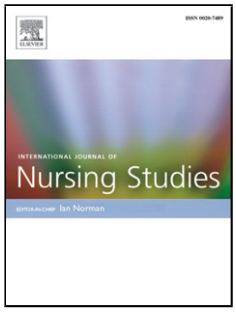

Title: Patient classification systems used to classify nursing intensity and assess nursing staffing resources in home health care - a scoping review

Authors: Jill Flo, Bjørg Landmark, Siri Tønnessen, Lisbeth Fagerström

PII: S0020-7489(19)30133-6
DOI: https://doi.org/10.1016/j.ijnurstu.2019.05.009
Reference: NS 3361

To appear in:

Received date: 14 December 2018
Revised date: 12 May 2019
Accepted date: 14 May 2019

Please cite this article as: Flo J, Landmark B, Tønnessen S, Fagerström L, Patient classification systems used to classify nursing intensity and assess nursing staffing resources in home health care - a scoping review, *International Journal of Nursing Studies* (2019), https://doi.org/10.1016/j.ijnurstu.2019.05.009

Patient classification systems used to classify nursing intensity and assess nursing staffing resources in home health care - a scoping review

Jill Flo[1], Bjørg Landmark[2], Siri Tønnessen[3], Lisbeth Fagerström[4]

[1]Faculty of Health and Social Sciences, Department of Nursing and Health Sciences, University College of Southeastern Norway, Drammen, Norway

[2]Faculty of Health and Social Sciences, Department of Nursing and Health Sciences, University College of Southeastern Norway, Drammen, Norway

[3]Faculty of Health and Social Sciences, Department of Nursing and Health Sciences, University College of Southeastern Norway, Vestfold, Norway

[4]Faculty of Health and Social Sciences, Department of Nursing and Health Sciences, University College of Southeastern Norway, Drammen, Norway

Faculty of education and welfare studies, Åbo Akademi, Vasa, Finland

*Corresponding author: Jill Flo, University College of Southeastern Norway, Box 7053, 3007 Drammen, Norway, E-mail: Jill.Flo@usn.no

Abstract

Objectives: To identify the patient classification systems used to classify nursing intensity in the assessment of nursing staffing resources currently used in home health care, with a special emphasis on validity, reliability and staff allocation.

Design: Scoping review of internationally published and grey literature, based on a methodological framework by Arksey and O'Malley.

Data sources: Searches of the electronic databases Cinahl, Medline, Embase and SweMed, the websites Google and Google Scholar and hand searches of reference lists occurred. Eligibility criteria included (A) a focus on patient classification systems measuring nursing intensity and workload in home health care and (B) published in English between January 2007 and March 2019. In level one testing two team members screened titles and abstracts, in level two testing two team members determined which papers should undergo a full text review. Data were extracted using structured extraction by one team member and verified by two other members.

Results: Thirteen peer-reviewed articles and grey literature documents were identified, from Canada, Ireland, the UK, the USA, Scotland, Turkey and the Netherlands. Four patient classification systems had been tested for both validity and reliability. Validity was tested through face validity, predictive validity, concurrent validity or content validity index. Reliability was tested through stability, internal consistency, observer agreement or inter rater reliability. One patient classification system had been tested only for reliability, through interrater reliability and observer agreement. Two patient classification systems had been evaluated through summative evaluation; one qualitatively through focus group interviews and one through semi-structured interviews. Only one patient classification system had been validity and reliability tested and evaluated. Overall, the patient classification systems in the included papers (13)

were considered to have benefits and to be appropriate for the measurement of patients' needs, workload and allocation of staff, although specific information was not always given.

Conclusion: Little has been published on validity or reliability tested patient classification systems linked to staffing allocation in home health care in the past decade. Limited research was seen where a patient classification system was considered to be fully operational in home health care.

Keywords: Home health care; nursing intensity; patient classification system, resource allocation; workload

1. Introduction

The population of older people in developed countries is estimated to increase from today's levels (WHO, 2016), and as people age their health needs become more chronic and complex, with several chronic diagnoses (The National Board of Health and Welfare, 2018, WHO, 2016). Concurrently, a shift from hospital-based care to community-based nursing can be discerned (Eurostat, 2016). Given that a larger older population will increase stress on health and long-term care systems, models for well-functioning municipal health and care services are being sought (Bloom et al., 2015).

The conceptualization, definition and content of home health care varies throughout the world, and numerous terms are used to describe the concept of care given in the home, including homecare services, home help services, home care, home based care, community health care and district nursing (Bing-Jonsson et al., 2013, Bôas and Shimizu, 2015, Drennan, 2019, Nielsen and Jørgensen, 2016). The nursing skill mix required in home health care and the titles given those providing such care differ throughout the world: in Norway, registered nurses and assistant nurses provide such care (Bing-Jonsson et al., 2013); in Belgium, home nurse and health care assistants (De Vliegher et al., 2014, De Vliegher et al., 2016); in Denmark, home health care givers (Nielsen and Jørgensen, 2016); in Sweden, home care assistants and district nurses (Craftman et al., 2018, Nordmark et al., 2015); in the United Kingdom (UK) and New Zealand, district nurses (Drennan, 2019, Walker and Hendry, 2009); in the United States of America (USA) registered nurses, licensed practice nurses, home health aides and certified nursing assistants (Luo et al., 2013); and in Ireland, public health nurses (Brady et al., 2007).

The common denominator in home health care is that nursing staff provide care in the home to people with a large variety of care needs. Nursing staff's activity profile includes direct patient care, surveillance/monitoring, administration, and psychosocial care and, irrespective of country, the provision of multidimensional care. Travel time is also a component of home health care (De

Vliegher et al., 2014). Today, home health care nursing staff perform more skilled nursing work and tasks and provide more complex and technical nursing care than before (De Vliegher et al., 2014, Drennan, 2019, The National Board of Health and Welfare, 2018). To meet both present and future needs in home health care, the systematic and regular monitoring of patients' need for help, care, support and interventions is needed, through work force planning tools or patient classification systems, so as to ensure the correct allocation of staff.

While the terms nursing work, nursing workload, patient dependency and nursing intensity are frequently used to describe the same or similar concepts (Morris et al., 2007), nursing workload is related to and defined by the amount or level of work that each nurse carries out.

The primary purpose underlying the use of a patient classification system is to facilitate nursing staff's ability to respond to the variable nature of nursing care (Edwardson and Giovannetti, 1994, Huckabay and Skonieczny, 1981). Patient classification systems also provide staff with a tool whereby nursing personnel resources can be determined and correctly allocated (Edwardson and Giovannetti, 1994, Giovannetti, 1984).

Two basic types of patient classification systems exist, prototype and factor evaluation systems. Prototype evaluation systems are characterized by descriptions of typical patients (prototypes) or typical nursing tasks in every patient class (Brady et al., 2007, Giovannetti, 1979, Giovannetti and Johnson,

1990, Saba, 2002). Used when patient classification systems were first developed, common prototype evaluation systems are still common but have to a certain extent been replaced by factor evaluation systems (Brady et al., 2007, Rauhala, 2008). Factor evaluation systems are characterized by the rating of several critical indicators or factors of nursing care and patients' individual characteristics (Brady et al., 2007, Saba, 2002). A fully operational patient classification system must include at least six essential elements: (1) A tool to predict nursing care requirements for the individual patient, (2) Validation of the amount of care given, (3) Evaluation of the patterns of care delivery, (4) Revalidation of the amount of care, (5) Relation of nursing care requirements to staff resource allocation, (6) A method for monitoring reliability over time (DeGroot, 1989).

Some of the most well-known patient classification systems used in home health care have been developed in the USA, Canada or the UK, for example: the Omaha system, the Clinical Care Classification (CCC) system (Saba, 2017, Saba, 1992, Saba, 2002), the Resident Assessment Instrument-Home Care (RAI-HC) (Carpenter and Hirdes, 2013), the Resource Utilization Groups (RUG-III), the Resident Assessment Instrument (RAI) and associated Minimum Data Set (MDS) (Björkgren et al., 1999, Brown, 2001), the Community Health Intensity Rating scale (CHIRS), (Hays et al., 1999, Peters, 1988), workforce planning and development data and methods (WP&D) (Hurst, 2006) and the Easley-Storfjell

Patient Classification Instrument for Caseload/Workload Analyses (CL/WLA) (Albrecht, 1991, Storfjell et al., 1997).

Scoping reviews provide an overview of existing evidence and thus broader knowledge of a subject. While a scoping review does not include quality appraisal (Arksey and O'Malley, 2005), the selection criteria used here were designed to capture all relevant research in order to provide a broad evidence base. Scoping reviews of nurse practitioners in primary health care (Grant et al., 2017), nurse practitioners' workload in primary health care (Martin-Misener et al., 2016), school nurse workload (Endsley, 2017) and nursing workforce planning and forecasting research (hospital) (Squires et al., 2017) were found. Literature reviews of caseload management in district nursing (Roberson, 2016), patient dependency (Walker and Hendry, 2009), safe-staffing models in care homes (Mitchell et al., 2017) and nursing hours per patient per day (Min and Scott, 2016) were also found. However, no scoping review following the framework of Arksey and O'Malley (Arksey and O'Malley, 2005) and including the potential size and scope of available research literature and grey literature (Grant and Booth, 2009) was seen with regard to the use and development of patient classification systems in home health care.

The aim of this scoping review was to identify the patient classification systems used to classify nursing intensity in the assessment of nursing staffing resources

in home health care during the most recent decade (2007-2017, updated March 2019), with a special emphasis on validity, reliability and staff allocation.

We focused on three of the system elements that DeGroot considered to be essential for a fully operational patient classification system (DeGroot, 1989), specifically that the tool can predict nursing care requirements, the tool is validated and reliability tested, and the nursing care requirements are in line with staff resource allocation. The following interrelated research questions guided the scope of the study: (1) What is the target population for patient classification systems used in home health care? (2) Which instruments have been developed to assess nursing care requirements for the individual patient's needs and nursing intensity? (3) Which patient classification systems used in home health care have been tested for validity or reliability? (4) Can patient classification systems be used for the allocation of staff in home health care?

2. Design and Method

To review the available literature on patient classification systems used in home health care, a scoping review design based on a framework by Arksey and O'Malley (Arksey and O'Malley, 2005) and further enhanced by Levac, Colquhoun and O'Brien (Levac et al., 2010) and the methodology for Scoping Reviews (The Joanna Briggs Institute, 2015) was chosen. Steps 1-5 of the Arksey and O'Malley methodological framework were followed during the course of this study: (i) Identifying the research question, (ii) Identifying

relevant studies, (iii) Study section, (iv) Charting the data, (v) Collating, summarizing and reporting the results (Arksey and O'Malley, 2005).

Given the more recent shift from hospital-based to community-based care and the subsequent demand for well-functioning home health care, the decision was made to search for literature from the most recent decade (2007-2017, updated 2019).

2.1 Identifying relevant literature

The aim of a scoping review is to be as comprehensive as possible in identifying published or unpublished studies and reviews (Arksey and O'Malley, 2005). The search strategy in this study included an electronic database search, a web search (Google/Google scholar, websites) and a hand search of relevant journals. The searches resulted in a total of 1247 records from electronic databases (534 in Cinahl, 509 in Medline, 173 in Embase and 31 in SweMed) and 56 records from "grey literature" and key journals. After duplicates were removed, 1040 records were seen (Figure 1).

2.2 Electronic database search

Thesaurus/medical subject headings (MeSH) and keywords comprised the search terms used during the electronic database search, with MeSH headings "exploded" where possible. Boolean operators (OR, AND) were used to narrow and expand the search. The search criteria were limited to peer-reviewed articles, English language and date of publication 2007-2017, performed in

August 2017 and updated in March 2019. The search terms were nursing intensity, patient classification, patient acuity, care dependency, workload, measurement, workload and nurse patient ratio: all combined with home care services, home health care or community health nursing.

2.3 Google and Google Scholar

To ensure a broader perspective, an additional grey literature search was carried out in November 2017, with a subsequent update in March 2019, of the Google/Google Scholar websites. The same search terms were used as during the electronic database search.

2.4 Key journals

Key journals were hand-searched, because electronic databases may be incomplete (Arksey and O'Malley, 2005). Journals were identified using Pub Med PubReMiner, and seven international journals were found to be of interest. A hand search of the following journals occurred: British Journal of Community Nursing, Community Practice, Journal of Clinical Nursing, International Journal of Nursing Studies, Journal of Advanced Nursing, Journal of Nursing Management and Home Healthcare Now. Searches were restricted to two years, from January 2016 to November/December 2017 and updated in March 2019. When journals contained relevant papers, the hand search restrictions were extended to three years.

The first author (X), a health sciences librarian with expertise in web-based information systems and systematic review methodology, conducted both the electronic database and website searches. The results were uploaded to a web-based reference management program (EndNote X8).

2.5 Selecting the literature

Selecting literature can be considered an iterative process in that it involves searching, refining the search strategy and reviewing articles for study inclusion (Levac et al., 2010). At the start of the review process the authors (XX) discussed inclusion and exclusion criteria and concluded that, for inclusion, papers must have a focus on home health care/home care services in the community and a classification system or tool whereby nursing intensity, patient classification, patient acuity, care dependency, workload and/or nurse patient ratio were measured. Thus papers not meeting the aforementioned criteria were excluded. The identification phase resulted in 1247 papers, with 1040 remaining after duplicates were removed (Figure 1). In level one testing, the screening phase, two authors (XX) screened titles and abstracts, first independently reviewing each record's title/abstract then together deciding on inclusion or exclusion, resulting in 55 papers. When selecting the literature, the inclusion and exclusion criteria noted above were followed. In level two testing, two authors (XX) first independently reviewed each paper's abstract/full text then together decided on inclusion or exclusion, resulting in 16 papers. Thereafter, three

authors (XXX) first independently engaged in full text reviews of the 16 papers to determine relevance and then together decided on inclusion or exclusion, with three papers being excluded because they did not include a description of a workload instrument or patient classification system used in home health care. Thus, 13 papers were included in this scoping review.

2.6 Charting the data

Charting data is also an iterative process in that it involves extracting data from the included studies (Levac et al., 2010). Data extracted from each paper (n = 13) in this scoping review included author(s), year of publication, country of origin, study location, population, sample size and context, type of instrument/tool, and whether the instrument/tool was validity or reliability tested or evaluated (Table 1).

3. Results

Thirteen papers met the inclusion criteria, published in Canada (3), Ireland (3), the UK (2), the USA (2), Scotland (1), Turkey (1) or the Netherlands (1) (Table 1).

3.1 Target population

The target populations included care for adults (Bowers and Durrant, 2014), older people aged 65 years and over (Byrne et al., 2007), older people's service (Kane, 2014), people aged 18 years and over (Poss et al., 2008) and people aged

zero to 70 years and older (Çelebioğlu et al., 2007). In some papers exact population ages were not given; instead the researchers noted that home health care nurses provided services for entire and diverse populations (Chapman et al., 2017, Grafen and Mackenzie, 2015, Hawes et al., 2007, Kane, 2014). The populations included individuals with complex clinical care needs and individuals in need of palliative care (Cawthorn and Rybak, 2008, Collister et al., 2014). In one source, the population was not specified (Storfjell et al., 2017).

3.2 Tools or instruments used to measure patients' needs and nurses' workload

Different types of instruments and workload systems were seen (Table 2).

Minimal information was included on the Domiciliary care system in the community (DominiC), an electronic workload management tool (Bowers and Durrant, 2014). Using DominiC, visiting district nurses identified patients' needs and recorded clinical nursing needs, estimates of frequencies of future visits, length of visits in minutes, referral source, pertinent risk assessment information, key quality and care outcome indicators and the staff skills needed to meet individual patient needs.

The Community Client Need Classification System (CCNCS) was described as a dependency-based workload system with seven assessment criteria and additional weighing (Brady et al., 2008, Byrne et al., 2007). The seven assessment criteria included: (1) Nursing assessment, (2) Physical care requirements, (3) Teaching needs and health promotion, (4) Carer and family

support, (5) Case/Care management, (6) Psycho-social support, (7) Environmental factors. Using the CCNCS, nurses scored patients' need levels from 1 (low need) to 5 (high need). Total needs were thereafter rated on a scale: 1 (0-7), 2 (8-14), 3 (15-21), 4, (22-28) and 5 (29-40). The travel time per visit (>20 minutes) was also recorded.

The Client Audit Community Care Workload Assessment Tool (Cawthorn and Rybak, 2008) allowed nurses to collect basic patient demographic information and diagnosis, with each patient assessed using the following criteria: identified nursing needs, frequency of visits, length of home visit, how far the patient resides from the home office, coordination time required, client stability, coping skill. The weighting score was: 8-13 = low intervention, 14-21 = moderate intervention and >22 = high intervention.

The Community Health Intensity Rating Scale (CHIRS) was used to assess the intensity of need for care for people/families in a community (Çelebioğlu et al., 2007). In the study included here, the original tool was translated into Turkish and an abbreviated CHIRS scoring sheet developed because of the tool's lengthiness. A Nursing Assessment section included 15 parameters, 91 items and 974 subitems encompassing 4 domains: (1) Environmental (finances and physical environment/safety), (2) Psychosocial (community networking, family system, emotional/mental response, individual growth and development), (3) Physical (sensory function, respiratory/circulatory, neuromusculoskeletal,

reproductive, digestion/elimination, structural integrity), (4) Health behavior (nutrition, personal habits and health management). Scores ranged from 0 (no need for care) to 4 (extreme need for care) and were used as indicators of the intensity of health care needs.

Little specific information was given about the Caseload Classification Tool for Community Nursing (Chapman et al., 2017), which was developed to assess twelve care need domains, including clinically evidenced-based care plans and three levels of care (routine, additional and significant). The tool focused on assessment, social circumstances, liaison, dignity and respect, intervention, band of staff providing care, expected visit time and length of stay on the caseload.

The Caseload Intensity Tool (CIT) was used to determine client intensity (Collister et al., 2014). Scores for each client were summarized into a six-category Client Intensity Scale: (A) Clients requiring minimal services, (B) Clinical condition: stability and predictability, (C) Clinical condition: complexity factors, (D) Care response: treatment and therapy process, (E) Care response: client/family care plan process and (F) Care response: system care plan process. The CIT was divided into three levels, level 1 = 1 score, level 2 = 2 scores and level 3 = 3 scores (1 score minimum and 15 maximum), and provided a raw score. The CIT raw score was then converted into a Client Intensity Scale, with 1 score becoming 1 (minimal), 4-6 becoming 2 (basic), 7-10 becoming 3 (moderate), 11-13 becoming 4 (significant) and 14-15 becoming

5 (extreme). Each staff member's caseload score was determined from the weighted sum of her/his caseload multiplied by the number of clients, for each of the Client Intensity Scale levels.

The Scottish Community Nursing Workload Measurement Tool enabled nurses to record and report their actual workload using six activity categories (Grafen and Mackenzie, 2015). The categories were: (1) Face-to face contact (everything occurring when the patient is present), (2) Non-face-to-face context (everything occurring when the patient is not present), (3) Home visits and planned sessions (routine, group and mass clinics/sessions), (4) Associated workload (management, administration, meetings and professional development activity), (5) Travel (including walking, waiting and parking time), and (6) Exception reporting (extraordinary events, for example adverse weather, car breakdown). There were four defined levels of intervention, ranging from straightforward (level 1) to complex (level 2). The tool captured workload on selected days, included a section for recording daily working hours and could be used daily or on an occasional basis over a 10-day period.

The Resident Assessment Instrument for Home Care (RAI-HC) was used to highlight issues related to functioning and quality of life and can be used to guide the development of individualized service plans (Hawes et al., 2007). The RAI-HC consisted of two elements, the Minimum Data Set-Home Care (MDS-HC) and Clinical Assessment Protocols (CAPs). The MDS-HC was used to

collect standardized patient information and asses 19 key domains related to function, health, social support and service use. The CAPs consisted of 30 problem-focused protocol areas that encompassed common risks for home health care clients, organized under the following areas: functional performance, sensory performance, mental health, bladder management, health problem/syndromes, and service oversight.

The Electronic Caseload Analysis Tool (eCAT) was described as a knowledge-based tool used to facilitate the caseload analysis process (Kane, 2014). There were five user levels (caseload holder, team leader, operational manager, service manager and commissioner), but only the caseload holder could enter data. The eCAT included eight categories: (1) Demography, (2) Caseload size, (3) Visiting patterns, (4) Reason for visit, (5) Dependency measure, (6) Caseload throughput, (7) Location of care and (8) Reviews.

Based on the theory of Virginia Henderson, the Care Dependency Scale (CDS) was used to provide standardized and meaningful descriptions of the type and intensity of nursing care in a holistic way (Kottner et al., 2010). The tool was used to measure 15 items (categories), rated using scores ranging from 1 (completely dependent) to 5 (completely independent). The scores from all 15 categories were added together, yielding sum scores that ranged from 15–75: *low sum scores* indicated *high care of dependency* and *high sum scores* indicated *independency.*

The Resource Utilization Groups Version III for Home Care (RUG-III/HC) (Poss et al., 2008) was described as a case-mix system that included seven hierarchical levels: (1) Special rehabilitation, (2) Extensive services, (3) Special care, (4) Clinically complex, (5) Impaired cognition, (6) Behavior problems and (7) Reduced physical functions. Each hierarchical level had different criteria, for example: (2) Extensive services includes tracheostomy, respirator, respiratory therapy while (3) Special care includes, for example, stage 3 or 4 pressure ulcer, enteral tube feeding, diagnosis of Multiple Sclerosis, treatment for burns, radiation treatment, intravenous, fever.

The Easley-Storfjell Instruments for Caseload/Workload analysis (CL/WLA) (Storfjell et al., 2017) were designed to help home health care managers plan, monitor and evaluate nursing activities. The analysis process consisted of four steps. Step 1: Analyze each case to predict the number of visits required and determine complexity of nursing care. Step 2: Chart time and complexity ratings on visual graph. Step 3: Calculate time for noncaseload work requirements or duties. Step 4: Summarize findings and compare number of required visits with workload analysis projection. In step 1, the number of visits was rated from 1 to 4, with 1 = 1 visit or fewer per month, 2 = 2-3 visits per months, 3 = 2-3 visits per week and 4 = 3-5 visits per week. The tool was used to determine complexity of care based on the assessment of six variables: (A) Clinical judgment required (assessment needs), (B) Teaching needs, (C) Physical care

needs (technical procedures), (D) Psychosocial support needs, (E) Coordination and care management needs, (F) Number and severity of problems.

3.3 Measuring patients' needs, nursing care requirements and assessment levels

All of the workload instruments and patient classification systems included here measured patients' needs and/or nursing requirements. How patients' needs were described and measured varied, encompassing a range from dependency, intensity, primary needs, human needs to nurse problems. Patients' needs were measured and scored in different ways, but the rating of needs from low to high was most common. Various low to high descriptors were seen in the CCNCS (low need to high need) (Brady et al., 2008, Byrne et al., 2007), the Client Audit Community Care Workload Tool (low intervention to high intervention) (Cawthorn and Rybak, 2008), the CHIRS (no need for care to extreme need for care) (Çelebioğlu et al., 2007), the Caseload classification tool (routine, additional to significant) (Chapman et al., 2017), the CIT (minimal to extreme) (Collister et al., 2014), The Scottish Community Nursing Workload Measurement Tool (straightforward to complex) (Grafen and Mackenzie, 2015), the RAI-HC (independent to severe impairment) (Hawes et al., 2007) and the Easley-Storfjell CL/WLA (minimal complexity to very great complexity) (Storfjell et al., 2017). In only one paper did the information on how needs were rated differ, regarding the CDS instrument (completely dependent to completely independent) (Kottner et al., 2010). In the papers on the RUG-III/HC (Poss et

al., 2008) and DominiC (Bowers and Durrant, 2014), no information on how needs were rated was given.

3.4 Instruments tested for validity

Validity refers to the degree to which an instrument measures what it is supposed to measure (Pallant, 2015, Polit and Beck, 2004). The terms face validity and content validity are technical terms used during the judgment of whether a scale looks reasonable (Streiner et al., 2015). Still, face validity should not to be used to provide primary evidence of an instrument's validity; other types of validity should also be demonstrated (Polit and Beck, 2004). Different types of validity testing exist and can be used to assess various aspects for certain contexts, such as criterion-related validity (predictive- and concurrent validity) or construct validity (Kline, 2005). The most common types of validity testing seen in the papers included here were face-, content- and predictive validity. Only five instruments were or had been tested for validity (Brady et al., 2008, Çelebioğlu et al., 2007, Collister et al., 2014, Hawes et al., 2007, Poss et al., 2008) (Table 2).

The CCNCS was tested for face validity and content validity (clarity, relevance and representativeness) by expert public health nurses and tested for predictive validity through the recording of the time per client per week for the four weeks of the study (Brady et al., 2008). The CCNCS' content validity was assessed through the measurement of Content Validity Index (CVI), with an overall CVI

of 0.99, indicating a high level of agreement between the public health nurses. Its predictive validity was calculated using the Kruskal-Wallis test, Jonckheere-Terpstra test and Mann Whitney U-test (Brady et al., 2008).

The Turkish language version of the CHIRS was face validity tested by six experts (Çelebioğlu et al., 2007). The language in the Turkish CHRIS was found to be suitable and its scale comprehensive but very long. Predictive validity was measured through the correlation of the Total Scale Score (TSS), including the total number of household members, total number of home visits during the last year, total number of visits to any health institution and participants' self-evaluation score of health care needs. Positive and significant correlation was seen between the TSS and participants' self-evaluation (Çelebioğlu et al., 2007).

The content validity of the CIT was determined using a careful selection of key factors from literature and staff participants as experts in the development of the tool's content (Collister et al., 2014). Its face validity was tested through the collection of qualitative data from the study participants, and its concurrent validity was also tested. The RUG-III/HC was validated using cost episodes on the individual level over a 13-week period in Canada (Poss et al., 2008). While not validity tested in the study included in this review, the RAI-HC has been tested for content validity and convergent validity in earlier studies (Hawes et al., 2007).

3.5 Instruments tested for reliability

Reliability is a fundamental way to reflect the amount of error in an instrument (Streiner et al., 2015) and refers to the accuracy and consistency of the instrument (Polit and Beck, 2004). Many different methods exist whereby one can test an instrument's reliability, including stability, internal consistency and/or equivalence (Kline, 2005, Streiner et al., 2015), and different methods were seen in the papers included here.

Five instruments were or had been tested for reliability (Brady et al., 2008, Çelebioğlu et al., 2007, Collister et al., 2014, Hawes et al., 2007, Kottner et al., 2010) (Table 2). The CCNCS was reliability tested for stability and consistency (Brady et al., 2008), and two scenarios developed: a pre- and post-testing scenario with one child case and one older patient case, rated by public health nurses. After using the CCNCS in practice for 10 days, the public health nurses rated the scenarios again (10-day-interval ratings). Cohen's kappa for the child case was 0.593 (moderate agreement) and for the older patient case 0.704 (substantial agreement). The internal consistency was measured by assessing the percentage agreement between public health nurses regarding client level of need. Despite different aims and focus, the reliability findings seen in the two CCNCS studies (Brady et al., 2008, Byrne et al., 2007) were similar.

The reliability of the Turkish language version of the CHIRS was determined using the total mean scale score (TSS) (26.66 ±5.32); the Cronbach's alpha coefficient for overall scale (0.525), environmental domain (0.503),

psychosocial domain (0.404), psychological domain (0.416), health behavior domain (0.453); and the Guttman split-half coefficient for internal consistency (0.629) (Çelebioğlu et al., 2007).

The CIT was internal reliability tested using Cronbach's alpha (range 0.79 to 0.95) and its interrater reliability tested by seventy-five staff, who scored case-study clients the same or within one client intensity score level 71% of the time (Collister et al., 2014). Its stability was tested by Wilcoxon paired test to determine differences in whether the tool was used at different times, with analysis indicating that the score did not change significantly when the time between tests was brief but did change significantly when the time was, on average, 38 days (Collister et al., 2014).

The RAI-HC had passed the test for reliability in earlier multiple trials, with a weighted kappa of 0.7 or more in dual to independent assessment of randomly selected home health care clients in the USA, Canada, Japan, Australia and the Czech Republic (Hawes et al., 2007).

The CDS was first tested for reliability (Kottner et al., 2010) by the nurses primarily responsible for a patient's care. One group of nurses conducted a first assessment, with a second group of nurses, blinded from the first data collection, conducting a second assessment with an interval of 1-3 days. The intra class correlation coefficient was calculated to indicate interrater reliability based on ANOVA (0.67-0.88) and simple kappa coefficient (0.61-0.79). Proportions of

observed agreement was calculated (0.76-0.94) and standard errors of measurement (0.28-0.56) was used to indicate interrater agreement (Kottner et al., 2010) between items. In general, the agreement and reliability values were high, although some items had the lowest agreement coefficients, for example recreational activities and mobility.

3.6 General evaluation of the instruments

Several (5) of the instruments were not validity or reliability tested but instead evaluated from different angles (Bowers and Durrant, 2014, Cawthorn and Rybak, 2008, Chapman et al., 2017, Grafen and Mackenzie, 2015, Kane, 2014).

The DominiC was considered a credible workforce tool and showed that more health care assistants and assistant practitioners were needed (Bowers and Durrant, 2014), although no information was given regarding the methods used for evaluation. One challenge during the development of the DominiC was that the project team appeared to lack both sufficient time and computer programming skills. A further challenge was the entering of accurate, timely data into the electronic system, as some of the staff did not realize the importance of capturing data.

Nurses (44) used the CCNCS to carry out caseload analysis as a method of evaluation, also recording the amount of care time spent with each client. It was seen that the total nursing time increased when the level of patient dependency

increased (Kruskal-Wallis: H=236·648, d.f.=4, p<0.001) and that the majority of clients had low to moderate levels of need (77%) (Byrne et al., 2007).

The benefits and limitations of the Client Audit Community Care Workload Assessment Tool as experienced during its implementation were presented: for example, that it facilitated understanding of workload and requirements associated with each patient and that it was work intensive and difficult to fit into a normal working day (Cawthorn and Rybak, 2008). There was no explicit mention of an evaluation.

Evaluation of the Caseload Classification Tool occurred over a 3-month period. The evaluation included analyses of the data produced (Chapman et al., 2017) and semi-structured interviews with staff on their experiences of using the tool. However, information on the number of participants or methodological issues was not included. It was determined that using a standardized approach can improve safety and quality of information, confirm the appropriate workforce and allow a more transparent picture of caseloads (Chapman et al., 2017).

The Scottish Community Nursing Workload Measurement Tool (Grafen and Mackenzie, 2015) was developed by community nursing staff. One challenge was the securement of the engagement of all community nursing staff, even though the tool was considered easy and simple to use. Another challenge was the setting, due to the differences between nursing in the community as opposed to an inpatient area. Appropriate levels of education and training in use of the

tool were highlighted. No information was given about the method used for the evaluation of the tool.

The implementation of the eCAT was evaluated through a series of focus group interviews with a semi-structured interview design across all trusts (Kane, 2014), yet more specific information about the focus group interviews was not given. The eCAT can benefit patients' care by improving nurses' (caseload holders') performance and maximizing staff resources.

3.7 Patient classification systems as a tool for the allocation of staff in HHC

Overall, the patient classification systems in the included papers (13) were considered to have benefits and to be appropriate for the measurement of patients' needs, workload and allocation of staff, although specific information was not always given. The DominiC was considered a credible workforce tool in regard to the reporting of incoming referrals, the number of patients seen, patients' needs and whether demands can be met with current staffing levels. Nurses found that the DominiC offered a real-time overview of patients' needs and caseload holders could report annually on caseload profiles (Bowers and Durrant, 2014). The CCNCS was found to measure workload with a good degree of validity and reliability (Brady et al., 2008) and the relationship between level of need and nursing time was analyzed, with the instrument considered useful in predicting the number of nurses needed for patients receiving community nursing services (Byrne et al., 2007). The CHIRS and CIT

were considered to measure workload with a good degree of reliability and validity, but specific information on time registration or the recording of staff situation was not given. When used by trained nurses in a home health care setting, the CDS was considered to provide reliable and reproducible results (Kottner et al., 2010), but no information on the recording of staff situation was given. The RAI-HC was considered to include a high-quality assessment and care planning design for use with a variety of home health care populations when implemented in conjunction with federal- and state-regulated programs (Hawes et al., 2007), but no information on the recording of staff situation was given. The eCAT was assessed as verifying and providing contemporaneous information on district nursing caseloads in the context of population needs, nursing activity, staff resources, dependency and service design at each level of nursing practice, service delivery and commissioning (Kane, 2014). The RUG-III/HC was considered suitable for adults who use home health care services for 60 days or longer and useful for describing long-stay populations (Poss et al., 2008). The CCNCS was found to provide staff and management with valuable information and an understanding of work requirements, and after the weighting of each patient's score, staff requirements became more transparent (Cawthorn and Rybak, 2008).The CL/WLA system was considered viable in the planning, monitoring and evaluation of nursing activities, including the calculation of the number of home visits per nurse through the division of total time available by the average time needed for a home visit (Storfjell et al., 2017).

Of the patient classification instruments seen in this scoping review, seven were developed prior to being investigated in the studies included here (Brady et al., 2008, Byrne et al., 2007, Çelebioğlu et al., 2007, Hawes et al., 2007, Kane, 2014, Kottner et al., 2010, Poss et al., 2008, Storfjell et al., 2017). Five were developed and tested in pilot projects (Bowers and Durrant, 2014, Cawthorn and Rybak, 2008, Chapman et al., 2017, Collister et al., 2014, Grafen and Mackenzie, 2015), yet no information was given regarding whether these five were then actually implemented or in continuous use.

4. Discussion

Using a scientifically tested patient classification system, managers are able to balance patients' needs and nursing resources with regard to optimizing patients' care. To be considered fully operational, patient classification systems should include certain essential elements. Validity and reliability, two such elements, are extremely important for decision making and an evidenced-based leadership. According to Giovannetti, before an instrument can be used with confidence, both validity and reliability must be established (Giovannetti, 1979).

Thirteen papers were included in this scoping review, with the majority of studies being set in Canada, Ireland, the UK or the USA. Of the thirteen patient classification systems, four were both validity and reliability tested and one was only reliability tested. Although limited information was provided on the various patient classification systems' validity and reliability, they were nonetheless

considered to have benefits and be appropriate with regard to measuring patients' needs, workload and the allocation of staff. However, there seems to be a need to further illuminate the use of patient classification systems in home health care.

The patient classification systems' target populations

The patient classification systems used in small trusts or community-based settings were all sufficiently described in the relevant papers (Table 1). The DominiC was implemented in one trust in England (Bowers and Durrant, 2014), the Client Audit Community Care Workload assessment Tool in a rural facility in Canada (Cawthorn and Rybak, 2008), the Caseload Classification Tool in one locality in England (Chapman et al., 2017), the CCNCS in three communities in Ireland (Byrne et al., 2007), the CIT in five geographic zones in Canada (Collister et al., 2014) and the Scottish Community Nursing Workload Measurement Tool in 14 National Health Service Scotland boards (Grafen and Mackenzie, 2015). This shows a range of different patient classification systems being used in small as well as national and international trusts.

In the included papers, the target populations varied between zero (Çelebioğlu et al., 2007) to age 65 years and over (Byrne et al., 2007, Çelebioğlu et al., 2007). Patients with different needs of care and complexity were cared for in the community (Cawthorn and Rybak, 2008, Collister et al., 2014). The complexity of home health care has been well described, for example in terms of depressive

symptoms (Szczerbińska et al., 2012), frail older people with complex health problems (Skilbeck et al., 2018), or loneliness (Tomstad et al., 2017). Older people can also worry about being a burden to others or losing their self-government, or feel fear or anxiety arising from existential thoughts or the loss of social ties (Hafskjold et al., 2016), which reveals the complexity that home health care nurses and nurse managers must navigate.

Instruments used to assess nursing care requirements

In the papers included in this scoping review, there was a focus on instruments that measure patients' needs, although the exact needs measured were not always delineated and at times minimal information was provided (Bowers and Durrant, 2014, Chapman et al., 2017). A difference was also seen between how patients' needs were defined, assessed and scored in the various patient classification systems. Only occasionally were criteria outside of direct care needs included, for example frequency or length of home visits, how far the patient lived from the home health care office (Cawthorn and Rybak, 2008) or travel time and exception reporting (Grafen and Mackenzie, 2015). Assessment should include non-patient factors (contextual, organizational and staff-related factors) that affect nurses' total workload (Fagerström and Vainikainen, 2014), indirect care-related nursing activities (phone call, ordering medication), and non-patient activities (education, meeting) (Morris et al., 2007). Total nursing

time includes both nursing and non-nursing activities (Alghamdi, 2016), which is of great importance when discussing workload in home health care settings.

Instruments tested for validity and reliability

An important finding from this scoping review was that only five patient classification systems were validity and/or reliability tested (Table 2). DeGroot identified validity (the tool accurately and adequately predicts individual patient care requirements) and reliability (the tool consistently predicts patient care requirements) as being critical for a fully operational patient classification system (DeGroot, 1989). Only five patient classification systems seen here were validated (Brady et al., 2008, Çelebioğlu et al., 2007, Collister et al., 2014, Hawes et al., 2007, Poss et al., 2008). Of these, validity was tested in regard to content, predictive or concurrent validity, yet three instruments were tested for validity through face validity alone, which should not be used as evidence of an instrument's validity (Polit and Beck, 2004).

Only five instruments were reliability tested (Brady et al., 2008, Çelebioğlu et al., 2007, Collister et al., 2014, Hawes et al., 2007, Kottner et al., 2010), and the Revised Easley-Storfjell Patient Classification system (R-ESPCI) (Storfjell et al., 2017) was tested in earlier studies for validity and reliability (Anderson and Rokosky, 2001). In one integrative review, insufficient evidence on the reliability and validity testing of patient classification systems for inpatients in a medical/surgical setting was found (Fasoli and Haddock, 2010), which

corresponds with the findings seen here. Both the reliability and validity of the instruments seen here should be tested, taking DeGroot's six elements for a fully operational patient classification system into consideration (DeGroot, 1989).

Patient classification systems used for allocation of staff

Of the thirteen patient classification systems seen in this scoping review, some were well established and used worldwide, such as the RAI-HC (Hawes et al., 2007, Hirdes et al., 2008) or the RUG-III/HC (Björkgren et al., 2000, Carpenter et al., 1997). Still, the RUG-III/HC falls under the RAI-HC "umbrella" and is a documentation and assessment tool and not a daily patient classification system (InterRAI, 2018).

It is not only in hospital settings, where nurses report that a lack of time often results in undone nursing tasks and adverse events, that nurse staffing impacts the quality of care (Aiken et al., 2013, Cho et al., 2016). Nurse staffing also impacts the quality of care in nursing home settings, where better care is provided when nursing homes are high-staffed (Schnelle et al., 2004). In an home health care setting, nurses have described the implementation of clinical priorities, such as rationing care, when time is insufficient (Tønnessen et al., 2011). In home health care, it is essential that nurse staffing and nurse workload be measured and balanced in relation to nursing care.

In this scoping review, limited research was seen where a patient classification system was considered to be fully operational in home health care in accordance

with De Groot's six essential elements (DeGroot, 1989). This is of interest to administrators, researchers, clinicians, and policy makers who seek the further development of an operational patient classification system for use in a home health care setting.

5. Strengths and limitations

One strength is the rigorous use of a comprehensive scoping methodology in accordance with a framework by Arksey and O'Malley (Arksey and O'Malley, 2005). One limitation is that only English language articles were included, which can limit the information available. Also, policymakers, practitioners and service users were not contacted despite recommendations to do so (Levac et al., 2010), due to the extensive number of countries included. Lastly, the descriptions of the patient classification systems in the included papers were sometimes unclear or incomplete, so there is a risk that the interpretation fails to provide a correct picture of the different patient classification systems used in home health care.

6. Conclusions

Despite a growing aging population, associated with chronic conditions and an increased need for home health care, in the past decade little advancement has been made with regard to the investigation of validated systems whereby measurements of home health care nurses' workload are linked to staffing allocations. Differences existed regarding whether the patient classification

systems were validity or reliability tested. The detail to which the various patient classification systems were described varied, and there was a lack of information about the patient classification systems in relation to the allocation of staff. Limited research exists proving that patient classification systems are fully operational for use in home health care.

What is known about this topic

- The number of persons with chronic diseases and complex care needs in municipal healthcare services are increasing, and home health care is one of the fastest growing health care sectors today.
- The number of caseload, workload and patient classification systems used to measure nursing intensity and staff allocation in home health care is limited, and these instruments are relatively unexplored.

What the paper adds

- Different types of instruments were used to measure nursing intensity and nursing care requirements in regard to direct and indirect patient care, and patient care-related nursing activities and non-patient activities vary.
- Few of the patient classification systems seen here had been validity and/or reliability tested, and only one had been validity and reliability tested and evaluated.
- Research on patient classification systems that are considered fully operational in home health care is limited.

References

Aiken, L.H., Sloane, D.M., Bruyneel, L., Van den Heede, K., Sermeus, W., Consortium, R.C., 2013. Nurses' reports of working conditions and hospital quality of care in 12 countries in Europe. International Journal of Nursing Studies 50 (2), 143-153.

Albrecht, M.N., 1991. Home Health Care: Reliability and Validity Testing of a Patient-Classification Instrument. Public Health Nursing 8 (2), 124-131.

Alghamdi, M.G., 2016. Nursing workload: a concept analysis. Journal of nursing management 24 (4), 449-457.

Anderson, K.L., Rokosky, J.S., 2001. Evaluation of a home health patient classification instrument. Western Journal of Nursing Research 23 (1), 56-71.

Arksey, H., O'Malley, L., 2005. Scoping studies: towards a methodological framework. International journal of social research methodology 8 (1), 19-32.

Bing-Jonsson, P.C., Bjørk, I.T., Hofoss, D., Kirkevold, M., Foss, C., 2013. Instruments measuring nursing staff competence in community health care: A systematic literature review. Home Health Care Management & Practice 25 (6), 282-294.

Björkgren, M.A., Fries, B.E., Shugarman, L.R., 2000. A RUG-III case-mix system for home care. Canadian Journal on Aging/La Revue canadienne du vieillissement 19 (S2), 106-125.

Björkgren, M.A., Häkkinen, U., Finne-Soveri, U.H., Fries, B.E., 1999. Validity and reliability of Resource Utilization Groups (RUG-III) in Finnish long-term care facilities. Scandinavian Journal of Social Medicine 27 (3), 228-234.

Bloom, D.E., Chatterji, S., Kowal, P., Lloyd-Sherlock, P., McKee, M., Rechel, B., Rosenberg, L., Smith, J.P., 2015. Macroeconomic implications of population ageing and selected policy responses. The Lancet 385 (9968), 649-657.

Bôas, M.L.d.C.V., Shimizu, H.E., 2015. Time spent by the multidisciplinary team in home care: subsidy for the sizing of staff. Acta Paulista de Enfermagem 28 (1), 32-40.

Bowers, B., Durrant, K., 2014. Measuring safe staff levels in the community: the 'DominiC' workforce management tool. British Journal of Community Nursing 19 (2), 58-64.

Brady, A.M., Byrne, G., Horan, P., Griffiths, C., Macgregor, C., Begley, C., 2007. Measuring the workload of community nurses in Ireland: a review of workload measurement systems. Journal of nursing management 15 (5), 481-489.

Brady, A.M., Byrne, G., Horan, P., Macgregor, C., Begley, C., 2008. Reliability and validity of the CCNCS: a dependency workload measurement system. Journal of clinical nursing 17 (10), 1351-1360.

Brown, J., 2001. Nursing Home Resident Assessment Resource Utilization Groups.

Byrne, G., Brady, A., Horan, P., Macgregor, C., Begley, C., 2007. Assessment of dependency levels of older people in the community and measurement of nursing workload. Journal of Advanced Nursing 60 (1), 39-49.

Carpenter, G.I., Ikegami, N., Ljunggren, G., Carrillo, E., Fries, E., 1997. RUG-III and resource allocation: comparing the relationship of direct care time with patient characteristics in five countries. Age and Ageing 26 (suppl 2), 61-65.

Carpenter, I., Hirdes, J.P., 2013. Using interRAI assessment systems to measure and maintain quality of long-term care. In: A Good Life in Old Age? Monitoring and Improving Quality in Long-term Care. pp. 93-139.

Cawthorn, L., Rybak, L., 2008. Workload measurement in a community care program. Nursing Economic$ 26 (1), 45-48.

Cawthorn, L., Rybak, L., 2008. Workload measurement in a community care program. Nursing Economics 26 (1), 45-48.

Çelebioğlu, A., Özsoy, S.A., Peters, D., 2007. Study of the reliability and validity of the Community Health Intensity Rating Scale (CHIRS) in the Turkish community. Public Health Nursing 24 (3), 283-292.

Chapman, H., Kilner, M., Matthews, R., White, A., Thompson, A., Fowler-Davis, S., Farndon, L., 2017. Developing a caseload classification tool for community nursing. British Journal of Community Nursing 22 (4), 192-196.

Cho, E., Lee, N.-J., Kim, E.-Y., Kim, S., Lee, K., Park, K.-O., Sung, Y.H., 2016. Nurse staffing level and overtime associated with patient safety, quality of care, and care left undone in hospitals: a cross-sectional study. International journal of nursing studies 60, 263-271.

Collister, B., Slauenwhite, C.A., Fraser, K.D., Swanson, S., Fong, A., 2014. Measuring Home Care Caseloads: Development of the Caseload Intensity Tool. Home Health Care Management & Practice 26 (4), 239-249.

Craftman, Å.G., Grundberg, Å., Westerbotn, M., 2018. Experiences of home care assistants providing social care to older people: A context in transition. International journal of older people nursing, e12207.

De Vliegher, K., Declercq, A., Aertgeerts, B., Gosset, C., Heyden, I., Moons, P., 2014. The activity profile of home nurses: a systematic review. Home Health Care Management & Practice 26 (4), 257-265.

De Vliegher, K., Declercq, A., Aertgeerts, B., Moons, P., 2016. Health Care Assistants in Home Nursing: The Holy Grail or the Emperor's New Clothes? A Qualitative Study. Home Health Care Management & Practice 28 (1), 51-56.

DeGroot, H.A., 1989. Patient Classification System Evaluation, Part 1, Essential System Elements. Journal of Nursing Administration 19 (6), 30-35.

DeGroot, H.A., 1989. Patient Classification System Evaluation, Part Two, System Selection and Implementation. Journal of Nursing Administration 19 (7), 24-30.

Drennan, V.M., 2019. More care out of hospital? A qualitative exploration of the factors influencing the development of the district nursing workforce in England. Journal of health services research & policy 24 (1), 11-18.

Edwardson, S.R., Giovannetti, P.B., 1994. Nursing workload measurement systems. Annual review of nursing research 12 (1), 95-123.

Endsley, P., 2017. School nurse workload: A scoping review of acute care, community health, and mental health nursing workload literature. The Journal of School Nursing 33 (1), 43-52.

Eurostat, 2016. Healthcare resource statistics - beds.

Fagerström, L., Vainikainen, P., 2014. Nurses' experiences of nonpatient factors that affect nursing workload: a study of the PAONCIL instrument's nonpatient factors. Nursing research and practice 2014, 1-9.

Fasoli, D.R., Haddock, K.S., 2010. Results of an Integrative Review of Patient Classification Systems. Annual review of nursing research 28 (1), 295-316.

Giovannetti, P., 1984. Staffing methods—implications for quality. In, Measuring the quality of nursing care. Churchill Livingstone, Edinburg, pp. 123-150.

Giovannetti, P., 1979. Understanding patient classification systems. The Journal of nursing administration 9 (2), 4-9.

Giovannetti, P., Johnson, J.M., 1990. A new generation patient classification system. Journal of Nursing Administration 20 (5), 33-40.

Grafen, M., Mackenzie, F.C., 2015. Development and early application of the Scottish Community Nursing Workload Measurement Tool. British Journal of Community Nursing 20 (2), 89-92.

Grant, J., Lines, L., Darbyshire, P., Parry, Y., 2017. How do nurse practitioners work in primary health care settings? A scoping review. International journal of nursing studies 75, 51-57.

Grant, M.J., Booth, A., 2009. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Information & Libraries Journal 26 (2), 91-108.

Hafskjold, L., Eide, T., Holmström, I.K., Sundling, V., van Dulmen, S., Eide, H., 2016. Older persons' worries expressed during home care visits: Exploring the content of cues and concerns identified by the Verona coding definitions of emotional sequences. Patient education and counseling 99 (12), 1955-1963.

Hawes, C., Fries, B.E., James, M.L., Guihan, M., 2007. Prospects and pitfalls: use of the RAI-HC assessment by the Department of Veterans Affairs for home care clients. The Gerontologist 47 (3), 378-387.

Hays, B.J., Sather, L., Peters, D.A., 1999. Quantifying client need for care in the community: A strategy for managed care. Public Health Nursing 16 (4), 246-253.

Hirdes, J.P., Ljunggren, G., Morris, J.N., Frijters, D.H., Soveri, H.F., Gray, L., Björkgren, M., Gilgen, R., 2008. Reliability of the interRAI suite of assessment instruments: a 12-country study of an integrated health information system. BMC health services research 8 (1), 277.

Huckabay, L., Skonieczny, R., 1981. Patient classification systems: the problems faced. Nursing & health care: official publication of the National League for Nursing 2 (2), 89-102.

Hurst, K., 2006. Primary and community care workforce planning and development. Journal of advanced nursing 55 (6), 757-769.

InterRAI, 2018.

Kane, K., 2014. Capturing district nursing through a knowledge-based electronic caseload analysis tool (eCAT). British Journal of Community Nursing 19 (3), 116, 118-124.

Kline, T., 2005. Psychological testing: A practical approach to design and evaluation. Sage.

Kottner, J., Halfens, R., Dassen, T., 2010. Interrater reliability and agreement of the Care Dependency Scale in the home care setting in the Netherlands. Scandinavian Journal of Caring Sciences 24, 56-61.

Levac, D., Colquhoun, H., O'Brien, K.K., 2010. Scoping studies: advancing the methodology. Implementation Science 5 (1), 69.

Luo, H., Lin, M., Castle, N.G., 2013. The correlates of nursing staff turnover in home and hospice agencies: 2007 national home and hospice care survey. Research on Aging 35 (4), 375-392.

Martin-Misener, R., Kilpatrick, K., Donald, F., Bryant-Lukosius, D., Rayner, J., Valaitis, R., Carter, N., Miller, P.A., Landry, V., Harbman, P., 2016. Nurse practitioner caseload in primary health care: scoping review. International journal of nursing studies 62, 170-182.

Min, A., Scott, L.D., 2016. Evaluating nursing hours per patient day as a nurse staffing measure. Journal of nursing management 24 (4), 439-448.

Mitchell, G., Cousins, C., Burrows, R., Cousins, G., 2017. A review of safe-staffing models and their applicability to care homes. Journal of nursing management 25 (2), 157-162.

Morris, R., MacNeela, P., Scott, A., Treacy, P., Hyde, A., 2007. Reconsidering the conceptualization of nursing workload: literature review. Journal of advanced Nursing 57 (5), 463-471.

Nielsen, M.S., Jørgensen, F., 2016. Meaning creation and employee engagement in home health caregivers. Scandinavian journal of caring sciences 30 (1), 57-64.

Nordmark, S., Zingmark, K., Lindberg, I., 2015. Experiences and views of the discharge planning process among Swedish district nurses and home care

organizers: A qualitative study. Home Health Care Management & Practice 27 (3), 108-118.

Pallant, J., 2015. SPSS Survival Manual. Open University Press, Maidenhead.

Peters, D.A., 1988. Development of a community health intensity rating scale. Nursing Research 37 (4), 202-207.

Polit, D., Beck, C., 2004. Nursing research : principles and methods. Lippincott Williams & Wilkins, Philadelphia, Pa.

Poss, J.W., Hirdes, J.P., Fries, B.E., McKillop, I., Chase, M., 2008. Validation of Resource Utilization Groups version III for Home Care (RUG-III/HC): evidence from a Canadian home care jurisdiction. Medical care 46 (4), 380-387.

Rauhala, A., 2008. The Validity and Feasibility of Measurement Tools for Human Resources Management in Nursing - Case of the RAFAELA System. In: Department of Health Policy and Management. University of Kuopio, KUOPIO.

Roberson, C., 2016. Caseload management methods for use within district nursing teams: a literature review. British journal of community nursing 21 (5).

Saba, V., 2017. Clinical Care Classification (CCC) System: An Overview. In: Harris, M.D. (Ed.), Handbook of home health care administration. Jones & Bartlett Learning, Burlington, Massachusetts pp. 225-234.

Saba, V., 1992. Home health care classification. Caring: National Association for Home Care magazine 11 (5), 58-60.

Saba, V., 2002. Home health care classification system (HHCC): an overview. The Online Journal of Issues in Nursing: A Scholarly Journal of the American Nurses Association 7 (3).

Schnelle, J.F., Simmons, S.F., Harrington, C., Cadogan, M., Garcia, E., M Bates-Jensen, B., 2004. Relationship of nursing home staffing to quality of care. Health services research 39 (2), 225-250.

Skilbeck, J.K., Arthur, A., Seymour, J., 2018. Making sense of frailty: An ethnographic study of the experience of older people living with complex health problems. International journal of older people nursing 13 (1), e12172.

Squires, A., Jylhä, V., Jun, J., Ensio, A., Kinnunen, J., 2017. A scoping review of nursing workforce planning and forecasting research. Journal of nursing management 25 (8), 587-596.

Storfjell, J.L., Allen, C.E., Easley, C.E., 1997. Analysis and management of home health nursing caseloads and workloads. Journal of Nursing Administration 27 (9), 24-33.

Storfjell, J.L., Easley, C.A., Easley, C.E., 2017. Analysis and management of home health nursing caseloads and workloads: implication for productivity. In: Harris, M.D. (Ed.), Handbook of home health care

administration. Jones & Bartlett Learning, Burlington, Massachusetts pp. 427-442.

Streiner, D.L., Norman, G.R., Cairney, J., 2015. Health measurement scales: a practical guide to their development and use. Oxford University Press, USA.

Szczerbińska, K., Hirdes, J.P., Życzkowska, J., 2012. Good news and bad news: depressive symptoms decline and undertreatment increases with age in home care and institutional settings. The American Journal of Geriatric Psychiatry 20 (12), 1045-1056.

The Joanna Briggs Institute, 2015. Methodology for JBI Scoping Reviews.

The National Board of Health and Welfare, 2018. Care and caring about the elderly- progress report.

Tomstad, S., Dale, B., Sundsli, K., Sævareid, H.I., Söderhamn, U., 2017. Who often feels lonely? A cross-sectional study about loneliness and its related factors among older home-dwelling people. International journal of older people nursing 12 (4), e12162.

Tønnessen, S., Nortvedt, P., Førde, R., 2011. Rationing home-based nursing care: professional ethical implications. Nursing ethics 18 (3), 386-396.

Walker, J., Hendry, C., 2009. A review of literature on the measurement of patient dependency, workload and workload planning and allocation within the context of nursing in the community. Christchurch: The New Zealand Institute of Community Health Care.

WHO, 2016. Global strategy and action plan on ageing and health (2016-2020).

Fig. 1. Flow diagram of study selection

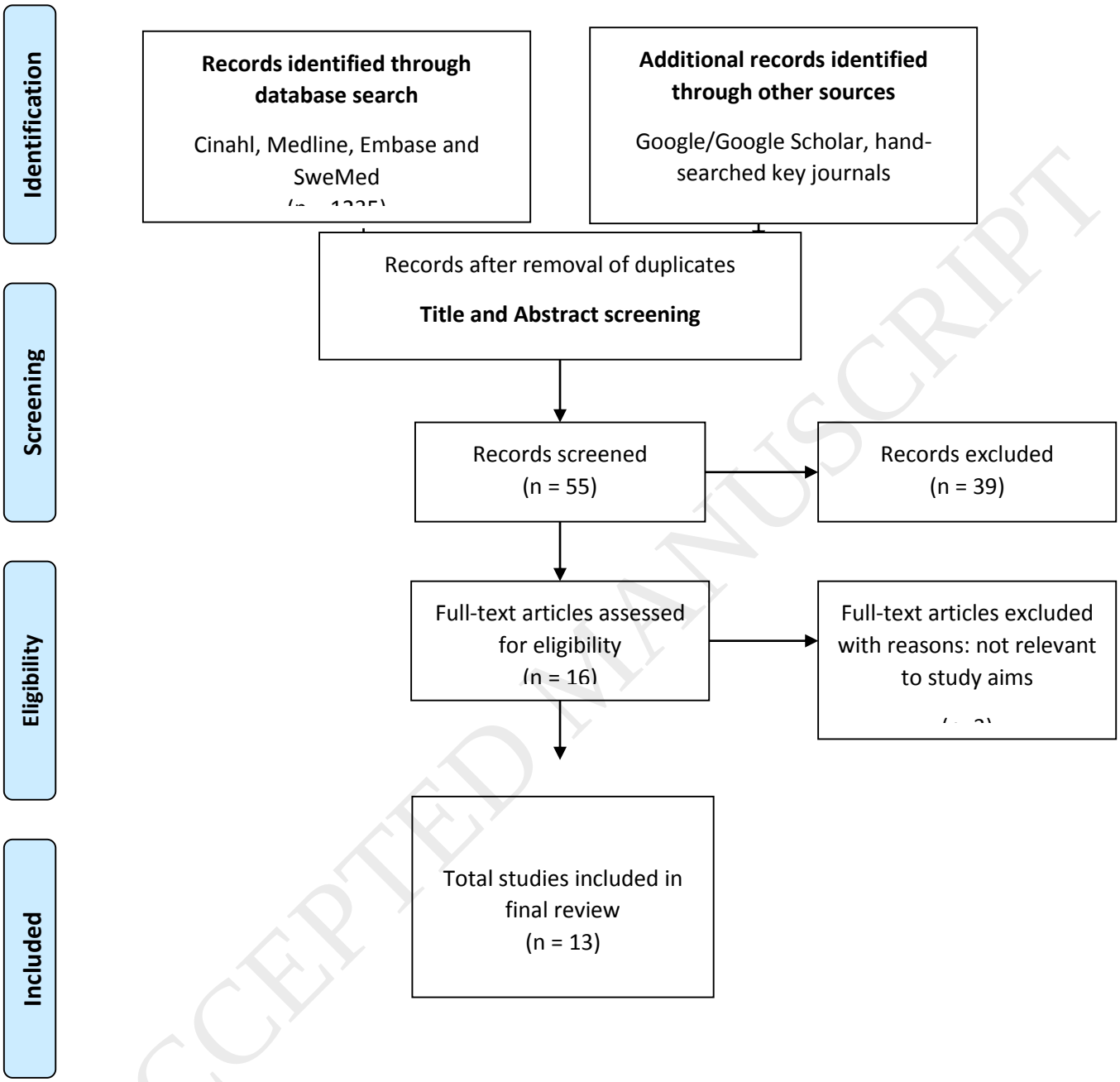

Table 1. A summary table demonstrating the key features of the 13 included studies.

Author(s)/date/countries	**Population/sample**	**Instrument**	**Validity/reliability/evaluated**
Bowers and Durrant (2014), UK	Pop. 283 650. Community nursing, 19 district nursing teams	Domiciliary care system in the Community (DominiC)	Evaluated
Brady et al. (2008), Ireland	3 community care areas 1) Pilot study with 9 PHNs, 2) 44 PHNs participated	Community Client Need Classification System (CCNCS)	Content validity, predictive validity, reliability: stability, internal consistency
Byrne et al. (2007), Ireland	3 community care areas 1) Pilot study with 9 PHNs, 2) 44 PHNs participated	Community Client Need Classification System (CCNCS)	Evaluated: dependency level and nursing time
Cawthorn and Rybak, (2008), Canada	Pop. 650 clients, RNs and LPNs	Client Audit Community Care Workload Assessment Tool	Evaluated

Çelebioğlu et al. (2007), Turkey	Pop. 372 families, one researcher	Community Health Intensity Rating Scale (CHIRS)	Face validity, predictive validity, reliability; internal consistency
Chapman et al. (2017), UK	Pop. >3000, 4 community nursing teams, 70 RNs, healthcare support workers, administrators	Caseload classification tool	Evaluated
Collister et al. (2014), Canada	5 geographical zones, mix of 97 clinical staff	Caseload Intensity tool (CIT)	Face validity, content validity, interrater reliability, internal consistency
Author/date/countries	Population/sample	Instrument	Validity/reliability/evaluated
Grafen and Mackenzie (2015), Scotland	Part of the national nursing and midwifery workload and workforce- planning program	The Scottish Community Nursing Workload Measurement Tool	Evaluated
Hawes et al. (2007), USA	A review of implementation challenges	The Resident Assessment	Validity, reliability tested in earlier studies; evaluated

		Instrument for home care (RAI-HC)	
Kane (2014), Ireland	Pop. 60 000-120 000 inhabitants per trust. 848 nursing staff	Electronic Caseload Analysis Tool (eCAT)	Evaluated
Kottner et al. (2010), Netherlands	Pop. 335 clients from 27 HHC agencies, year 2007 337 clients from 21 HHC agencies, year 2008	Care Dependency Scale (CDS)	Reliability tested
Poss et al. (2008), Canada	Pop. 21 578 clients	Resource Utilization Groups version III for Home Care (RUG-III/HC)	Validity tested
Storfjell et al. (2017), USA		The Easley-Storfjell Instrument for Caseload/Workload Analysis (CL/WLA)	Validity, reliability tested in earlier studies; evaluated

Table 2. A summary table showing author(s), year of publication, instrument/tool, categories/criteria for assessment, whether validity or reliability tested (*2 reports/papers were produced by different researchers from the same study)

Author(s), year of publication, instrument/tool	**Categories/criteria for assessment**	**Validity tested**	**Reliability tested**
Bowers and Durrant (2014) DominiC	Key functions: Centrally held patient details Records patients' needs and predicts visit length Gives patients a choice of timed visits Highlights risk when visiting Communicates reminders for future visits Allocates work geographically Reduces risk of missed/duplicated visits The same nurses for continuity of care Medical device management Records patient quality data Records staff roster Identifies future service shortfalls Increases transparency between teams Produces reports on services delivery and efficiency *More specific information about the tool was not given.*	No information was given	No information was given
***Brady et al. (2008)**	The seven assessment criteria included: 1) Nursing Assessment, 2) Physical Care	Content Validity Index	Stability through pre- and post-testing

***Byrne et al. (2007) CCNCS**	Requirements, 3) Teaching Needs and Health Promotion, 4) Carer and Family Support, 5) Case/Care Management, 6) Psycho-social Support, 7) Environmental Factors. Nurses scored patients' needs levels from 1 (low need) to 5 (high need). Total needs were thereafter rated on a scale: 1 (0-7), 2 (8-14), 3 (15-21), 4, (22-28) and 5 (29-40). The travel time per visit (>20 minutes) was also recorded.	(CVI) overall total CVI at 0.99. Predictive Validity: relationship between needs level and nursing time calculated using the Kruskal-Wallis test, Jonckheere-Terpstra test and Mann-Whitney U-test.	Child case (k 0.593) Older person case (k 0.704) Internal consistency by percentage agreement: Child case (91.1% recorded patient's level 4 and 5) Older person case (91-92% recorded patient's level 2 and 3)
Cawthorn and Rybak (2007). The Client Audit Community Care Workload Assessment tool	Identified nursing needs, frequency and length of home visit, how far the patient resides from the home office, coordination time required, client stability, client and/or caregiver's coping skill. The weighting score was: 8-13 = low intervention, 14-21 = moderate intervention and >22 = high intervention. In addition, the Regina Risk Inventory Tool and the Mini-Mental Status Evaluation were assessed, *but no further information was given.*	No information was given	No information was given

Çelebioğlu et al. (2007) **CHIRS**	Included 4 domains (with 15 parameters, 91 items and 975 sub-items): 1. Environmental 2. Psychosocial 3. Physical 4. Health behavior. Scores ranged from 0 (no need for care) to 4 (extreme need for care) and were used as indicators of the intensity of health care needs.	Face validity by experts: considered suitable, but too long. Predictive validity: Correlation of total scale score (TSS) and total number of household members (p=.001) No correlation between TSS and total number of visits (p=.30) Correlation between TSS and total number of health institutions visited (p=.001)	Internal consistency for the overall scale: Cronbach's alpha (0.525) Guttman split-half coefficient (0.629)
Chapman et al. (2007) **Caseload Classification Tool for Community Nursing**	The tool focused on assessment, social circumstances, liaison, dignity and respect, intervention, band of staff providing care, expected visit time and length of stay on the caseload. Included a 12 package of care with clinically evidenced-based care plans. *More specific information about the tool was not given.* Level of care from routine, additional to significant.	No information was given	No information was given

Collister et al. (2014) CIT	Included six categories: A: Clients requiring minimal services B: Clinical Condition: Stability and Predictability C: Clinical condition: Complexity Factors D: Care Response: Treatment and Therapy Process E: Care Response: Client/family Care Plan Process F: Care Response: System Care Plan Process The CIT was divided into three levels: level 1 = 1 score, level 2 = 2 scores and level 3 = 3 scores (1 score minimum and 15 maximum), resulting in a raw score that was subsequently converted into a Client Intensity Scale.	Content validity by literature, key factors and experts. Face validity by collecting qualitative data from the staff. Concurrent validity by analyzing the CIT score with staff's monthly recorded activity Stability: Wilcoxon test for paired samples showed that when time between repeated tests was brief, the score did not change significantly.	Internal reliability: Cronbach's alpha (0.79 / 0.95) Interrater reliability (71%)
Grafen and Mackenzie (2015) The Scottish Community Nursing Workload Measurement Tool	Included 6 activity categories: 1) Face-to face contact (everything occurring when the patient is present) 2) Non-face-to-face context (everything occurring when the patient is not present) 3) Home visits and planned sessions (routine, group and mass clinics/sessions) 4) Associated workload (management, administration, meetings and professional development activity)	No information was given	No information was given

	5) Travel (including walking, waiting and parking time) 6) Exception reporting (extraordinary events, e.g., adverse weather, car breakdown) Levels of intervention, ranging from straightforward (level 1) to complex (level 2).		
Hawes et al. (2007) **RAI-HC**	Consisted of two elements, MDS-HC and CAPs. MDS-HC: collect standardized patient information and assess 19 key domains related to function, health, social support and service use. The CAPs consisted of 30 problem-focused protocol areas that encompassed common risks for home care clients, organized under the following areas: 1. Functional performance 2. Sensory performance 3. Mental health 4. Bladder management 5. Health problem/syndromes 6. Service oversight	Tested in earlier studies (Morris et al. 1997)	Tested in earlier studies (Kwan, Chi, Lam, Lam and Shou, 2000, Landi et al. 2000)

Kane (2014) **eCAT**	Included eight categories: 1. Demography (total population) 2. Caseload size (number of patients) 3. Visiting patterns (8 variables) 4. Reason for visit (24 variables) 5. Dependency measure (ADL dependency and team dependency) 6. Caseload throughput 7. Location of care (6 variables) 8. Reviews (8 variables) *More specific information about the categories or the variables was not given*	No information was given	No information was given
Kottner et al. (2010) **CDS**	Included 15 items: eating and drinking, continence, body posture, mobility, day/night pattern, getting (un) dressed, body temperature, hygiene, avoidance of danger, communication, contact with others, sense of rules/values, daily activities, recreational activities, learning ability. Scores ranging from 1 (completely dependent) to 5 (completely independent). The scores from all 15 categories were summed up, yielding sum scores that ranged from 15–75; *low sum scores*		Interrater reliability (k 0.61-0.79) Agreement (0.67%-0.88 %)

	indicated a *high care of dependency* and *high sum scores* indicated *independency*.		
Poss et al. (2008) RUG-III/HC	Included seven hierarchical levels: 1. Special rehabilitation 2. Extensive services 3. Special care 4. Clinically complex 5 Impaired cognition 6. Behavior problems 7. Reduced physical functions ADL (eating, toilet use, transfer, bed, mobility) score 4-15 IADL (meal preparation, medication management, phone use) score 0-3	Validity: by cost episodes over a 13-week period. Individual client level billing records matched with assessment information from the RUG-III/HC.	
Storfjell et al. (2017) Easley–Storfjell caseload/workload analysis instruments	Assessment of six variables: 1. Clinical judgment required (assessment needs) 2. Teaching needs 3. Physical care needs (technical procedures) 4. Psychosocial support needs 5. Coordination and care management needs 6. Number and severity of problems.	Tested in earlier studies (Anderson and Rokosky, 2001)	Tested in earlier studies (Anderson and Rokosky, 2001)

	The rating was from minimal complexity to very great complexity.		