



Health insurance literacy assessment tools: a systematic literature review

Ana Cecilia Quiroga Gutiérrez¹

Received: 10 February 2021 / Accepted: 30 June 2021 / Published online: 3 August 2021
© The Author(s) 2021

Abstract

Aim This systematic literature review aimed to find and summarize the content and conceptual dimensions assessed by quantitative tools measuring Health Insurance Literacy (HIL) and related constructs.

Methods Using a Peer Review of Electronic Search Strategy (PRESS) and the PRISMA guideline, a systematic literature review of studies found in ERIC, Econlit, PubMed, PsycInfo, CINAHL, and Google Scholar was performed in April 2019. Measures for which psychometric properties were evaluated were classified based on the Paez et al. conceptual model for HIL and further assessed using COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) Risk of Bias checklist and criteria for good measurement properties.

Results Out of 123 original tools, only 19 were tested for psychometric and measurement properties; 18 of these 19 measures were developed and used in the context of Medicare. Four of the found measures tested for psychometric properties evaluated all four domains of HIL according to Paez et al.'s conceptual model; the rest of the measures assessed three (3), two (8), or one domain (4) of HIL.

Conclusion Most measurement tools for HIL and related constructs have been developed and used in the context of the USA health insurance system, primarily in Medicare, while there is a paucity of measurement tools for private health insurances and from other countries. Furthermore, there is a lack of conceptual consistency in the way HIL is measured. Standardization of HIL measurement is crucial to further understand the role and interactions of HIL in other insurance contexts.

Keywords Health insurance literacy · Health insurance · Health insurance decision making · Health insurance education · Health literacy

Introduction

The selection of a less-than-optimal plan for a specific individual, or the complete lack of coverage, can have severe financial and health consequences (Bhargava and Loewenstein 2015; Bhargava et al. 2015, 2017; Flores et al. 2017). Simultaneously, health insurance represents one of the most complex and costly products individuals consume (Paez et al. 2014). A survey commissioned in 2008 by eHealth, Inc., an online marketplace for health insurance in the USA, found

that most consumers lack a basic understanding of health insurance terms. Similarly, consumer testing in the USA found that low numeracy and confusion regarding cost-sharing terms can hinder optimal selection of a health plan (Quincy 2012a). A more recent study using data collected from nonelderly adults in the June 2014 wave of the Health Reform Monitoring Survey (HRMS) found that both literacy and numeracy were more likely to be lower for uninsured adults compared to insured ones, especially for those with lower income and eligibility for subsidized coverage. The authors concluded that “gaps in literacy and numeracy among the uninsured will likely make navigating the health care system difficult” (Long et al. 2014). Another study from the USA, found a negative association between insurance comprehension and odds of choosing a plan that was at least \$500 more expensive annually compared to the cheapest option, concluding that both health insurance comprehension and numeracy were critical skills in choosing

✉ Ana Cecilia Quiroga Gutiérrez
ana.quiroga@unilu.ch

¹ Department of Health Sciences and Medicine, Center for Health, Policy and Economics, University of Lucerne, Switzerland, Swiss Learning Health System, Frohburgstrasse 3, PO Box 4466, CH-6002 Lucerne, Switzerland

a health insurance plan that provides adequate risk protection (Barnes et al. 2015).

In light of these findings, and the introduction of the Affordable Care Act in the USA, the concept of health insurance literacy (HIL) has gained increasing attention over the past decade.

In 2009, McCormack and colleagues developed the first framework and instrument to measure HIL by integrating insights from the fields of health literacy and financial literacy. The instrument was originally developed to measure HIL of older adults in the USA. The study found that, in line with previous research, older adults, and those with lower education and income had lower levels of HIL (McCormack et al. 2009). In 2011, a roundtable sponsored by the Consumers Union in the USA that included experts from academia, advocacy, health insurance, and private research firms defined HIL as “the degree to which individuals have the knowledge, ability, and confidence to find and evaluate information about health plans, select the best plan for their own (or their family’s) financial and health circumstances, and use the plan once enrolled” (Quincy 2012b).

In 2014, the American Institute of Research in the USA developed a conceptual model for HIL and the Health Insurance Literacy Measure (HILM) (Paez et al. 2014). The model was developed using information collected from: a literature review, key informant interviews, and a stakeholder group discussion. The conceptual model consists of four domains of HIL: health insurance knowledge, information seeking, document literacy, and cognitive skills, with self-efficacy as an underlying domain. In contrast to the previously mentioned framework of McCormack and colleagues (McCormack et al. 2009), which focused on the Medicare population, Paez et al.’s conceptual model was developed and used in the context of private health insurance.

Even though research on HIL is still in its early stages, studies show that HIL may influence how individuals use and choose health insurance services (Loewenstein et al. 2013; Barnes et al. 2015), seek health insurance information (Tennyson 2011), and use health services (Edward et al. 2018; Tipirneni et al. 2018; James et al. 2018). So far, most research addressing HIL has taken place in the USA, with relatively little literature on the topic coming from other countries. Furthermore, early literature focused mainly on consumer’s understanding of their health insurance (Lambert 1980; Marquis 1983; Cafferata 1984; McCall et al. 1986; Garnick et al. 1993) resulting in the use of inconsistent terminology.

Given the still recent definition of HIL and inconsistency of its assessment at the international level, this review includes HIL and related constructs such as health insurance knowledge, understanding, familiarity, comprehension, and numeracy, with the goal of identifying assessment tools that cover the domains on Paez et al.’s conceptual model for HIL.

Specifically, this review aims to (1) identify and summarize the content of tools that aim to assess HIL and related constructs, (2) describe conceptual dimensions assessed by psychometrically tested measures and when possible, (3) briefly discuss the methodological quality of measurement development and psychometric properties the found tools.

Methods

After consulting with a librarian and performing a Peer Review of Electronic Search Strategy (PRESS)(McGowan et al. 2016), a systematic literature search was conducted in April 2019 to identify studies using quantitative and mixed methods tools aiming to assess HIL and related constructs. In order to identify measures assessing HIL and related constructs, search terms that were used included constructs that are described in the HIL frameworks previously introduced (McCormack et al. 2009; Paez et al. 2014). Examples of search terms include “health insurance literacy,” “health insurance,” “health plan,” “Medicare,” “knowledge,” “understand*,” or “instrument” (a full list of search terms and PRESS derived search strings used can be found on Online Resource 1). The following databases were searched: ERIC, Econlit, PubMed, PsycInfo, CINAHL, and Google Scholar. Screening of titles and abstracts, full texts; and data extraction were completed by two independent reviewers. Disagreements between reviewers were resolved by discussion until consensus was reached, and a third researcher was consulted when no consensus was attained.

The literature search was complemented through environmental scan and reference-harvesting, which involved reaching out to researchers for gray literature and finding relevant references from bibliographies of included studies, respectively. Screening, data extraction, and PRISMA diagram generation was completed using Distiller SR (DistillerSR n.a.). Statistical software Stata 16 was used to synthesize information and generate figures (StataCorp LLC 2019).

The review is in accordance with the PRISMA reporting guidelines for systematic reviews (Moher et al. 2009). The PRISMA diagram (Fig. 1) illustrates the number of articles found through different sources, those eliminated due to deduplication or exclusion, and the final count of studies included in the review.

Inclusion and exclusion criteria

The main inclusion and exclusion criteria are summarized in Table 1 using the PICoS mnemonic (Population, phenomenon of Interest, Context, Study design) (Lockwood et al. 2015).

Studies with any target population were included, except for those limited to health professionals, health insurance

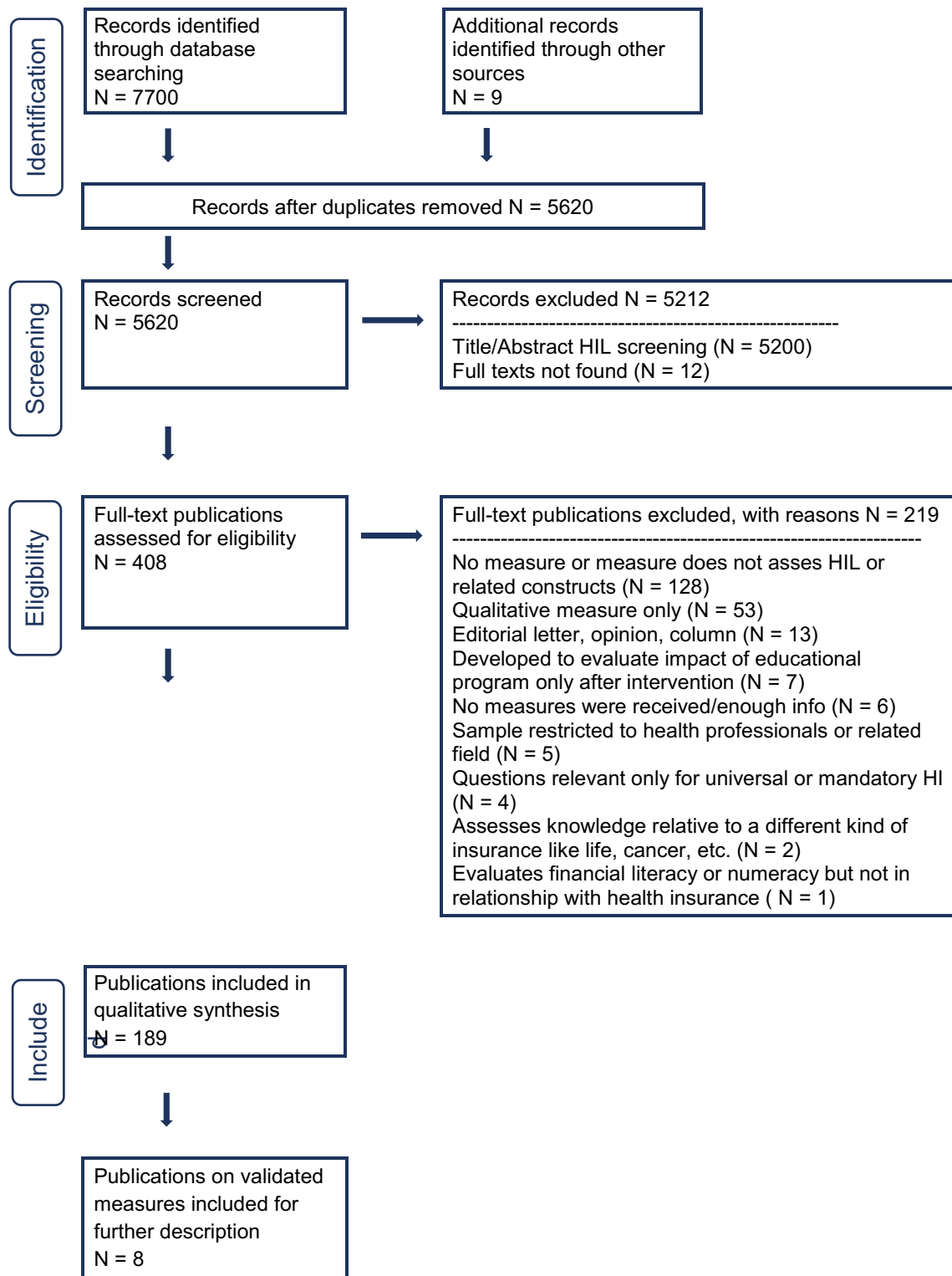


Fig. 1 PRISMA diagram

experts, or students of disciplines such as medicine, pharmacy, and nursing.

To be included, studies had to describe in sufficient detail the measure that was used to assess HIL or a related construct. Sufficient description included, but was not limited to,

whether the used measure was subjective (i.e., self-reported) or objective; number of items, HIL domains assessed by the measure according to the Paez. et al. conceptual model (Paez et al. 2014), and country. The review was limited to quantitative and mixed methods measures. Purely qualitative

Table 1 PICoS based inclusion and exclusion criteria

P – Population	Any target population except health professionals, health insurance experts, or students of disciplines such as medicine, pharmacy, and nursing.
I – Interest	Health Insurance Literacy and related constructs including: familiarity, comprehension, understanding of health insurance.
Co – Context	Any context except when questions refer only to awareness of universal or mandatory health insurance. Instruments asking about other aspects regarding universal or mandatory health insurance, such as utilization, services provided, coverage, etc., are included.
S – Study type	Quantitative and mixed methods studies. Purely qualitative measures are excluded.

measures were excluded. The search was not restricted to any time period or country. Inclusion of publications was limited to those published in English.

Cross-sectional, cohort and case-control studies were included. Studies which described an intervention to improve HIL or related constructs were included only if baseline levels were assessed.

In the case of studies in which a measure was used or mentioned but was not thoroughly described, authors were contacted. Those studies for which no further or sufficient information could be obtained were excluded.

Data extraction

Found measurement tools were categorized according to characteristics, such as originality of the measure: if it was a previously developed measure or if it was a mixed measure, meaning the measure had been previously developed but was supplemented with additional items or items were adapted from previously used measures. Further characteristics included the number of items, whether the measure was quantitative or used a mixed method such as coding of answers to open questions; if measures were objective and/or subjective; design of the study in which the measurement tool was used, e.g., surveys, interviews, focus groups, etc.

Measures for which psychometric properties were evaluated, were categorized using Paez. et al.'s conceptual model (Paez et al. 2014), by domains and number of domains assessed. This conceptual model was selected for several reasons. First, it was developed with purchasers of private health insurance in mind, rather than being focused on Medicare, which is the case of the model from McCormack et al. (Tseng et al. 2009). Second, it focuses on domains and domain-specific tasks that make up the concept of HIL, rather than showcasing factors that may explain different levels of HIL (Vardell 2017). Finally, the conceptual model by Paez et al. (2014) is the most recent attempt at conceptualizing and operationalizing the concept of HIL. Measures for which insufficient information was provided regarding the assessment of measurement properties were not included in this section.

Psychometrically tested measures were further assessed for methodological quality using the CONsensus-based Standards for the selection of health Measurement Instruments (COSMIN) Risk of Bias checklist (Terwee et al. 2018; Mokkink et al. 2018); a modular tool which provides standards to evaluate the methodological quality of studies on outcome measures. The checklist allows the assessment of the methodological quality of measure development, including content validity, structural validity, internal consistency, reliability, measurement error, criterion validity, hypotheses testing for construct validity, and responsiveness; for definitions of these measurement properties please refer to Online Resource 2.

The methodological quality of the studies on outcome measures was rated as very good, adequate, doubtful, or inadequate quality; according to COSMIN standards, only when sufficient information was available to do so. Overall rating for methodological quality was given by selecting the lowest or “worst score counts” rating for each of the domains assessed per COSMIN guidelines. Because there is no established gold-standard to assess HIL, and none of the measures were tested for cross-cultural validity, criterion validity and cross-cultural validity were not assessed. The levels of evidence of the overall study quality for measurement properties was determined using the GRADE approach specified in the COSMIN manual. The quality of the evidence was determined using a four-scale grading: high, moderate, low, and very low. Quality of measurement properties was assessed according to the COSMIN criteria for good measurement properties when possible (Mokkink et al. 2018).

Given that the main focus of this systematic literature review is the content and conceptual comparison of the found psychometrically tested measures, the methodological quality of the studies and quality of measurement properties are only briefly discussed.

Results

The PRISMA flow chart (Fig. 1) summarizes the results of the search process. The search yielded a total of 7700 publications

and 9 more were found through hand search, resulting in 7709 items. After deduplication, the abstracts of the remaining 5620 unique records were screened out of which 408 were included for full-text screening. After full-text screening, a total of 189 publications were selected for inclusion in the qualitative analysis and synthesis.

Out of these 189, 108 used original measures, 67 used previously developed instruments, and 14 combined previously developed measures with new items; 171 of the studies used quantitative methods while 18 used mixed methods, see for example De Gagne et al. (2015) and Desselle (2003).

The frequency of papers published related to the topic of HIL and similar constructs has increased over time, especially since 2015 (Fig. 2). Over 80% of these publications are from the USA (151), while the rest comes from other countries (see Fig. 3 for all countries).

Characteristics of measures

One hundred twenty-three original measures for HIL or related constructs were found. Out of these, 109 were quantitative (88.6%). When categorized based on the number of domains assessed according to the Paez et al. conceptual model, approximately 66% (72) of the original quantitative measures assessed only one domain, 19.3% (21) measured two domains, 8.3% (9) three, and 6.4% (7) assessed all four domains.

Out of the original quantitative measures assessing only one domain, most measures (68) evaluated knowledge.

The most used measures and questions were derived from the Medicare Current Beneficiary Survey (MCBS) (24), Health Insurance Literacy Measure (HILM) (13), Kaiser Family Foundation quiz (8), United States Household Health Reform Monitoring Survey (5), and S-TOFHLA for functional health literacy (5).

Psychometrically tested measures

Nineteen of the measures found, were evaluated for psychometric properties. These assessed at least one HIL domain according to the Paez et al. conceptual model (2014) and presented information on development and measurement properties. Eighteen of these measures were developed and used to evaluate Medicare beneficiaries’ knowledge in the USA. The most recent psychometrically tested measure found, the Health Insurance Literacy Measure (HILM), is the only one attempting to assess consumers’ ability to select and use private health insurance.

A summary of the measures for which psychometric properties were evaluated can be found in Table 2. Table 3 describes the HIL domains assessed with these measures according to the Paez et al. conceptual model (2014). Table 4 summarizes the evaluation of the measurement properties quality.

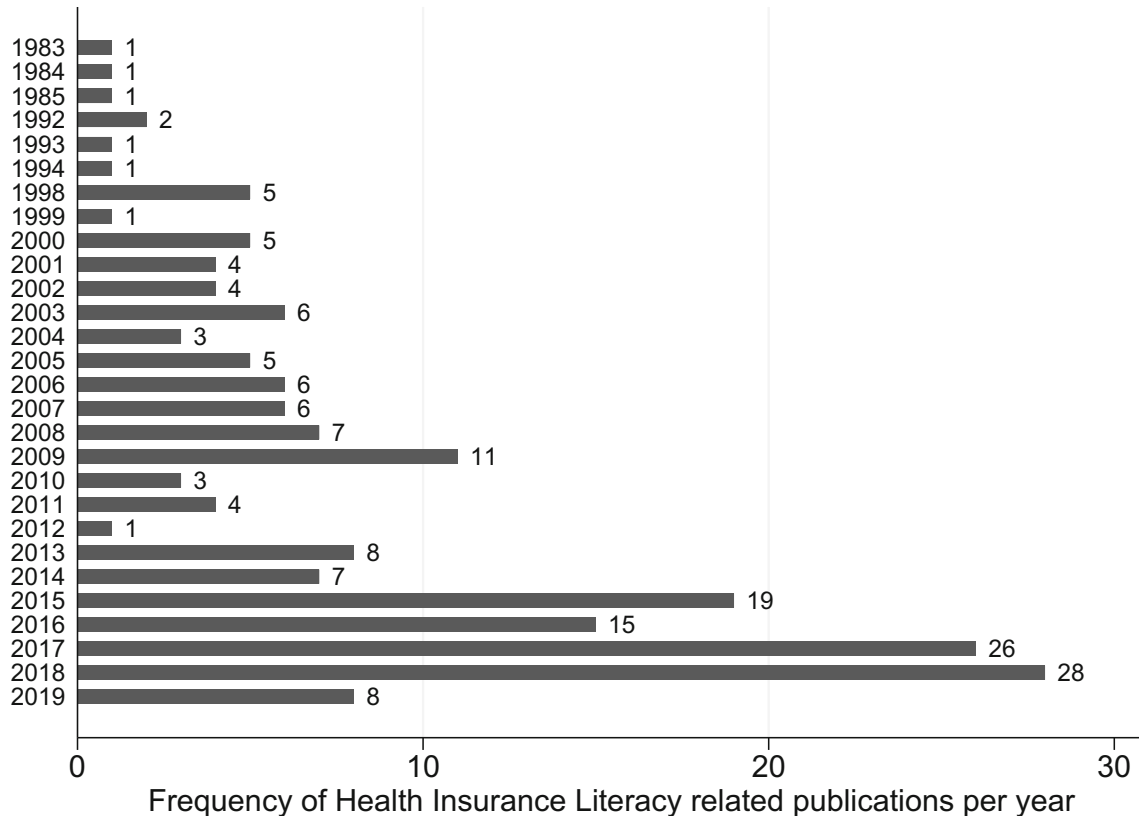


Fig. 2 Frequency of publications per year

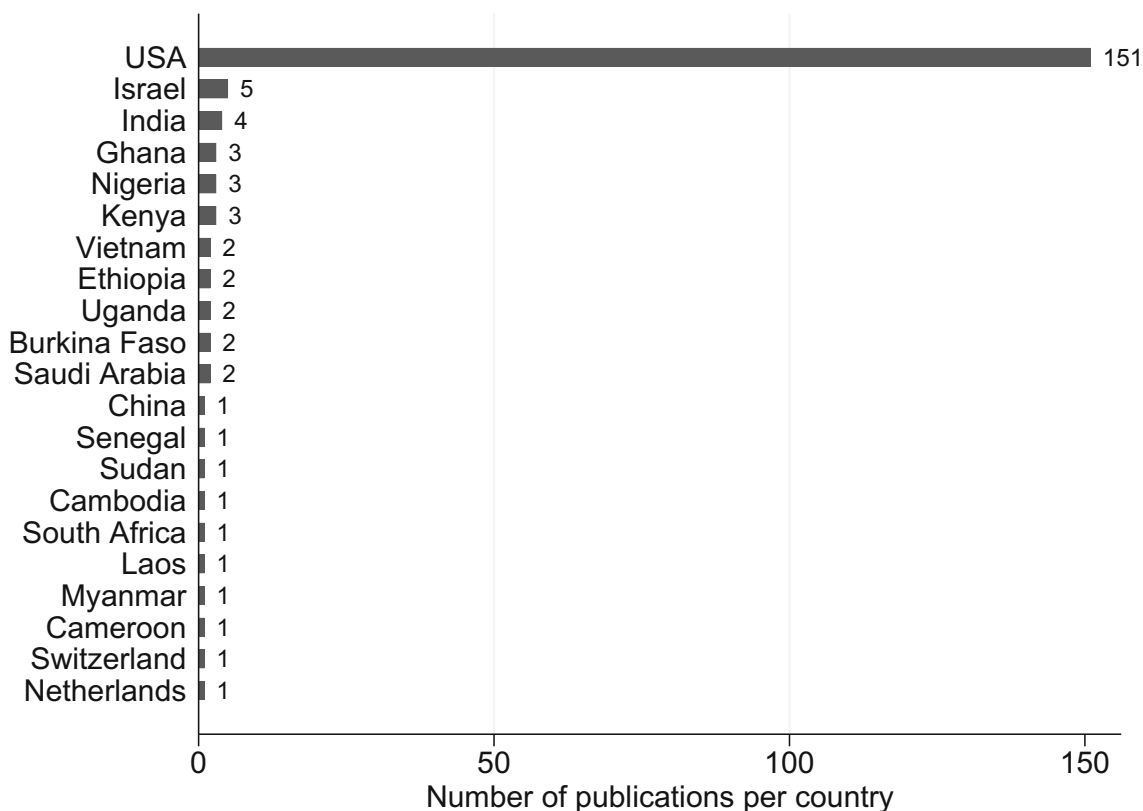


Fig. 3 Frequency of publications per country

Information on the methodological quality of measurement development, methodological quality of studies assessing measurement properties, and measurement properties of the found psychometrically tested measures can be found in Online Resources 3, 4, and 5, respectively. Level of evidence was moderate for all the following measures, which are listed in Table 2, as only one study of adequate quality is available for most of them. Quality of measurement development was deemed inadequate for all measures as there was no information on the measurement development, or it lacked the inclusion of a target population sample.

As measures developed in the context of Medicare and specifically the MCBS represent the most widely used in the found studies, these will be discussed in the following chronologically in the order in which they were used, followed by the assessment of the HILM.

1995–1998 MCBS measures

Six measures were created by using existing items from different waves of the MCBS from 1995 to 1998 (Bann et al. 2000). The aim of these measures was to evaluate the impact of the National Medicare Education Program (NMEP) project, an initiative to develop informational resources for Medicare.

The first two of these measures were single item questions that assessed the understandability of the Medicare program,

called the “Medicare Understandability Question” and the “Global Know-All-Need-To-Know Question.”

The “Know-All-Need-To-Know Index” consisted of five questions included in two rounds of the MCBS, in 1996 and 1998. It assessed how much individuals felt they knew about different aspects of Medicare, Medigap, benefits, health expenses, and finding and choosing health providers.

An eight-item true/false quiz, referred to as the “Eight-item Quiz” consisted of items that were included in one round in 1998. Items aimed to assess knowledge on Medicare options and managed care plans through true/false/not sure statements.

Similarly, a “four-item quiz,” and a “three-item quiz” were also generated by including true/false MCBS items used in 1998.

Results on measurement properties for the above described measures were presented in two papers (Bann et al. 2000; Bonito et al. 2000), which showed that the methodological quality varied between very good and adequate for structural validity, internal consistency, reliability, measurement error, construct validity, and responsiveness. Internal consistency for the “Know-All-Need-To-Know index” and the “Eight-item quiz” was determined as sufficient, while the “Four-item quiz” and “Three-item quiz” was insufficient. Content validity and structural validity were indeterminate for all measures

Table 2 Summary of psychometrically tested measures

HIL assessment tool (authors, year)	Construct(s)	Number of HIL domains assessed*	Target population	Context of Use	Mode of administration (e.g., self-report, interview-based, parent/proxy report etc)	Objective/ Subjective	Number of items ^a	Response options	Sample size
Medicare understandability question (Bann et al. 2000)	Medicare knowledge	1	US Medicare beneficiaries	MCBS	Computer-assisted personal interviewing	Subjective	1	Yes/No/Do not know	39,357
Global know-all-need-to-know question (C. Bann et al. 2000)	Medicare knowledge	1	US Medicare beneficiaries	MCBS	Computer-assisted personal interviewing	Subjective	1	Five-point scale from “Almost none of what I need to know” to “Just about everything I need to know”	14,438
Know-all-need-to-know index (C. Bann et al. 2000)	Medicare knowledge	2	US Medicare beneficiaries	MCBS	Computer-assisted personal interviewing	Subjective	5	Five-point scale from “Almost none of what I need to know” to “Just about everything I need to know”	26,409
Four-item quiz (C. Bann et al. 2000)	Medicare knowledge	1	US Medicare beneficiaries	MCBS	Computer-assisted personal interviewing	Objective	4	True/False	12,643
Three-item quiz (C. Bann et al. 2000)	Medicare knowledge	1	US Medicare beneficiaries	MCBS	Computer-assisted personal interviewing	Objective	3	True/False	13,964
Eight-item quiz (C. Bann et al. 2000)	Medicare knowledge	2	US Medicare beneficiaries	MCBS	Computer-assisted personal interviewing	Objective	8	True/False	14,532
Kansas city index (Bonito et al. 2000; McCormack et al. 2002)	Medicare knowledge	2	Medicare beneficiaries in the 10-county Kansas City metropolitan statistical area	Medicare	Computer-assisted telephone interview	Objective	15	Yes/No, Multiple-choice	2107
National Evaluation index (Bonito et al. 2000; McCormack and Uhrig 2003)	Medicare knowledge	2	US Medicare beneficiaries	Medicare	Mail survey	Objective/Subjective	22	True/False/Do not know, Yes/No/Do not know, multiple-choice	6154
2002 99-item questionnaire (MCBS) (Uhrig et al. 2002)	Knowledge of the Medicare program	4	US Medicare beneficiaries	MCBS	Computer-assisted personal interviewing	Objective/Subjective	99	True/False/Do not know, Yes/No/Do not know, multiple-choice	2634
Knowledge Quiz – Form A (Bann and McCormack 2005)	Medicare knowledge	2	US Medicare beneficiaries	MCBS	Computer-assisted personal interviewing	Objective	11	True/False	2634

Table 2 (continued)

HIL assessment tool (authors, year)	Construct(s)	Number of HIL domains assessed*	Target population	Context of Use	Mode of administration (e.g., self-report, interview-based, parent/proxy report etc)	Objective/ Subjective	Number of items ^a	Response options	Sample size
Knowledge Quiz – Form B (Bann and McCormack 2005)	Medicare knowledge	2	US Medicare beneficiaries	MCBS	Computer-assisted personal interviewing	Objective	11	True/False	2634
Knowledge Quiz – Form C (Bann and McCormack 2005)	Medicare knowledge	2	US Medicare beneficiaries	MCBS	Computer-assisted personal interviewing	Objective	11	True/False	2634
Health Literacy Quiz – Form A (Bann and McCormack 2005)	Health Literacy (insurance focused)	3	US Medicare beneficiaries	MCBS	Computer-assisted personal interviewing	Objective	11	True/False, Yes/No, Multiple choice	2634
Health Literacy Quiz – Form B (Bann and McCormack 2005)	Health Literacy (insurance focused)	4	US Medicare beneficiaries	MCBS	Computer-assisted personal interviewing	Objective	11	Multiple choice	2634
Health Literacy Quiz, Form C (Bann and McCormack 2005)	Health Literacy (insurance focused)	3	US Medicare beneficiaries	MCBS	Computer-assisted personal interviewing	Objective	11	True/False, Yes/No, Multiple choice	2634
PKI (Bann et al. 2003)	Medicare knowledge	3	US Medicare beneficiaries	MCBS	Computer-assisted personal interviewing	Subjective	5	Five-point scale from “almost none of what you need to know” to “just about everything you need to know.”	25,130
Seven-item quiz (Bann et al. 2003)	Medicare Knowledge	2	US Medicare beneficiaries	MCBS	Computer-assisted personal interviewing	Objective	7	True/False	16,982
Health Insurance Literacy items (MCBS) (McCormack et al. 2009)	Health Insurance Literacy	4	US Medicare beneficiaries	MCBS	Computer-assisted personal interviewing	Objective	27	True/False, Multiple-choice	1202
HILM (Paez et al. 2014)	Health Insurance Literacy	4	US adults 22 through 64, who were uninsured, enrolled in private insurance, or Medicaid	Private health insurance	Self-assessment survey	Subjective	21	Multiple-choice (level of confidence)	828

HIL domains based on the Health Insurance Literacy conceptual model by Paez et al. (2014)

Table 3 HIL domains assessed by psychometrically tested measures according to Paez et al. (2014)

HIL assessment tool	Knowledge	Information seeking	Document literacy	Cognitive skills
Medicare understandability question (Bann et al. 2000)	X			
Global know-all-need-to-know question (Bann et al. 2000)	X			
Know-all-need-to-know index (Bann et al. 2000)	X	X		
Four-item quiz (Bann et al. 2000)	X			
Three-item quiz (Bann et al. 2000)	X			
Eight-item quiz (Bann et al. 2000)	X	X		
Kansas city index (Bonito et al. 2000; McCormack et al. 2002)	X	X		
National Evaluation index (Bonito et al. 2000; McCormack and Uhrig 2003)	X	X		
2002 99-item questionnaire (MCBS) (Uhrig et al. 2002)	X	X	X	X
Knowledge Quiz – Form A (Bann and McCormack 2005)	X	X		
Knowledge Quiz – Form B (Bann and McCormack 2005)	X	X		
Knowledge Quiz – Form C (Bann and McCormack 2005)	X	X		
Health Literacy Quiz – Form A (Bann and McCormack 2005)	X	X	X	
Health Literacy Quiz – Form B (Bann and McCormack 2005)	X	X	X	X
Health Literacy Quiz, Form C (Bann and McCormack 2005)	X	X	X	
PKI (Bann et al. 2003)	X	X		X
Seven-item quiz (Bann et al. 2003)	X	X		
Health Insurance Literacy items (MCBS) (McCormack et al. 2009)	X	X	X	X
HILM (Paez et al. 2014)	X	X	X	X

included in this set. Hypothesis testing was sufficient as well as responsiveness for all measures.

Kansas City index and national evaluation index

The “Kansas City Index” and the “National Evaluation Index” were measures designed to evaluate the impact of Medicare information material on beneficiaries’ knowledge, all of the questions could be answered by consulting the Medicare & You handbook (Medicare and You n.a.).

The “Kansas City Index” (Bonito et al. 2000; McCormack et al. 2002) is a 15-item measure that was used to evaluate the impact of different interventions such as the distribution of the Medicare & You handbook, and the Medicare & You bulletin on Medicare beneficiary knowledge in the Kansas City metropolitan area.

The “National Evaluation Index” (Bonito et al. 2000; McCormack and Uhrig 2003) is a 22-item index that reflects Medicare related knowledge in seven different content areas: awareness of Medicare options, access to original Medicare, cost implications of insurance choices, coverage/benefits, plan

Table 4 Quality of measurement properties according to COSMIN criteria

HIL measure	Internal consistency	Content validity	Construct validity			Responsiveness
			Structural validity	Hypothesis testing	Cross cultural validity	
Medicare Understandability Question	NA	?	?	+	NA	+
Global Know-All-Need-To-Know Question	NA	?	?	+	NA	+
Know-All-Need-To-Know Index	+	?	?	+	NA	+
Four-Item Quiz	–	?	?	+	NA	+
Three-Item Quiz	–	?	?	+	NA	+
Eight-Item Quiz	+	?	?	+	NA	+
Kansas city index	+	?	?	+	NA	+
National Evaluation Index	+	?	?	+	NA	+
2002 99-item questionnaire (MCBS)	?	±	+	+	NA	+
Knowledge Quiz - Form A	–	?	?	+		+
Knowledge Quiz - Form B	–	?	?	+		+
Knowledge Quiz - Form C	–	?	?	+		+
Health Literacy Quiz - Form A	–	?	?	+	NA	+
Health Literacy Quiz - Form B	–	?	?	+	NA	+
Health Literacy Quiz - Form C	–	?	?	+	NA	+
PKI	+	?	+	+	NA	+
Seven-item Quiz	+	?	+	+	NA	+
Health Insurance Literacy Items (MCBS)	+	?	–	+	NA	+
HILM	+	±	–	+	NA	+

Assessment of quality of measurement properties based on COSMIN criteria for good measurement properties (Terwee et al. 2018; Mokkink et al. 2018). Each result is rated either as (+) sufficient, (–) insufficient, (?) indeterminate or (NA) not applicable

rules/restrictions, availability of information, and beneficiary rights.

Methodological quality for the assessment of measurement properties was either very good or adequate for structural validity, internal consistency, reliability, construct validity, and responsiveness. Methodological quality for the assessment of measurement error was considered doubtful both for the “Kansas City Index” and the “National City Index” since the time period between responses was approximately two years (Bonito et al. 2000), as it is recommended to be measured over shorter intervals, especially to assess test-retest reliability (Terwee et al. 2007).

Regarding the quality of measurement properties, both the “Kansas City Index” and the “National Evaluation Index” met sufficient criteria for internal consistency, hypothesis testing, and responsiveness. Quality of content validity and structural validity was indeterminate.

2002 Questionnaire, knowledge quizzes, and health literacy quizzes

Uhrig et al. (Uhrig et al. 2002) developed a questionnaire composed of 99 questions based on recommendation from

Bonito et al. (Bonito et al. 2000) to develop a knowledge index using Item Response Theory (IRT), which would allow measuring and tracking Medicare beneficiaries’ knowledge over time. Questions were cognitively tested and calibrated using IRT to develop six alternative forms of quizzes: three on Medicare knowledge and three on health literacy specific for Medicare beneficiaries focused on insurance-related terminology and scenarios (Bann and McCormack 2005). These six measures were generated to demonstrate how the different items in the previously generated questionnaire could be used to create dynamic quizzes to be integrated in the MCBS, using different questions throughout the waves of the survey, yet providing comparable results for longitudinal studies and allowing quizzes to be updated when items became irrelevant or there were policy changes.

The development of the questionnaire items described by Uhrig et al. (Uhrig et al. 2002) was based on a background research, review of existing Medicare informational materials and knowledge surveys, and discussions with experts in the field. Questions were generated and selected to cover five knowledge areas: eligibility for and structure of Original Medicare, Medicare+Choice (an alternative model of Medicare), plan choices and health plan decision-making,

information and assistance, and Medigap/Employer-sponsored supplemental insurance. Four additional question categories were included: self-reported knowledge, health literacy, cognitive abilities, and other non-knowledge items. The questionnaire was pilot tested within the MCBS and the resulting data was used to evaluate the psychometric properties of the items and to calibrate them using IRT. After completing item calibration, the authors generated the three alternative forms of “knowledge” quizzes and three alternative forms of “health literacy” quizzes (Bann and McCormack 2005). Items were selected by the authors so that each of the forms would contain at least one item from each of the identified content areas, items with similar total correct percentages, high slopes, a variety of difficulty levels, as well as items that seem relevant from a policy standpoint. The three knowledge measures contained one Original Medicare section and one Medicare+Choice section, covering both factors in each of the forms. Similarly, the three “Health Literacy” measures contained sections for two different factors, a terminology, and a reading comprehension one.

Methodological quality on measurement properties for the 99-item questionnaire is adequate for structural validity. Methodology for structural validity, construct validity, comparison with other instruments, and comparison between subgroups was adequate for all quizzes, while quality for Internal consistency, reliability, and measurement error was very good.

In regard to measurement properties, content validity was inconsistent for the 99-item questionnaire, while it met the sufficient criteria for structural validity, hypothesis testing, and responsiveness. The “Knowledge” and “Health Literacy” measures forms A, B, and C met the sufficient criteria for hypothesis testing and responsiveness, while internal consistency was insufficient, and content validity and structural validity was indeterminate.

Perceived knowledge index (PKI) and seven-item quiz (2003)

In a paper published in 2003, Bann et al. used data from rounds of the MCBS 1998 and 1999 to develop two different knowledge indices to measure knowledge of Medicare beneficiaries. The first of these indices is the “Perceived Knowledge Index” (PKI), a five-item measure constructed from questions included in two different rounds of the MCBS. The questions assessed how much beneficiaries felt they knew about different aspects of Medicare. The second index was a seven-item quiz, made up of questions also included in two different rounds of the MCBS, which used true/false questions to test objective knowledge of Medicare (Bann et al. 2003).

For both the PKI and the “Seven-item quiz” the methodological quality of the following were adequate: analysis of structural validity, internal consistency, known-groups

validity, and comparison between subgroups was very good, while also having convergent validity and comparison with other instruments.

Both PKI and the “Seven-item quiz” met sufficient criteria for internal consistency, hypothesis testing, and responsiveness. Content validity was indeterminate.

HIL framework and health insurance literacy items (MCBS) (2009)

In 2009, McCormack et al. published a paper in which they described the development of a framework and measurement instrument for HIL in the context of Medicare (McCormack et al. 2009). The framework was built based on a literature review and additional key studies. It integrates consumer and health care system variables that would be associated with HIL and the navigation of the health care and health insurance systems. Items were developed to operationalize the framework and cognitively tested and eventually fielded in the MCBS national survey.

Methodological quality for the analysis of structural validity, internal consistency, known-groups validity, and comparison between subgroups was considered to be very good. Reliability, convergent validity, and comparison with other instruments had adequate methodological quality.

The measure met sufficient measurement properties criteria for internal consistency, hypothesis and responsiveness was met. Structural validity was insufficient and content validity was indeterminate.

Health insurance literacy measure (HILM) (2014)

The Health Insurance Literacy Measure (HILM), a self-assessment measure that was constructed based on formative research and stakeholder guidance. The conceptual basis of the measure is the Health Insurance Literacy conceptual model by Paez et al. (Paez et al. 2014), which was previously described.

The HILM was developed using a four-stage process. First a conceptual model for HIL was constructed. Once the conceptual model was finalized, two pools of items were created to operationalize its domains. Stage three consisted of cognitive testing of the items. Finally, the last stage involved field-testing of the HILM to develop scales and establish its validity.

When assessing methodological quality of the study on measurement properties, comprehensibility for the HILM was set to doubtful while relevance was adequate. Structural validity, internal consistency, reliability, measurement error, known groups validity, comparison between subgroups and comparison before and after intervention were very good. Convergent validity and comparison with other instruments were inadequate.

Concerning measurement properties, evidence for content validity was found to be inconsistent. Criteria for structural validity was insufficient, while sufficient criteria for internal consistency, hypothesis testing, and responsiveness was met.

Discussion

This review demonstrates the wide variety of instruments that have been used to assess HIL and related constructs. While the increasing body of evidence around HIL and related constructs, such as health insurance knowledge or understanding has provided valuable insights into the topic, further steps are needed to improve the quality and value of gathered data and reduce waste in research (Ioannidis et al. 2014).

Over 80% of the studies that were found through this review were carried out in the USA. Reforms and expansion of social health programs in the USA such as Medicare have played an important role in the growing interest in HIL. This is reflected in the fact that most of the instruments for which psychometric properties were evaluated identified through this review have been developed and used in the context of Medicare, which consists of a very specific population exclusive to the USA and its health insurance system. Only one of the psychometrically tested instruments, as well as its underlying conceptual model (Paez et al. 2014) was developed and used in the context of private health insurance. However, it was still exclusive to the USA health insurance system and its population.

Furthermore, the domains and aspects assessed by these measures focus mostly on knowledge. This is the only domain assessed by all found psychometrically tested measures. Out of the 19 found psychometrically tested measures, four assessed all domains of HIL according to the conceptual model of Paez et al. (2014), while 12 of them evaluate only one or two domains. Most of these measures not only ignore important aspects that are associated with the way people navigate the health insurance system and make health insurance and healthcare decisions, but they also make it difficult to compare results across studies and populations.

None of the measures reviewed included a sample of the target population in the development phase, which is important to evaluate comprehensiveness. As a result, the quality of measurement development was rated inadequate. Similarly, while assessment of reliability and measurement error was mentioned by some of the studies, little information was available to determine the quality of measurement properties for each individual measure.

A valid and reliable instrument tool to assess HIL may not only help to accurately measure an individual's or population's HIL level but can also provide ways of identifying vulnerable groups and guide the development and implementation of effective interventions to

facilitate health system navigation, health insurance utilization and access. For example, deeper understanding of HIL could inform health insurance design and choice architecture to facilitate optimal health insurance selection (Barnes et al. 2019).

Given that existing HIL measurement tools are context-specific, further steps may require adaptation of current definitions of HIL, its conceptualization and operationalization according to specific health and health insurance systems. The McCormack et al. 2009 framework for HIL and the Paez et al. (2014) conceptual model for HIL might provide a foundation for measurement development and research in other contexts but should be adapted and instruments should be tested for validity.

Also, there are currently not enough standardized measures that would allow assessing HIL across different contexts and health insurance systems. The translation and cultural validation of the HILM could represent a viable solution to further research on specific health insurance systems. However, some of its limitations are that it is a self-administered and self-reported instrument, which may bring respondent-bias and provide a subjective perception of one's own HIL rather than an objective assessment of it.

There are some limitations to this systematic literature review. First, the search was restricted to studies or papers published in English that used quantitative or mixed methods instruments to assess HIL and related constructs. Therefore, some relevant instruments may have been missed. Second, even though the COSMIN checklist and guideline for systematic reviews is a valuable tool for the evaluation and critical appraisal of outcome measures, it was originally developed to assess the methodological quality of health outcome measures, and may not be ideal for evaluating instruments assessing HIL.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10389-021-01634-7>.

Acknowledgements The author would like to thank in particular Prof. Dr. Stefan Boes, Dr. Sarah Mantwill, Aljoscha Benjamin Hwang, Daniella Majakari, and Tess Bardy for their contribution to this review.

Authors' contributions Ana Cecilia Quiroga Gutierrez contributed to the study conception and design. Material preparation, screening of titles and abstracts, screening of full texts, data collection, synthesis and analysis were performed by Ana Cecilia Quiroga Gutierrez. A second independent reviewer (non-author), Daniella Majakari performed screening of titles and abstracts as well as full texts. A third independent reviewer (non-author), Tess Bardy, performed data collection. Peer Review Search Strategy was conducted by Aljoscha Benjamin Hwang (non-author). Work was critically revised by Sarah Mantwill (non-author) and Prof. Dr. Stefan Boes (non-author).

Funding Open Access funding provided by Universität Luzern.

Data availability Data resulting from the systematic review can be requested from the corresponding author.

Code availability Code for figure generation can be requested to the corresponding author.

Declarations

Conflict of interest The author has no relevant financial or non-financial interests to disclose.

Ethics approval Not applicable. Given that this study is a systematic literature review, it involved no human subjects or animals.

Consent to participate Not applicable.

Consent for publication Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bann C, McCormack L (2005) Measuring knowledge and health literacy among medicare beneficiaries. <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Reports/Research-Reports-Items/CMS062191.html>
- Bann C, Lissy K, Keller S, et al (2000) Analysis of medicare beneficiary baseline knowledge data from the medicare current beneficiary survey knowledge index technical note. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Reports/downloads/berkman_2003_7.pdf
- Bann CM, Terrell SA, McCormack LA, Berkman ND (2003) Measuring beneficiary knowledge of the Medicare program: a psychometric analysis. *Health Care Financ Rev* 24:111–125
- Barnes A, Hanoch Y, Rice T (2015) Determinants of coverage decisions in health insurance marketplaces: consumers' decision-making abilities and the amount of information in their choice environment. *Health Serv Res* 50:58–80. <https://doi.org/10.1111/1475-6773.12181>
- Barnes AJ, Karpman M, Long SK et al (2019) More intelligent designs: comparing the effectiveness of choice architectures in US health insurance marketplaces. *Organ Behav Hum Decis Process*. <https://doi.org/10.1016/j.obhdp.2019.02.002>
- Bhargava S, Loewenstein G (2015) Choosing a health insurance plan: complexity and consequences. *JAMA* 314:2505–2506. <https://doi.org/10.1001/jama.2015.15176>
- Bhargava S, Loewenstein G, Sydnor J (2015) Do individuals make sensible health insurance decisions? Evidence from a menu with dominated options. NBER Working Papers 21160, National Bureau of Economic Research
- Bhargava S, Loewenstein G, Benartzi S (2017) The cost of poor health (plan choices) & prescriptions for reform. *Behav Sci* 3:12
- Bonito A, Bann C, Kuo M, et al (2000) Analysis of baseline measures in the Medicare current beneficiary survey for use in monitoring the national medicare education program. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Reports/Downloads/Bonito_2000_7.pdf. Accessed 23 Apr 2020
- Cafferata GL (1984) Knowledge of their health insurance coverage by the elderly. *Med Care* 22:835–847
- De Gagne J, PhD D, Oh J et al (2015) A mixed methods study of health care experience among Asian Indians in the southeastern United States. *J Transcult Nurs* 26:354–364. <https://doi.org/10.1177/1043659614526247>
- Desselle SP (2003) Consumers' lack of awareness on issues pertaining to their prescription drug coverage. *J Health Soc Policy* 17:21–39
- DistillerSR (n.a.) Evidence Partners, Ottawa, Canada. <https://www.evidencepartners.com>. Accessed April 2019–July 2020
- Edward J, Morris S, Mataoui F et al (2018) The impact of health and health insurance literacy on access to care for Hispanic/Latino communities. *Public Health Nurs* 35:176–183. <https://doi.org/10.1111/phn.12385>
- Flores G, Lin H, Walker C et al (2017) The health and healthcare impact of providing insurance coverage to uninsured children: A prospective observational study. *BMC Public Health* 17:Article number: 553. <https://doi.org/10.1186/s12889-017-4363-z>
- Garnick D, Hendricks A, Thorpe K et al (1993) How well do Americans understand their health coverage? *Health Aff* 12:204–212. <https://doi.org/10.1377/hlthaff.12.3.204>
- Ioannidis JPA, Greenland S, Hlatky MA et al (2014) Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 383:166–175. [https://doi.org/10.1016/S0140-6736\(13\)62227-8](https://doi.org/10.1016/S0140-6736(13)62227-8)
- James TG, Sullivan MK, Dumeny L, et al (2018) Health insurance literacy and health service utilization among college students. *J Am Coll Heal* 68:200–206. <https://doi.org/10.1080/07448481.2018.1538151>
- Lambert ZV (1980) Elderly consumers' knowledge related to Medigap protection needs. *J Consum Aff* 14:434–451. <https://doi.org/10.1111/j.1745-6606.1980.tb00680.x>
- Lockwood C, Munn Z, Porritt K (2015) Qualitative research synthesis: methodological guidance for systematic reviewers utilizing meta-aggregation. *JBIEvid Implement* 13:179–187. <https://doi.org/10.1097/XEB.0000000000000062>
- Loewenstein G, Friedman JY, McGill B et al (2013) Consumers' misunderstanding of health insurance. *J Health Econ* 32:850–862. <https://doi.org/10.1016/j.jhealeco.2013.04.004>
- Long S, Shartzer A, Polity M (2014) Low Levels of Self-Reported Literacy and Numeracy Create Barriers to Obtaining and Using Health Insurance Coverage. <http://apps.urban.org/features/hrms/briefs/Low-Levels-of-Self-Reported-Literacy-and-Numeracy.html>. Accessed 11 Feb 2019
- Marquis MS (1983) Consumers' knowledge about their health insurance coverage. *Health Care Financ Rev* 5:65–80
- McCall N, Rice T, Sangl J (1986) Consumer knowledge of Medicare and supplemental health insurance benefits. *Health Serv Res* 20:633–657
- Mccormack L, Uhrig J (2003) How does beneficiary knowledge of the Medicare program vary by type of insurance? *Med Care* 41:972–978
- McCormack LA et al (2002) Health insurance knowledge among Medicare beneficiaries. *Health Serv Res* 37:43–63
- McCormack L, Bann C, Uhrig J et al (2009) Health insurance literacy of older adults. *J Consum Aff* 43:223–248. <https://doi.org/10.1111/j.1745-6606.2009.01138.x>
- McGowan J, Sampson M, Salzwedel DM et al (2016) PRESS peer review of electronic search strategies: 2015 guideline statement. *J Clin Epidemiol* 75:40–46. <https://doi.org/10.1016/j.jclinepi.2016.01.021>

- Medicare & You | Medicare (n.a.) <https://www.medicare.gov/medicare-and-you>. Accessed 8 Oct 2020b
- Moher D, Liberati A, Tetzlaff J et al (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 6:e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Mokkink LB, Prinsen C, Patrick D et al (2018) COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 27:1147–1157. <https://doi.org/10.1007/s11136-018-1798-3>
- Paez KA, Mallery CJ, Noel H et al (2014) Development of the health insurance literacy measure (HILM): conceptualizing and measuring consumer ability to choose and use private health insurance. *J Health Commun* 19:225–239. <https://doi.org/10.1080/10810730.2014.936568>
- Quincy L (2012a) What's behind the door: consumer difficulties selecting health plans. In: Consumers union. <https://consumersunion.org/research/whats-behind-the-door-consumer-difficulties-selecting-health-plans/>. Accessed 17 Jul 2018
- Quincy L (2012b) Measuring health insurance literacy: a call to action. https://consumersunion.org/pub/Health_Insurance_Literacy_Roundtable_rpt.pdf. Accessed 17 Jul 2018
- StataCorp LLC (2019) Stata statistical software: release 16. StataCorp LLC, College Station
- Tennyson S (2011) Consumers' insurance literacy: evidence from survey data. *Finan Serv Rev* 20:165–179
- Terwee CB, Bot SDM, de Boer MR et al (2007) Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 60:34–42. <https://doi.org/10.1016/j.jclinepi.2006.03.012>
- Terwee CB, Prinsen CAC, Chiarotto A et al (2018) COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res* 27:1159–1170. <https://doi.org/10.1007/s11136-018-1829-0>
- Tipimene R, Politi MC, Kullgren JT et al (2018) Association between health insurance literacy and avoidance of health care services owing to cost. *JAMA Netw Open* 1:e184796. <https://doi.org/10.1001/jamanetworkopen.2018.4796>
- Tseng C-W, Dudley RA, Brook RH et al (2009) Elderly patients' knowledge of drug benefit caps and communication with providers about exceeding caps. *J Am Geriatr Soc* 57:848–854. <https://doi.org/10.1111/j.1532-5415.2009.02244.x>
- Uhrig J, Squire C, McCormack L, et al (2002) Questionnaire Development Final Report 110
- Vardell EJ (2017) Health insurance literacy: how people understand and make health insurance purchase decisions. Ph.D., the University of North Carolina at Chapel Hill

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.