with HL (measured by using each of the adapted 3HLSQs) in individuals enrolled in consumer-driven health plans. MacLeod *et al*[37] reported similar findings for older people with inadequate HL, who incurred higher medical costs than peers with high HL. Interestingly, the authors found these results in both healthy elderly and those with an impaired health status. Both Hardie *et al*[19] and Mantwill *et al*[32] found a negative association between medication costs and HLL (measured by adapted versions of the first and third HLSQs, respectively), in older people and in persons with type 2 diabetes. Hardie[19] also found a negative association between the total amount of medical and medication costs and HL measured by the first and third of the adapted 3HLSQs. Howard *et al*[18] reported no significant difference in costs incurred for medication costs by older people with different HL levels, as measured using the S-TOFHLA.

Outpatient costs

Rasu *et al*[31] found higher outpatient costs for individuals with below basic HL, compared with those with above basic HL, as measured by the NAAL. Similarly,

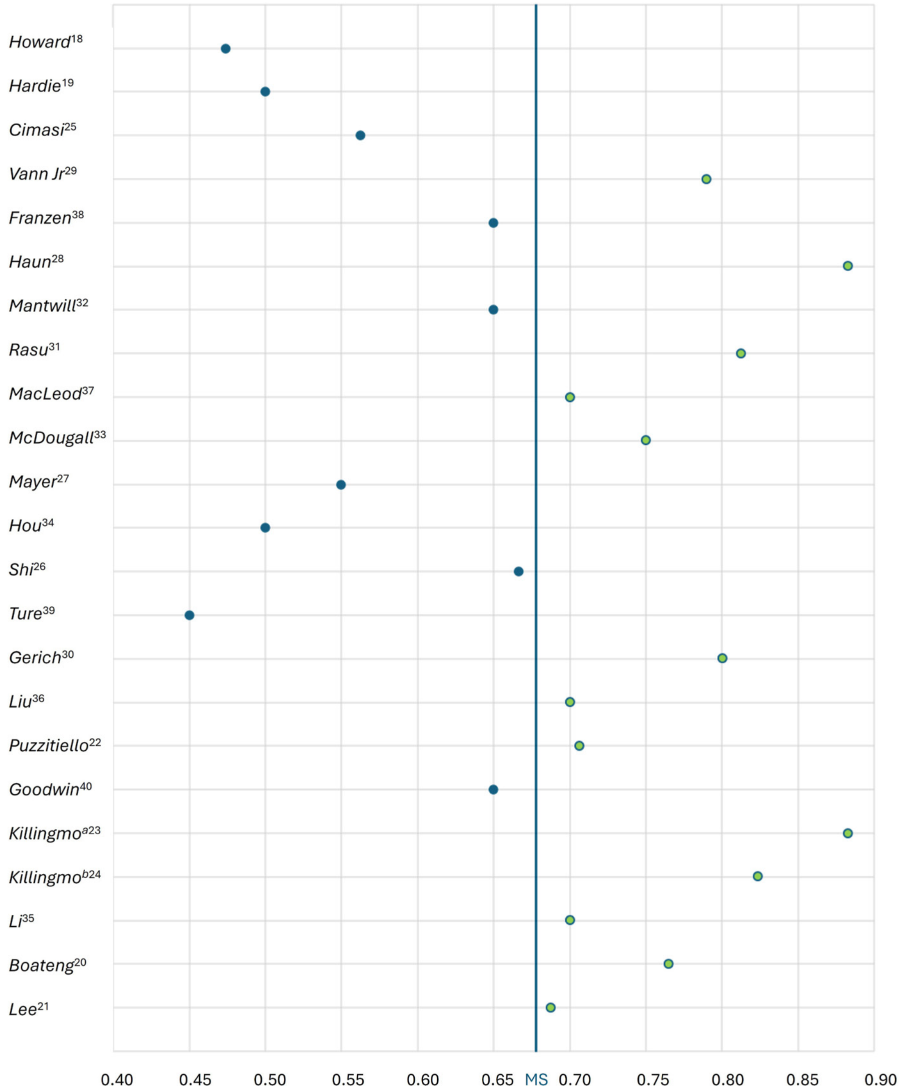

Figure 2 Scoring of studies according to critical appraisal.

Franzen *et al*[38] reported that persons with lower levels of HL (measured by one of the adapted 3HLSQ) incurred higher outpatient costs in both the years considered for the analysis. In contrast, Howard *et al*[18] found that elderly people with adequate HL spend more on outpatient visits than those with marginal HL. Regarding preventive costs, Hou *et al*[34] found that older people with higher HL levels (HLS-EU-Q47) are more likely to self-pay for vaccination.

Inpatient costs

Hardie *et al*[19] found a negative association between costs related to admission or admission facilities and HL. The association was found either when measuring HL by using the whole 3HLSQ, adapted version, or when using one of the single questions, (except when using the second question, in assessing admission costs). Ture *et al*[39] found no statistically significant differences in pulmonary admission costs incurred by individuals with COPD with different HL levels (HLS-EU-Q47). Howard *et al*[18] reported higher inpatient costs for older people with inadequate HL. Similarly, Goodwin *et al*[40] reported a higher risk of incurring higher child hospitalisation costs for parents with an HL score (SILS) lower than the highest. Franzen *et al*[38] reported no statistically significant association between HLL and inpatient costs for insured people. Shi[26] found a negative association between HLL of US residents (derived from the Health Literacy Map) and costs related to end-of-life care. Finally, Cimasi[25] reported an inverse correlation between HLL of Missouri residents (derived from the NAAL) and preventable hospitalisation charges.

Emergency costs

Both Rasu *et al*[31] and Shi *et al*[26] reported higher emergency costs for US residents with low HLL. Similar findings were reported by Howard *et al*,[18] referring to older people. Similarly, Hardie[19] reported a negative association between the HLL of enrolees in a CDHP and emergency costs. Goodwin *et al*[40] reported a higher risk of incurring higher emergency costs for children whose parents had a HL score lower than the highest.

Table 5 Synthesis of evidence

		Treatment costs				Outpatient costs		Inpatient costs						ED costs			
Study	Subscale /subgroup / subanalysis	Medical	Medication / pharmacy	Medical +pharmacy	Self-paid vaccination (pneumonia)	Outpatient	Office visits	Admission	Admission facilities	Inpatient	End -of -life care (last 6 months)	End-of-life care (terminal)	Preventable hospitalisation	Emergency	O-O-P	Financial hardship	Total costs
Howard *et al*[18]	Marginal HL (ref. adequate)					−350$*								+80$*			
	Inadequate HL (ref. adequate)									+1543$**				+108$*			+1551$**
Hardie *et al*[19]	3HLSQ (score)							C: −0.06**	C: −0.06**					C: −0.15***			C: −0.05***
	1Q (score)	C: −0.18**	C: −0.13*	C: −0.17**				C: −0.34***	C: −0.42***					C: −0.51***			C: −0.22***
	2Q (score)	C: −0.16*							C: −0.32*					C: −0.34**			C: −0.18***
	3Q (score)	C: −0.29***		C: −0.23***				C: −0.52***	C: −0.50***					C: −0.96***			C: −0.27***
Cimasi *et al*[25]													not clearly defined				
Franzen *et al*[38]	Lowest HL score (2010)					β: 0.097*											β: 0.088*
	Lowest HL score (2011)					β: 0.119**											β: 0.118**
Haun *et al*[28]	Inadequate HL (ref: adequate)																β: 2.06***
	Marginal HL (ref: adequate)																β: 1.42***
Mantwill and Schulz[32]	(2010)		B: −445.880*														
	(2011)		B: −455.005*														
Rasu *et al*[31]	below basic HL vs above basic HL		+2452$*				+565$*							+51$*			
MacLeod *et al*[37]	Inadequate HL vs high HL (sicker sample)	+1267$***															
	Inadequate HL vs high HL (healthier sample)	+487$***															
McDougall *et al*[33]	Low HL (ref. high HL)															OR: 5.41*	
Hou *et al*[34]	Limited HL (ref. sufficient HL) - No				OR: 1.28*												
Shi *et al*[26]	High HL (ref. Low HL)										β: −0.18**	β: −0.30**					
Ture *et al*[39]	Inadequate HL vs adequate HL													$43.09 vs $26.66 **			225.03$ vs 165.39$*
Liu *et al*[36]	Inadequate vs adequate HL														AME: −0.023*		
Goodwin *et al*[40]	sensitivity analysis (All other scores vs highest HL score)									RaR: 2.26 (95%CI: 1.49 to 3.44)				RaR: 2.88 (95%CI: 1.63 to 5.07)			
Li *et al*[35]															β: −4.97*		

*p<0.05,**p<0.01, ***p<0.001.
$, US dollars; AME, average marginal effect; B, unstandardised regression coefficient; C, undefined coefficient; RaR, rate ratio; RR, risk ratio; β, standardised regression coefficient.

Out-of-pocket costs

Liu *et al*[36] reported higher out-of-pocket costs for Chinese residents (Ningbo province) with inadequate HL compared with those with adequate HL, as measured by the NHLS. Similarly, Li *et al*[53] reported a negative association between HL of Chinese residents (measured using the same scale) and out-of-pocket costs.

Financial hardship

McDougall and colleagues[33] were the only ones to investigate the relationship between HLL (measured by one of the 3HLSQ) and financial hardship in cancer survivors, reporting a positive association.

Total costs

Eight studies investigated the relationship between HLL and the sum of the different HCCs considered for the analysis, with 5 of them reporting statistically significant results. Howard *et al*[18] and Ture *et al*[39] reported significantly higher total HCC incurred by older people and COPD patients with low HLL. Hardie *et al*[19] reported a highly significant negative association between HLL and total HCC in individuals enrolled in a CDHP. Franzen *et al*[38] found that Swiss insured people with low HL are more likely to incur total HCC than their peers with higher HL. Haun[28] reported a higher risk of incurring higher costs for US veterans with marginal and inadequate HL, compared with those with adequate HL, as measured by the BRIEF.

Non-significant results

Eight studies did not report any statistically significant results. Vann *et al*[29] reported non-statistically significant associations between HLL of child caregivers (measured by the REALD-30 and the NVS) and costs related to children's oral care (preventive, restorative and emergency). Gerich *et al*[30] reported no statistically significant associations between HL profiles of the general population—based on the scores of the HLS-EU-Q16 and the five questions investigating health-related knowledge—and pharmacy costs. Both studies by Killingmo and colleagues[23 24] reported no statistically significant associations between HL (measured by one item of the MSK-HQ) and total HCC in workers on sick leave due to musculoskeletal disorders. Boateng *et al*[20] found no statistically significant differences in household costs incurred for the malaria treatment of children under 5 years of age whose caregivers had different HL profiles. Lee *et al*[21] reported no statistically significant association between HLL in the general population (measured by the HLS-EU-Q16) and out-of-pocket expenses. Finally, two studies conducted by Mayer *et al*[27] and Puzzitiello *et al*[22] reported no relationship found between HLL and HCC. Specifically, Mayer and colleagues[27] evaluated the differences in out-of-pocket costs and other costs related to medication, radiological procedures and inpatient and outpatient procedures in subjects with chronic pain and different HL scores (adapted version of the 3HLSQ). Puzzitiello and colleagues[22] evaluated the association between inpatient costs for shoulder arthroplasty and HLL of patients, as measured by the BRIEF. However, both studies did not report any statistical outcome for the evaluated relationship.

DISCUSSIONS

Our review aimed at collecting and synthesising the evidence available on the impact that health literacy may have on healthcare costs sustained by citizens, insurance companies and healthcare systems. We identified 23 studies that investigated the relationship between HLL and HCC in a variety of populations in different countries. Fifteen studies reported statistically significant results, 9 of which were of average or high quality. Findings from most of the studies were coherent and revealed that individuals with low HLL, even from different populations, incur or are more likely to incur higher HCC than their peers with higher HLL. This was reported for generic medical costs, for treatment-related, inpatient, outpatient and emergency costs, for medical out-of-pocket expenses and financial hardship, with greater evidence for inpatient and emergency costs. Two studies reported a positive association between HLL and economic outcomes. In one case, the economic outcome was related to the use of preventive services (ie, self-paid vaccination),[34] which can reduce the risk of incurring higher expenses due to infectious disease. It is worth noting that, in this case, participants with higher health literacy were more likely to have a family doctor and to be able to self-manage illnesses.[34] In the other case, a positive association was reported between HLL and costs linked to outpatient visits.[18] However, the same study showed that individuals with marginal and inadequate HLL incur higher emergency-related costs than individuals with adequate HL and that inpatient and total costs are higher for individuals with inadequate HL. These inverse trends in outpatient costs, on the one hand, and emergency costs, on the other, have already been shown to have a causal relationship.[55]

Previous systematic reviews investigated the association between HL and the use of healthcare services. One found evidence supporting the association between HLL and the use of cancer screening.[56] Others explored the association between HL and vaccine adherence or disease self-management, with inconclusive results.[57–59] Evidence was found to support the association between HL and medication adherence,[60 61] which is known to improve outcomes and quality of life, reducing HCC. This latter research is more in line with the results of our review.

Higher quality evidence was available for financial hardship, medication, outpatient visits, emergency and total costs. Average quality evidence was available for inpatient costs (three studies), medication, outpatient, emergency, generic medical and total costs, and for out-of-pocket expenses.

According to the critical appraisal, only five studies were judged to be of high quality (≥0.80). This is because several studies were found to be at risk of bias. Most of the included studies used instruments that were not validated but only adapted from validated instruments. The choice of a non-valid instrument can increase the risk of measurement bias and affect the study results. Some studies did not collect direct measures of the assessed variables. Instead, they estimated HLL or HCC, based on the available dataset. Although it allows researchers to conduct studies at a population level, this may affect the reliability of the measure.

We found a variety of instruments used to assess HL. Contrary to our expectations, the most used HL measure was the 3HLSQs by Chew and colleagues. This instrument is not frequently used in studies conducted on HL; however, it is quick to administer.[62] The review revealed that brief HL measures are used in specific contexts to routinely screen for HLL.[22 28] These instruments (often made of a single item) are easy and fast to use. On the one hand, that is likely to increase the respondent compliance. On the other hand, they measure just one or a few dimensions of HL, raising concerns about their effectiveness in detecting the real HLL of respondents.[29 41 42] However, quickly assessing the HLL of individuals attending a healthcare facility can help adopt useful strategies to properly manage those identified as being at risk of misunderstanding health information.

There was a great heterogeneity in the methodologies adopted for cost analysis (table 3). Most of the studies did not clarify the chosen perspective, leaving the reader in doubt whether the analysed costs were covered by the patient, the insurance company or the healthcare system. It is worth noting that healthcare systems are complex and diverse from country to country, and what can be considered normal in one country cannot be in another one. Hence, every study assessing healthcare costs should clearly define who is the payer.

Another recurrent risk of bias was linked to the selection of participants. Half of the identified studies seemed to use appropriate selection strategies (eg, random selection, multi-stage selection, etc.), while only eight studies chose an appropriate sample frame (table 4). Choosing an unrepresentative sample frame can nullify the effort made to properly select subjects. Some studies tried to overcome this bias by limiting their findings to the selected sample. This strategy cannot be meaningfully applied when statistical significance is defined. The purpose of defining a level of statistical significance is to determine whether a result found depends on chance and, therefore, whether that result can be extended to a population homogeneous to the one examined. Computing the significance level of the results without extending them to a target population is inconsistent.

Non-respondents can also affect the representativeness of the sample. Only two out of 23 studies properly addressed this issue by categorising non-respondents and reporting their main characteristics, while two studies did it partially.

We have noticed a growing trend in the publication of studies on the subject. Although the first study was published in 2005, over half of the studies that we included were published in the last 4 years. This testifies to the growing interest in research in the field of health literacy, reported by a trend analysis by Qi *et al*[63] (Nutbeam 2020). Qi and colleagues mapped the distribution of research studies on HL by country, revealing that the USA is the leading country in this research field, followed by Australia. This is consistent with our findings, as 11 studies out of 23 were conducted in the USA (table 1). However, from our review, it emerged that other countries have begun to show interest in research on the economic impact of HL. Trends of HL research in European countries are probably influenced by the Action Network on Measuring Population and Organizational Health Literacy (M-POHL), an international network created to measure and improve health literacy in populations and organisations.[11] The M-POHL started its activities in 2018 and, since then, has developed and validated a variety of instruments for measuring HL, two of which are used in 4 of the studies we identified (table 2).

Strengths and limitations

To the best of our knowledge, this is the first systematic review aimed at identifying studies that analysed and quantified the direct relationship between HLL and HCC. We adopted a blind double screening approach to maximise the efficacy of the review process.[64] The set of selected databases should cover a large part of the available scientific literature on the topic. However, some articles may have been left out as they were not indexed in the selected databases or belonged to grey literature. Furthermore, we could not retrieve two articles. In addition, the great heterogeneity of the studies did not allow us to perform a meta-analysis.

Nevertheless, we believe that our review can provide a better understanding of the deep relationship between HL and HCC, supporting interventions aimed at reducing health expenditure by increasing individual HL. It also provides an overview of the methodologies adopted to investigate economic dimensions of health literacy, highlighting the lack of uniformity in the assessment of HL and HCC. Our findings can also help identify gaps and future research needs.

Conclusion

Our study aimed at providing an overview of currently available evidence on the association between health literacy and healthcare costs sustained by citizens, insurance companies and health providers. We found evidence of varying strength supporting a negative association between health literacy and a variety of healthcare costs. Research, policies and practices aimed at reducing healthcare expenditure should consider this association. Future research should clarify whether this association is causal in nature.

Contributors SN and FT conceptualised the research. FT, AI and RM screened and selected the articles to be included. FT, AI and RM extracted the data. FT wrote the original draft. All authors reviewed and edited the draft. RM is the guarantor of the work.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data sharing not applicable as no datasets generated and/or analysed for this study.

ORCID iD

Riccardo Mastrantonio https://orcid.org/0000-0002-0465-3784

REFERENCES

1 Nutbeam D. Health Promotion Glossary. *Health Promot Int* 1998;13:349–64.
2 Nutbeam D, Muscat DM. Health Promotion Glossary 2021. *Health Promot Int* 2021;36:1578–98.
3 Gessler D, Juraskova I, Sansom-Daly UM, *et al*. Clinician-patient-family decision-making and health literacy in adolescents and young adults with cancer and their families: A systematic review of qualitative studies. *Psychooncology* 2019;28:1408–19.
4 Papadakos JK, Hasan SM, Barnsley J, *et al*. Health literacy and cancer self-management behaviors: A scoping review. *Cancer* 2018;124:4202–10.
5 Miller TA. Health literacy and adherence to medical treatment in chronic and acute illness: A meta-analysis. *Patient Educ Couns* 2016;99:1079–86.
6 Hyvert S, Yailian A-L, Haesebaert J, *et al*. Association between health literacy and medication adherence in chronic diseases: a recent systematic review. *Int J Clin Pharm* 2023;45:38–51.
7 Athanasopoulou K, Mentis M, Vathi-Sarava P, *et al*. Health literacy of older adults with musculoskeletal problems: A systematic review. *Int J Orthop Trauma Nurs* 2024;55:101127.
8 Firmino RT, Ferreira FM, Martins CC, *et al*. Is parental oral health literacy a predictor of children's oral health outcomes? Systematic review of the literature. *Int J Paediatr Dent* 2018.
9 Riemann L, Lubasch JS, Heep A, *et al*. The Role of Health Literacy in Health Behavior, Health Service Use, Health Outcomes, and Empowerment in Pediatric Patients with Chronic Disease: A Systematic Review. *Int J Environ Res Public Health* 2021;18:12464.
10 Zaidman EA, Scott KM, Hahn D, *et al*. Impact of parental health literacy on the health outcomes of children with chronic disease globally: A systematic review. *J Paediatr Child Health* 2023;59:12–31.
11 The HLS19 Consortium of the WHO Action Network M-POHL. International Report on the Methodology, Results, and Recommendations of the European Health Literacy Population Survey 2019-2021 (HLS19) of M-POHL. Austrian National Public Health Institute, Vienna, 2021. Available: https://m-pohl.net/sites/m-pohl.net/files/inline-files/HLS19%20International%20Report.pdf
12 Eichler K, Wieser S, Brügger U. The costs of limited health literacy: a systematic review. *Int J Public Health* 2009;54:313–24.
13 Palumbo R. Examining the impacts of health literacy on healthcare costs. An evidence synthesis. *Health Serv Manage Res* 2017;30:197–212.
14 Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev* 2021;10:89.
15 Downes MJ, Brennan ML, Williams HC, *et al*. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ Open* 2016;6:e011458.
16 Schnitzler L, Roberts TE, Jackson LJ, *et al*. A consensus-based checklist for the critical appraisal of cost-of-illness (COI) studies. *Int J Technol Assess Health Care* 2023;39:e34.
17 Glisic M, Raguindin PF, Gemperli A, *et al*. A 7-Step Guideline for Qualitative Synthesis and Meta-Analysis of Observational Studies in Health Sciences. *Public Health Rev* 2023;44:1605454.
18 Howard DH, Gazmararian J, Parker RM. The impact of low health literacy on the medical costs of Medicare managed care enrollees. *Am J Med* 2005;118:371–7.
19 Hardie NA, Kyanko K, Busch S, *et al*. Health literacy and health care spending and utilization in a consumer-driven health plan. *J Health Commun* 2011;16 Suppl 3:308–21.
20 Ofori Boateng M, Asuman D, Kugbey N, *et al*. Health literacy and household financial loss on malaria treatment for children under five in Ghana: a patients' perspective. *Int Health* 2025;17:77–83.
21 Lee J, Min HS, Ryu DH. The association between health literacy, private health insurance, and medical expenditure in South Korea. *BMC Health Serv Res* 2025;25:106.
22 Puzzitiello RN, Colliton EM, Swanson DP, *et al*. Patients with limited health literacy have worse preoperative function and pain control and experience prolonged hospitalizations following shoulder arthroplasty. *J Shoulder Elbow Surg* 2022;31:2473–80.
23 Killingmo RM, Tveter AT, Pripp AH, *et al*. Modifiable prognostic factors of high societal costs among people on sick leave due to musculoskeletal disorders: findings from an occupational cohort study. *BMJ Open* 2024;14:e080567.
24 Killingmo RM, Rysstad T, Maas E, *et al*. Modifiable prognostic factors of high societal costs among people on sick leave due to musculoskeletal disorders: a replication study. *BMC Musculoskelet Disord* 2024;25:990.
25 Cimasi RJ, Sharamitaro AR, Seiler RL. The association between health literacy and preventable hospitalizations in Missouri: implications in an era of reform. *J Health Care Finance* 2013;40:1–16.
26 Shi K, Hung P, Wang SY. Associations Among Health Literacy, End-of-Life Care Expenditures, and Rurality. *J Rural Health* 2021;37:517–25.
27 Mayer S, Spickschen J, Stein KV, *et al*. The societal costs of chronic pain and its determinants: The case of Austria. *PLoS One* 2019;14:e0213889.
28 Haun JN, Patel NR, French DD, *et al*. Association between health literacy and medical care costs in an integrated healthcare system: a regional population based study. *BMC Health Serv Res* 2015;15:249.
29 Vann WF Jr, Divaris K, Gizlice Z, *et al*. Caregivers' health literacy and their young children's oral-health-related expenditures. *J Dent Res* 2013;92:55S–62S.
30 Gerich J, Moosbrugger R, Heigl C. Health literacy and age-related health-care utilisation: a multi-dimensional approach. *Ageing Soc* 2022;42:1538–59.
31 Rasu RS, Bawa WA, Suminski R, *et al*. Health Literacy Impact on National Healthcare Utilization and Expenditure. *Int J Health Policy Manag* 2015;4:747–55.
32 Mantwill S, Schulz PJ. Low health literacy associated with higher medication costs in patients with type 2 diabetes mellitus: Evidence from matched survey and health insurance data. *Patient Educ Couns* 2015;10:S0738-3991(15)30015-X.
33 McDougall JA, Banegas MP, Wiggins CL, *et al*. Rural Disparities in Treatment-Related Financial Hardship and Adherence to Surveillance Colonoscopy in Diverse Colorectal Cancer Survivors. *Cancer Epidemiol Biomarkers Prev* 2018;27:1275–82.
34 Hou W-H, Kuo KN, Chen M-J, *et al*. Simple scoring algorithm to identify community-dwelling older adults with limited health literacy: a cross-sectional study in Taiwan. *BMJ Open* 2021;11:e045411.
35 Li L, Li Y, Nie X, *et al*. The association between health literacy and medical out-of-pocket expenses among residents in China. *J Public Health (Berl)* 2024;32:621–9.
36 Liu L, Huang J, Li G, *et al*. The economic costs of limited health literacy in China: evidence from China's National Health Literacy Surveillance data. *BMC Health Serv Res* 2022;22:521.

37 MacLeod S, Musich S, Gulyas S, *et al*. The impact of inadequate health literacy on patient satisfaction, healthcare utilization, and expenditures among older adults. *Geriatr Nurs* 2017;38:334–41.
38 Franzen J, Mantwill S, Rapold R, *et al*. The relationship between functional health literacy and the use of the health system by diabetics in Switzerland. *Eur J Public Health* 2014;24:997–1003.
39 Ture DA, Demirci H, Sengoren Dikis O. The relationship between health literacy and disease specific costs in subjects with chronic obstructive pulmonary disease (COPD). *Aging Male* 2020;23:396–402.
40 Goodwin EJ, Zaniletti I, Solano J, *et al*. Parental Health Literacy and Acute Care Utilization in Children With Medical Complexity. *Hosp Pediatr* 2024;14:e426–31.
41 Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. *Fam Med* 2004;36:588–94.
42 Wallace LS, Rogers ES, Roskos SE, *et al*. Brief report: screening items to identify patients with limited health literacy skills. *J Gen Intern Med* 2006;21:874–7.
43 Sørensen K, Van den Broucke S, Pelikan JM, *et al*. Measuring health literacy in populations: illuminating the design and development process of the European Health Literacy Survey Questionnaire (HLS-EU-Q). *BMC Public Health* 2013;13:948.
44 Pelikan JM, Ganahl K. Measuring Health Literacy in General Populations: Primary Findings from the HLS-EU Consortium's Health Literacy Assessment Effort. *Stud Health Technol Inform* 2017;240:34–59.
45 Haun J, Noland-Dodd V, Varnes J, *et al*. Testing the BRIEF Health Literacy Screening Tool. *Fed Pract* 2009;24–31.
46 Hill JC, Kang S, Benedetto E, *et al*. Development and initial cohort validation of the Arthritis Research UK Musculoskeletal Health Questionnaire (MSK-HQ) for use across musculoskeletal care pathways. *BMJ Open* 2016;6:e012331.
47 Baker DW, Williams MV, Parker RM, *et al*. Development of a brief test to measure functional health literacy. *Patient Educ Couns* 1999;38:33–42.
48 Weiss BD, Mays MZ, Martz W, *et al*. Quick assessment of literacy in primary care: the newest vital sign [published correction appears in Ann Fam Med. *Ann Fam Med* 2006;3:514–22.
49 Lee JY, Rozier RG, Lee S-YD, *et al*. Development of a word recognition instrument to test health literacy in dentistry: the REALD-30--a brief communication. *J Public Health Dent* 2007;67:94–8.
50 Morris NS, MacLean CD, Chew LD, *et al*. The Single Item Literacy Screener: evaluation of a brief instrument to identify limited reading ability. *BMC Fam Pract* 2006;7:21.
51 Osborne RH, Batterham RW, Elsworth GR, *et al*. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). *BMC Public Health* 2013;13:658.
52 Gerich J, Moosbrugger R. Subjective Estimation of Health Literacy-What Is Measured by the HLS-EU Scale and How Is It Linked to Empowerment? *Health Commun* 2018;33:254–63.
53 Li Y. Introduction of 2012 Chinese residents health literacy monitoring program. *Chin J Health Educ* 2014;30:563–5.
54 Shen M, Hu M, Liu S, *et al*. Assessment of the Chinese Resident Health Literacy Scale in a population-based sample in South China. *BMC Public Health* 2015;15:637.
55 Timmins L, Peikes D, McCall N. Pathways to reduced emergency department and urgent care center use: Lessons from the comprehensive primary care initiative. *Health Serv Res* 2020;55:1003–12.
56 Kim K, Han HR. Potential links between health literacy and cervical cancer screening behaviors: a systematic review. *Psychooncology* 2016;25:122–30.
57 Lorini C, Santomauro F, Donzellini M, *et al*. Health literacy and vaccination: A systematic review. *Hum Vaccin Immunother* 2018;14:478–88.
58 Siena LM, Isonne C, Sciurti A, *et al*. The Association of Health Literacy with Intention to Vaccinate and Vaccination Status: A Systematic Review. *Vaccines (Basel)* 2022;10:1832.
59 Zhang NJ, Terry A, McHorney CA. Impact of health literacy on medication adherence: a systematic review and meta-analysis. *Ann Pharmacother* 2014;48:741–51.
60 Capoccia K, Odegard PS, Letassy N. Medication Adherence With Diabetes Medication: A Systematic Review of the Literature. *Diabetes Educ* 2016;42:34–71.
61 Jeganathan C, Hosseinzadeh H. The Role of Health Literacy on the Self-Management of Chronic Obstructive Pulmonary Disease: A Systematic Review. *COPD* 2020;17:318–25.
62 Tavousi M, Mohammadi S, Sadighi J, *et al*. Measuring health literacy: A systematic review and bibliometric analysis of instruments from 1993 to 2021. *PLoS One* 2022;17:e0271524.
63 Qi S, Hua F, Xu S, *et al*. Trends of global health literacy research (1995-2020): Analysis of mapping knowledge domains based on citation data mining. *PLoS One* 2021;16:e0254988.
64 Stoll CRT, Izadi S, Fowler S, *et al*. The value of a second reviewer for study selection in systematic reviews. *Res Synth Methods* 2019;10:539–45.